Introduction to Quantum Probability

J.M. Swart

April 5, 2017

Contents

Preface

It is a fact of everyday life that our knowledge about the world around us is always incomplete and imperfect. We may feel pretty sure that we locked the front door when we left our house this morning, but less sure about how much milk there is left in our fridge. A mathematical theory that deals with such incomplete knowledge is probability theory. Since the early 1930-ies, in particular since the monograph of Kolmogorov [Kol33], probability theory is based on measure theory. Incomplete knowledge about a physical system is described by a probability space $(\Omega, \mathcal{F}, \mu)$, where Ω is a set, called the state space, $\mathcal{F}$ is a σ-algebra on Ω, and μ is a probability measure on $\mathcal{F}$.

At the same time when Kolmogorov's monograph laid the axiomatic basis for probability theory as it had been around since the times of Fermat, physicists were discovering a whole new type of probability theory. With the arrival of the Copenhagen interpretation of quantum mechanics, it became clear that quantum mechanics, at its heart, is a theory about probabilities, and that these probabilities do not fit into Kolmogorov's scheme. In order to describe incomplete knowledge about a quantum physical system, instead of a probability space $(\Omega, \mathcal{F}, \mu)$, physicists use a pair $(\mathcal{A}, \rho)$ where $\mathcal{A}$ is a C^*-algebra and ρ is a positive linear form on $\mathcal{A}$. If $\mathcal{A}$ is noncommutative, then these 'quantum probability spaces' do not correspond to anything classical, and put a severe strain on our imagination.

The aim of the present course is to make acquaintance with this quantum probability formalism, its interpretation, its difficulties, and its applications. Prerequisites for this course are elementary knowledge of complex numbers and linear algebra. It is helpful if one has some familiarity with the basic concepts of probability theory such as independence, conditional probabilities, expectations, and so on.

Sections marked with * can be skipped at a first reading.

Chapter 1

Linear spaces

1.1 Linear spaces

Let $\mathbb{K}$ denote either $\mathbb{R}$ or $\mathbb{C}$.[1] By definition, a *linear space* (or *vector space*) over $\mathbb{K}$ is a set $\mathcal{V}$, with a special element $0 \in \mathcal{V}$ called *origin*, on which an addition $(\phi, \psi) \mapsto \phi + \psi$ and multiplication with scalars $(a, \phi) \mapsto a\phi$ are defined, such that

$$\begin{aligned}
&\text{(i)} && (\phi + \psi) + \chi = \phi + (\psi + \chi),\\
&\text{(ii)} && \phi + \psi = \psi + \phi,\\
&\text{(iii)} && \phi + 0 = \phi,\\
&\text{(iv)} && (ab)\phi = a(b\phi),\\
&\text{(v)} && 0\phi = 0,\\
&\text{(vi)} && 1\phi = \phi,\\
&\text{(vii)} && a(\phi + \psi) = a\phi + a\psi,\\
&\text{(viii)} && (a + b)\phi = a\phi + b\phi
\end{aligned}$$

for all $\phi, \psi, \chi \in \mathcal{V}$ and $a, b \in \mathbb{K}$.

A subset of $\mathcal{V}$ that is closed under addition and multiplication with scalars is called a linear subspace. By definition, the *span* of a subset $\mathcal{W} \subset \mathcal{V}$ is the linear subspace defined as[2]

$$\operatorname{span}(\mathcal{W}) := \{a_1\phi(1) + \cdots + a_n\phi(n) : \phi(1), \ldots, \phi(n) \in \mathcal{W}\}$$

We say that $\mathcal{W}$ *spans* the linear subspace $\operatorname{span}(\mathcal{W})$. We say that a linear space $\mathcal{V}$ is *finite dimensional* if there exists a finite set $\mathcal{W}$ such that $\mathcal{V} = \operatorname{span}(\mathcal{W})$.

[1] In fact, more generaly, all of Section 1.1 is true when $\mathbb{K}$ is division ring, but we will not need this generality.

[2] In these lecture notes, the symbol $\subset$ means: subset of (and possibly equal to). Thus, in particular, $A \subset A$.

A finite collection $\{\phi(1), \ldots, \phi(n)\}$ of elements of a linear space $\mathcal{V}$ is called *linearly independent* if the equation

$$a_1\phi(1) + \cdots + a_n\phi(n) = 0$$

has no other solutions than $a_1 = a_2 = \cdots = a_n = 0$. If moreover $\{\phi(1), \ldots, \phi(n)\}$ spans $\mathcal{V}$ then we call $\{\phi(1), \ldots, \phi(n)\}$ a *basis* for $\mathcal{V}$. Let $\{e(1), \ldots, e(n)\}$ be a basis for $\mathcal{V}$. Then for every $\phi \in \mathcal{V}$ there exist unique $\phi_1, \ldots, \phi_n \in \mathbb{K}$ such that

$$\phi = \phi_1 e(1) + \cdots + \phi_n e(n).$$

Thus, given a basis we can set up a linear isomorphism between our abstract vector space $\mathcal{V}$ and the concrete linear space $\mathbb{K}^n := \{(\phi_1, \ldots, \phi_n) : \phi_i \in \mathbb{K}\ \forall i = 1, \ldots, n\}$. We call $(\phi_1, \ldots, \phi_n)$ the *coordinates* of ϕ with respect to the basis $\{e(1), \ldots, e(n)\}$. Note that if we want to label a collection of vectors in $\mathcal{V}$, such as $\{\phi(1), \ldots, \phi(n)\}$, then we put the labels between brackets to distinguish such notation from the coordinates of a vector with respect to a given basis.

It can be shown that every finite dimensional linear space has a basis. (Note that this is not completely straightforward from our definitions!) If $\mathcal{V}$ is finite dimensional, then one can check that all bases of $\mathcal{V}$ have the same number of elements n. This number is called the *dimension* $\dim(\mathcal{V})$ of $\mathcal{V}$. *From now on, all linear spaces are finite dimensional, unless stated otherwise.*

Let $\mathcal{V}, \mathcal{W}$ be linear spaces. By definition, a map $A : \mathcal{V} \to \mathcal{W}$ is called *linear* if

$$A(a\phi + b\psi) = aA\phi + bA\psi \qquad (a, b \in \mathbb{K},\ \phi, \psi \in \mathcal{V}).$$

We denote the space of all linear maps from $\mathcal{V}$ into $\mathcal{W}$ by $\mathcal{L}(\mathcal{V}, \mathcal{W})$. In an obvious way $\mathcal{L}(\mathcal{V}, \mathcal{W})$ is itself a linear space. If $A \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, $\{e(1), \ldots, e(n)\}$ is a basis for $\mathcal{V}$, and $\{f(1), \ldots, f(m)\}$ is a basis for $\mathcal{W}$, then

$$(A\phi)_i = \sum_{j=1}^{n} A_{ij}\phi_j \qquad (i = 1, \ldots, m),$$

where ϕ_j $(j = 1, \ldots, n)$ and $(A\phi)_i$ $(i = 1, \ldots, m)$ are the coordinates of ϕ and $A\phi$ with respect to $\{e(1), \ldots, e(n)\}$ and $\{f(1), \ldots, f(m)\}$, respectively, and

$$\begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix}$$

is the *matrix* of A with respect to the bases $\{e(1), \ldots, e(n)\}$ and $\{f(1), \ldots, f(m)\}$. The numbers $A_{ij} \in \mathbb{K}$ are called the *entries* of A.

Exercise 1.1.1 If $A \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ and $B \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, then show that

$$(AB)_{ij} = \sum_k A_{ik} B_{kj}.$$

The *kernel* and *range* of a linear operator $A \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ are defined by

$$\begin{aligned} \mathrm{Ker}(A) &:= \{\phi \in \mathcal{V} : A\phi = 0\}, \\ \mathrm{Ran}(A) &:= \{A\phi : \phi \in \mathcal{V}\}. \end{aligned}$$

One has

$$\dim(\mathrm{Ker}(A)) + \dim(\mathrm{Ran}(A)) = \dim(\mathcal{V}).$$

If a linear map $A : \mathcal{V} \to \mathcal{W}$ is a bijection then one can check that its inverse A^{-1} is also linear. In this case we call A *invertible*. A linear map $A : \mathcal{V} \to \mathcal{W}$ is invertible if and only if $\mathrm{Ker}(l) = \{0\}$ and $\mathrm{Ran}(l) = \mathcal{W}$. This is equivalent to $\mathrm{Ker}(l) = \{0\}$ and $\dim(\mathcal{V}) = \dim(\mathcal{W})$.

For any linear space $\mathcal{V}$, we write $\mathcal{L}(\mathcal{V}) := \mathcal{L}(\mathcal{V}, \mathcal{V})$ for the space of all linear maps $A : \mathcal{V} \to \mathcal{V}$. We also call such linear maps *linear operators*. We define the *commutator* of two operators $A, B \in \mathcal{L}(\mathcal{V})$ by

$$[A, B] := AB - BA,$$

and we say that A and B commute if $[A, B] = 0$, i.e., if $AB = BA$.

By definition, the *trace* of a linear operator in $\mathcal{L}(\mathcal{V})$ is given by

$$\mathrm{tr}(A) := \sum_{i=1}^{n} A_{ii}.$$

Here A_{ij} denotes the matrix of A with respect to any basis $\{e(1), \ldots, e(n)\}$ of $\mathcal{V}$; it can be shown that the definition of the trace is independent of the choice of the basis. The trace is linear and satisfies

$$\mathrm{tr}(AB) = \mathrm{tr}(BA) \qquad (A \in \mathcal{L}(\mathcal{V}, \mathcal{W}),\ B \in \mathcal{L}(\mathcal{W}, \mathcal{V})).$$

By definition, an *eigenvector* of a linear operator $A \in \mathcal{L}(\mathcal{V})$ is a vector $\psi \in \mathcal{V}$, $\psi \neq 0$, such that

$$A\psi = \lambda\psi$$

for some $\lambda \in \mathbb{K}$. The constant λ is called the *eigenvalue* corresponding to the eigenvector ψ. By definition,

$$\sigma(A) := \{\lambda \in \mathbb{K} : \lambda \text{ is an eigenvalue of } A\}$$

is called the *spectrum* of A.

Exercise 1.1.2 Show that $\sigma(A) = \{\lambda \in \mathbb{K} : (\lambda - A) \text{ is not invertible}\}$.

The following proposition holds only for linear spaces over the complex numbers.

Proposition 1.1.3 (Nonempty spectrum) *Let $\mathcal{V} \neq \{0\}$ be a linear space over $\mathbb{C}$ and let $A \in \mathcal{L}(\mathcal{V})$. Then $\sigma(A)$ is not empty.*

Proof (sketch) The eigenvalues of A can be found by solving the equation $\det(A-\lambda) = 0$. Here $\det(A - \lambda)$ is a polynomial of order $\dim(\mathcal{V})$ which, as we know, has $\dim(\mathcal{V})$ complex roots. ∎

A linear operator is called *diagonalizable* if there exists a basis $\{e(1), \ldots, e(n)\}$ for $\mathcal{V}$ consisting of eigenvectors of A. With respect to such a basis, the matrix of A has the *diagonal form* $A_{ij} = \lambda_i \delta_{ij}$, where λ_i is the eigenvalue corresponding to the eigenvector $e(i)$, and

$$\delta_{ij} := \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

1.2 Inner product spaces

Let $\mathcal{H}$ be a linear space over $\mathbb{K} = \mathbb{R}$ of $\mathbb{C}$. By definition, an *inner product* on $\mathcal{H}$ is a map $(\phi, \psi) \mapsto \langle\phi|\psi\rangle$ from $\mathcal{H} \times \mathcal{H}$ into $\mathbb{K}$ such that

$$\begin{array}{lll} \text{(i)} & \langle\phi|a\psi+b\chi\rangle = a\langle\phi|\psi\rangle + b\langle\phi|\chi\rangle & (\phi,\psi,\chi \in \mathcal{H},\ a,b \in \mathbb{C}), \\ \text{(ii)} & \langle\phi|\psi\rangle = \langle\psi|\phi\rangle^* & (\phi,\psi \in \mathcal{H}), \\ \text{(iii)} & \langle\phi|\phi\rangle \geq 0 & (\phi \in \mathcal{H}), \\ \text{(iv)} & \langle\phi|\phi\rangle = 0 \ \Rightarrow \ \phi = 0. & \end{array}$$

Here a^* denotes the complex conjugate of a complex number a. A linear space that is equipped with an inner product is called an *inner product space*. By definition,

$$\|\psi\| := \sqrt{\langle\psi|\psi\rangle} \qquad (\psi \in \mathcal{H})$$

is the *norm* associated with the inner product $\langle\cdot|\cdot\rangle$. Two vectors ϕ, ψ are called *orthogonal* if $\langle\phi|\psi\rangle = 0$. A basis $\{e(1), \ldots, e(n)\}$ of $\mathcal{H}$ is called *orthogonal* if $\langle e(i)|e(j)\rangle = 0$ for all $i \neq j$. It is called *orthonormal* if in addition $\langle e(i)|e(i)\rangle = 1$ for all i. Every inner product space has an orthonormal basis.

Dirac's [Dir58] bracket notation is a clever way to 'decompose' the inner product $\langle\psi|\phi\rangle$ on an inner product space $\mathcal{H}$ into two parts, $\langle\psi|$ and $|\phi\rangle$, which Dirac called a bra and a ket, so that together they form a bra(c)ket $\langle\phi|\psi\rangle$. For any $\psi \in \mathcal{H}$, define operators $\langle\psi| \in \mathcal{L}(\mathcal{H}, \mathbb{K})$ and $|\psi\rangle \in \mathcal{L}(\mathbb{K}, \mathcal{H}) \cong \mathcal{H}$ by

$$\begin{array}{ll} \langle\psi|\phi := \langle\psi|\phi\rangle & (\phi \in \mathcal{H}), \\ |\psi\rangle\lambda := \lambda\psi & (\lambda \in \mathbb{K}). \end{array}$$

Then for any $\phi, \psi \in \mathcal{H}$, the composition $\langle\phi|\,|\psi\rangle$ is an operator in $\mathcal{L}(\mathbb{K}, \mathbb{K}) \cong \mathbb{K}$ that can be associated with the number $\langle\phi|\psi\rangle \in \mathbb{K}$. Here we write $\cong$ to indicate that two linear spaces are in a natural way isomorphic.

If $\{e(1), \ldots, e(n)\}$ is an orthonormal basis of $\mathcal{H}$ and $\phi \in \mathcal{H}$, then the coordinates of ϕ with respect to this basis are given by

$$\phi_i = \langle e(i)|\phi\rangle.$$

If $\mathcal{H}_1, \mathcal{H}_2$ are inner product spaces with orthonormal bases $\{e(1), \ldots, e(n)\}$ and $\{f(1), \ldots, f(m)\}$, respectively, and $A \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, then the matrix of A with respect to these bases is given by

$$A_{ij} = \langle f(i)|A|e(j)\rangle.$$

One has

$$A = \sum_{ij} A_{ij}|f(i)\rangle\langle e(j)|.$$

Note that $\langle e(j)| \in \mathcal{L}(\mathcal{H}_1, \mathbb{K})$ and $|f(i)\rangle \in \mathcal{L}(\mathbb{K}, \mathcal{H}_2)$, so the composition $|f(i)\rangle\langle e(j)|$ is an operator in $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. In particular, for the identity map $1 \in \mathcal{L}(\mathcal{H}) = \mathcal{L}(\mathcal{H}, \mathcal{H})$ one has the useful relation

$$1 = \sum_i |e(i)\rangle\langle e(i)|.$$

If $\mathcal{H}_1, \mathcal{H}_2$ are inner product spaces and $A \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, then there exists a unique *adjoint* $A^* \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$ of A, such that

$$\langle\phi|A\psi\rangle_2 = \langle A^*\phi|\psi\rangle_1 \qquad (\phi \in \mathcal{H}_2,\ \psi \in \mathcal{H}_1),$$

where $\langle\cdot|\cdot\rangle_1$ denotes the inner product in $\mathcal{H}_1$ and $\langle\cdot|\cdot\rangle_2$ denotes the inner product in $\mathcal{H}_2$. It is easy to see that

$$(aA + bB)^* = a^*A^* + b^*B^* \qquad (A, B \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2),\ a, b \in \mathbb{K}),$$

i.e., $A \mapsto A^*$ is *colinear*, and

$$(A^*)^* = A.$$

If $A \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ and $B \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_3)$ then one has

$$(AB)^* = B^*A^*.$$

Exercise 1.2.1 Let $A \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ and let $\{e(1), \ldots, e(n)\}$ and $\{f(1), \ldots, f(n)\}$ be orthonormal bases for $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. Show that the matrix of A^* is given by

$$A^*_{ij} = (A_{ji})^*.$$

Exercise 1.2.2 We can view $\mathbb{K}$ in a natural way as a (one-dimensional) inner product space with inner product $\langle a|b\rangle := a^*b$. Show that for any inner product space $\mathcal{H}$ and $\phi \in \mathcal{H}$,

$$|\phi\rangle^* = \langle\phi|.$$

Exercise 1.2.3 Let $\mathcal{H}_1, \mathcal{H}_2$ be inner product spaces and let $A \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. Show that $\langle\phi|A^* = \langle A\phi|$ for all $\phi \in \mathcal{H}_1$.

Exercise 1.2.4 Let $\mathcal{H}_1, \mathcal{H}_2$ be inner product spaces and let $A, B \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. Show that

$$\mathrm{tr}(A^*B) = \sum_{ij} (A_{ji})^* B_{ji}.$$

Show that $\langle A|B\rangle := \mathrm{tr}(A^*B)$ defines an inner product on $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$.

An operator $A \in \mathcal{L}(\mathcal{H})$ is called *normal* if it commutes with its adjoint, i.e.,

$$AA^* = A^*A.$$

The following theorem holds only for inner product spaces over $\mathbb{C}$.

Theorem 1.2.5 (Diagonalization of normal operators) *Assume that $\mathcal{H}$ is an inner product space over $\mathbb{C}$. Then an operator $A \in \mathcal{L}(\mathcal{H})$ is normal if and only if there exists an orthonormal basis $\{e(1), \ldots, e(n)\}$ and complex numbers $\lambda_1, \ldots, \lambda_n$ such that*

$$A = \sum_{i=1}^{n} \lambda_i |e(i)\rangle\langle e(i)|. \tag{1.1}$$

Note that (1.1) says that the matrix of A with respect to $\{e(1), \ldots, e(n)\}$ is diagonal, i.e., $A_{ij} = \lambda_i \delta_{ij}$. The constants $\lambda_1, \ldots, \lambda_n$ (some of which may be the same) are the eigenvalues of A.

Proof of Theorem 1.2.5 (sketch) If A is normal, then we claim that

$$\|A^*\phi\| = \|A\phi\| \qquad (\phi \in \mathcal{H}).$$

Indeed, this follows by writing $\|A^*\phi\|^2 = \langle A^*\phi|A^*\phi\rangle = \langle\phi|AA^*\phi\rangle = \langle\phi|A^*A\phi\rangle = \langle A\phi|A\phi\rangle$. Next, we claim that

$$A\phi = \lambda\phi \quad \text{implies} \quad A^*\phi = \lambda^*\phi.$$

Indeed, $A\phi = \lambda\phi$ implies $(A-\lambda)\phi = 0$. Since $A-\lambda$ is normal, it follows that $\|(A-\lambda)^*\phi\| = \|(A-\lambda)\phi\| = 0$ and hence $A^*\phi = \lambda^*\phi$.
By Proposition 1.1.3, each $A \in \mathcal{L}(\mathcal{H})$ has at least one eigenvector ϕ, say $A\phi = \lambda\phi$. We claim that A maps the space $\{\phi\}^\perp := \{\psi \in \mathcal{L}(\mathcal{H}) : \langle\psi|\phi\rangle = 0$ into itself. Indeed, $\langle\psi|\phi\rangle = 0$ implies $\langle A\psi|\phi\rangle = \langle\psi|A^*\phi\rangle = \lambda^*\langle\psi|\phi\rangle = 0$. Applying Proposition 1.1.3 to the restriction of A to the smaller space $\{\phi\}^\perp$, we see that A must have another eigenfuction in $\{\phi\}^\perp$. Repeating this process, we arrive at an orthogonal basis of eigenvectors. Normalizing yields an orthonormal basis. ∎

If $\mathcal{H}_1, \mathcal{H}_2$ are inner product spaces and $U \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, then we say that U is *unitary* if

$$\langle U\phi|U\psi\rangle_2 = \langle\phi|\psi\rangle_1 \qquad (\phi, \psi \in \mathcal{H}_1),$$

i.e., U preserves the inner product.

Exercise 1.2.6 Let $\mathcal{H}_1, \mathcal{H}_2$ be inner product spaces and $U \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. Assume that $\mathcal{H}_1$ and $\mathcal{H}_2$ have the same dimension. Show that an operator $U \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ is unitary if and only if U is invertible and $U^{-1} = U^*$. Hint: consider the image under U of an orthonormal basis of $\mathcal{H}_1$.

Note that since any invertible operator in $\mathcal{L}(\mathcal{H}) = \mathcal{L}(\mathcal{H}, \mathcal{H})$ commutes with its inverse, Exercise 1.2.6 shows that unitary operators in $\mathcal{L}(\mathcal{H})$ are normal.

Exercise 1.2.7 Let $\mathcal{H}$ be an inner product space over $\mathbb{C}$. Show that an operator $U \in \mathcal{L}(\mathcal{H})$ is unitary if and only if U is of the form

$$U = \sum_{i=1}^{n} \lambda_i |e(i)\rangle\langle e(i)|$$

where $\{e(1), \ldots, e(n)\}$ is an orthonormal basis of $\mathcal{H}$ and $\lambda_1, \ldots, \lambda_n$ are complex numbers such that $|\lambda_i| = 1$ for $i = 1, \ldots, n$.

An operator $A \in \mathcal{L}(\mathcal{H})$ is called *hermitian* or *self-adjoint* if $A = A^*$. In coordinates with respect to an orthonormal basis, this means that $A_{ij} = (A_{ji})^*$. Obviously, hermitian operators are normal.

Exercise 1.2.8 Let $\mathcal{H}$ be an inner product space over $\mathbb{C}$ with orthonormal basis $\{e(1), \ldots, e(n)\}$, and let

$$A = \sum_{i=1}^{n} \lambda_i |e(i)\rangle\langle e(i)|$$

be a normal operator on $\mathcal{H}$. Show that A is hermitian if and only if the eigenvalues λ_i are real.

Remark For hermitian operators, Theorem 1.2.5 also holds for inner product spaces over $\mathbb{R}$. To see this, let $\mathcal{H}$ be an inner product space over $\mathbb{R}$ and let $A \in \mathcal{L}(\mathcal{H})$ be hermitian. We need to show that A has at least one eigenvector. By the same arguments used in the proof of Theorem 1.2.5, one can then construct an orthonormal basis of such eigenvectors. To see that A has at least one eigenvector, let $\{e(1), \ldots, e(n)\}$ be an orthonormal basis for $\mathcal{H}$. The matrix of A with respect to this basis satisfies $A_{ij} = (A_{ji})^*$, so we may view A as a hermitian operator on the space $\mathbb{C}^n := \{(\phi_1, \ldots, \phi_n) : \phi_i \in \mathbb{C}\ \forall i\}$. By Proposition 1.1.3, A has at least one eigenvector $\phi = (\phi_1, \ldots, \phi_n) \in \mathbb{C}^n$. We are done if we can show that A has an eigenvector in $\mathbb{R}^n$. To this aim, we observe that if $\phi \in \mathbb{C}^n$ is an eigenvector with eigenvalue λ, then its complex conjugate $\phi^* := (\phi_1^*, \ldots, \phi_n^*)$ is also an eigenvector, with the same eigenvalue. Indeed, $(A\phi^*)_i = \sum_j A_{ij}\phi_j^* = (\sum_j A_{ij}\phi_j)^* = (\lambda\phi_i)^* = \lambda\phi_i^*$, where we have used that λ is real. It follows that $\mathrm{Re}(\phi) := (\phi^* + \phi)/2$ and $\mathrm{Im}(\phi) := (i\phi^* - i\phi)/2$ also satisfy $A\mathrm{Re}(\phi) = \lambda\mathrm{Re}(\phi)$ and $A\mathrm{Im}(\phi) = \lambda\mathrm{Im}(\phi)$. Since at least one of these vectors must be nonzero, we have found a real eigenvector for A.[3]

An operator $A \in \mathcal{L}(\mathcal{H})$ is called *positive* if and only if A is hermitian and all its eigenvalues are nonnegative. We define a partial order on the space of all hermitian operators by

$$A \leq B \ \Leftrightarrow\ B - A \text{ is positive.}$$

Let $\mathcal{H}$ be an inner product space and let $\mathcal{F} \subset \mathcal{H}$ be a linear subspace of $\mathcal{H}$. Let

$$\mathcal{F}^\perp := \{\phi \in \mathcal{H} : \langle\phi|\psi\rangle = 0\ \forall \psi \in \mathcal{F}\}.$$

denote the *orthogonal complement* of $\mathcal{F}$. Then each vector $\phi \in \mathcal{H}$ can in a unique way be written as

$$\phi = \phi' + \phi'' \qquad (\phi' \in \mathcal{F},\ \phi'' \in \mathcal{F}^\perp).$$

We call ϕ' the *orthogonal projection* of ϕ on the subspace $\mathcal{F}$, and write

$$\phi' =: P_{\mathcal{F}}\phi.$$

One can check that $P_{\mathcal{F}}^* = P_{\mathcal{F}} = P_{\mathcal{F}}^2$. The next exercise shows that conversely, every operator with these properties is of the form $P_{\mathcal{F}}$.

Exercise 1.2.9 Let $\mathcal{H}$ be an inner product space and assume that $P \in \mathcal{L}(\mathcal{H})$ satisfies $P^* = P = P^2$. Show that there exists a linear subspace $\mathcal{F} \subset \mathcal{H}$ such that

[3] Our proof shows that for any linear space $\mathcal{V}$ over $\mathbb{R}$, we may without loss of generality assume that $\mathcal{V}$ is embedded in (i.e., a subspace of) a complex inner product space $\mathcal{W}$ on which is defined a colinear bijection $\phi \mapsto \phi^*$ such that $\mathcal{V} = \{\phi \in \mathcal{W} : \phi^* = \phi\}$. Such a space $\mathcal{W}$ is called a *complexification* of $\mathcal{V}$.

$P = P_{\mathcal{F}}$. Hint: since P is hermitian, we can write $P = \sum_i \lambda_i |e(i)\rangle\langle e(i)|$. Consider $\mathcal{F} := \text{span}\{e(i) : \lambda_i = 1\}$.

In view of Exercise 1.2.9, we call any operator $P \in \mathcal{L}(\mathcal{H})$ such that $P^* = P = P^2$ a *projection*. Obviously, projections are hermitian operators.

By definition, a *partition of the identity* is a finite set of projections $\{P_1, \ldots, P_m\}$ such that

$$\sum_{i=1}^{m} P_i = 1 \quad \text{and} \quad P_i P_j = 0 \quad (i \neq j).$$

If $\mathcal{F}_1, \ldots, \mathcal{F}_m$ are subspaces of $\mathcal{H}$, then $P_{\mathcal{F}_1}, \ldots, P_{\mathcal{F}_m}$ is a partition of the identity if and only if $\mathcal{F}_1, \ldots, \mathcal{F}_m$ are mutually orthogonal and span $\mathcal{H}$. In terms of partitions of the identity, we can formulate Theorem 1.2.5 slightly differently.

Theorem 1.2.10 (Spectral decomposition) *Let $\mathcal{H}$ be an inner product space over $\mathbb{C}$ and let $A \in \mathcal{L}(\mathcal{H})$ be normal. For each $\lambda \in \sigma(A)$, let*

$$\mathcal{F}_\lambda := \{\phi \in \mathcal{H} : A\phi = \lambda\phi\}$$

denote the eigenspace corresponding to the eigenvalue λ. Then $\{P_{\mathcal{F}_\lambda} : \lambda \in \sigma(A)\}$ is a partition of the unity and

$$A = \sum_{\lambda \in \sigma(A)} \lambda P_{\mathcal{F}_\lambda}.$$

Using the spectral decompositon, one can define a 'functional calculus' for normal operators. If $\mathcal{H}$ is a complex inner product space, $A \in \mathcal{L}(\mathcal{H})$, and $f : \mathbb{C} \to \mathbb{C}$ is a function, then one defines a normal operator $f(A)$ by

$$f(A) := \sum_{\lambda \in \sigma(A)} f(\lambda) P_{\mathcal{F}_\lambda}.$$

Exercise 1.2.11 Let $\mathcal{H}$ be an inner product space over $\mathbb{C}$ and let $A \in \mathcal{L}(\mathcal{H})$ be a normal operator. Let $a_0, \ldots, a_n \in \mathbb{C}$ and let $p : \mathbb{C} \to \mathbb{C}$ be the polynomial $p(z) := a_0 + a_1 z + \cdots + a_n z^n$. Let $p(A)$ be defined with the funcional calculus for normal operators. Show that $p(A) = a_0 1 + a_1 A + \cdots + a_n A^n$.

Exercise 1.2.12 Let $\mathcal{H}$ be an inner product space over $\mathbb{C}$ and let $A \in \mathcal{L}(\mathcal{H})$ be a normal operator. Let $f : \mathbb{C} \to \mathbb{C}$ be the function $f(z) := z^*$ and let $f(A)$ be defined with the funcional calculus for normal operators. Show that $f(A) = A^*$.

Exercise 1.2.13 Let $\mathcal{H}$ be an inner product space and $A \in \mathcal{L}(\mathcal{H})$. By definition, $e^A := \sum_{n=0}^{\infty} \frac{A^n}{n!}$. (Since $\mathcal{H}$ is finite dimensional, it is not hard to see that the infinite sum converges.) In the special case that A is normal, show that e^A, defined with the funcional calculus for normal operators, coincides with our previous definition of e^A.

Exercise 1.2.14 Let A be a hermitian operator. Show that e^{iA} (defined with the functional calculus for normal operators) is a unitary operator.

Exercise 1.2.15 Let $\mathcal{H}$ be an inner product space over $\mathbb{C}$ and $A \in \mathcal{L}(\mathcal{H})$. Show that A is hermitian if and only if $\langle\phi|A|\phi\rangle$ is real for all $\phi \in \mathcal{H}$.

Exercise 1.2.16 Let $\mathcal{H}$ be an inner product space over $\mathbb{C}$ and $A \in \mathcal{L}(\mathcal{H})$. Show that the following conditions are equivalent.

(1) A is a positive operator.

(2) $\langle\phi|A|\phi\rangle$ is real and nonnegative for all $\phi \in \mathcal{H}$.

(3) There exists a $B \in \mathcal{L}(\mathcal{H})$ such that $A = B^*B$.

Exercise 1.2.17 Let $\mathcal{H}$ be an inner product space over $\mathbb{C}$, let $\mathcal{F}, \mathcal{G}$ be inear subspaces of $\mathcal{H}$, and let $P_{\mathcal{F}}, P_{\mathcal{G}}$ denote the orthogonal projection operators on $\mathcal{F}, \mathcal{G}$, respectively. Prove that $P_{\mathcal{F}} \leq P_{\mathcal{G}}$ if and only if $\mathcal{F} \subset \mathcal{G}$. Hint: you can use that for any $\psi \in \mathcal{H}$, one has $\|P_{\mathcal{G}}\psi\| \leq \|\psi\|$ with equality if and only if $\psi \in \mathcal{G}$.

Exercise 1.2.18 Let $\mathcal{H}$ be an inner product space over $\mathbb{C}$. Let $P, Q \in \mathcal{L}(\mathcal{H})$ be projection operators and assume that $(1 - P)Q = 0$. Prove that $Q \leq P$.

1.3 Dual, quotient, sum, and product spaces*

Dual spaces

Let $\mathcal{V}$ be a linear space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. By definition,

$$\mathcal{V}' := \mathcal{L}(\mathcal{V}, \mathbb{K})$$

is the *dual* of $\mathcal{V}$. The elements of $\mathcal{V}'$ (usually denoted by l) are called *linear forms* on $\mathcal{V}$. The dual space $\mathcal{V}'$ has the same dimension as $\mathcal{V}$. If $\{e(1), \ldots, e(n)\}$ is a basis for $\mathcal{V}$ then the linear forms $\{f(1), \ldots, f(n)\}$ given by

$$f(i)(e(j)) := \delta_{ij}$$

form a basis of $\mathcal{V}'$, called the *dual basis* of $\{e(1), \ldots, e(n)\}$. There exists a natural isomorphism between $\mathcal{V}$ and its double dual:

$$\mathcal{V} \cong \mathcal{V}''.$$

Here we map a $\phi \in \mathcal{V}$ to the linear form $L_\phi \in \mathcal{L}_\phi(\mathcal{V}', \mathbb{K})$ given by

$$L_\phi(l) := l(\phi) \qquad (l \in \mathcal{V}').$$

Since the kernel of the map $\phi \mapsto L_\phi$ is zero and $\mathcal{V}$ and $\mathcal{V}''$ have the same dimension, this is a linear isomorphism. Note that since $\mathcal{V}$ and $\mathcal{V}'$ have the same dimension, there also exist (many) linear isomorphisms between $\mathcal{V}$ and $\mathcal{V}'$. However, if $\dim(\mathcal{V}) > 1$, it is not possible to choose a 'natural' or 'canonical' linear isomorphism between $\mathcal{V}$ and $\mathcal{V}'$, and therefore we need to distinguish these as different spaces.

If $\mathcal{V}_1, \mathcal{V}_2$ are linear spaces and $A \in \mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$, then by definition its *dual* is the linear map $A' \in \mathcal{L}(\mathcal{V}_2', \mathcal{V}_1')$ defined by

$$A'(l) := l \circ A \qquad (l \in \mathcal{V}_2'),$$

where $\circ$ denotes composition.

If $\mathcal{H}$ is an inner product space then the map $\phi \mapsto \langle\phi|$ is a colinear bijection from $\mathcal{H}$ to $\mathcal{H}'$. In particular,

$$\mathcal{H}' = \{\langle\phi| : \phi \in \mathcal{H}\}.$$

If $\mathcal{H}_1, \mathcal{H}_2$ are inner product spaces and $A \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, then its dual A' is the map

$$A'(\langle\phi|) = \langle A^*\phi| \qquad (\phi \in \mathcal{H}_2).$$

Quotient spaces

Let $\mathcal{V}$ be a linear space over $\mathbb{K}$ and let $\mathcal{W}$ be a linear subspace of $\mathcal{V}$. For any $\phi \in \mathcal{V}$ write $\phi + \mathcal{W} := \{\phi + \psi : \psi \in \mathcal{W}\}$. Then the *quotient space*

$$\mathcal{V}/\mathcal{W} := \{\phi + \mathcal{W} : \phi \in \mathcal{V}\}$$

is a linear space with zero element $0 + \mathcal{W}$ and

$$a(\phi + \mathcal{W}) + b(\psi + \mathcal{W}) := (a\phi + b\psi) + \mathcal{W} \qquad (a, b \in \mathbb{K},\ \phi, \psi \in \mathcal{V}).$$

Exercise 1.3.1 Show that linear combinations in $\mathcal{V}/\mathcal{W}$ are well-defined, i.e., if $\phi + \mathcal{W} = \tilde{\phi} + \mathcal{W}$ and $\psi + \mathcal{W} = \tilde{\psi} + \mathcal{W}$, then $(a\phi + b\psi) + \mathcal{W} = (a\tilde{\phi} + b\tilde{\psi}) + \mathcal{W}$.

Exercise 1.3.2 Let $l : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ be the *quotient map* $l(\phi) := \phi + \mathcal{W}$. Show that $\mathrm{Ker}(l) = \mathcal{W}$ and $\mathrm{Ran}(l) = \mathcal{V}/\mathcal{W}$. Show that

$$\dim(\mathcal{V}) = \dim(\mathcal{V}/\mathcal{W}) + \dim(\mathcal{W}).$$

Exercise 1.3.3 Let $l : \mathcal{V}_1 \to \mathcal{V}_2$ be a linear map. Show that there exists a natural linear isomorphism

$$\mathcal{V}_1/\mathrm{Ker(A)} \cong \mathrm{Ran}(A),$$

Exercise 1.3.4 Let $\mathcal{V}_3 \subset \mathcal{V}_2 \subset \mathcal{V}_1$ be linear spaces. Show that there exists a natural linear isomorphism

$$(\mathcal{V}_1/\mathcal{V}_2) \cong (\mathcal{V}_1/\mathcal{V}_3)/(\mathcal{V}_2/\mathcal{V}_3).$$

The direct sum

Let $\mathcal{V}_1, \ldots, \mathcal{V}_n$ be linear spaces over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. By definition, the *direct sum* of $\mathcal{V}_1, \ldots, \mathcal{V}_n$ is the space

$$\mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_n := \{(\phi(1), \ldots, \phi(n)) : \phi(1) \in \mathcal{V}_1, \ldots, \phi(n) \in \mathcal{V}_n\},$$

which we equip with a linear structure by putting

$$a(\phi(1), \ldots, \phi(n)) + b(\psi(1), \ldots, \psi(n)) := (a\phi(1) + b\psi(1), \ldots, a\phi(n) + b\psi(n)).$$

If $\mathcal{V}$ is some linear space and $\mathcal{V}_1, \ldots, \mathcal{V}_n$ are linear subspaces of $\mathcal{V}$ such that every $\phi \in \mathcal{V}$ can in a unique way be written as $\phi = \phi(1) + \cdots + \phi(n)$ with $\phi(1) \in \mathcal{V}_1, \ldots, \phi(n) \in \mathcal{V}_n$, then there is a natural isomorphism $\mathcal{V} \cong \mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_n$, given by

$$\phi(1) + \cdots + \phi(n) \mapsto (\phi(1), \ldots, \phi(n)).$$

Also in this case, we say that $\mathcal{V}$ is the *direct sum* of $\mathcal{V}_1, \ldots, \mathcal{V}_n$. We often look at a direct sum in this way. Thus, we often view $\mathcal{V}_1, \ldots, \mathcal{V}_n$ as linear subspaces of $\mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_n$, and write $\phi(1) + \cdots + \phi(n)$ rather than $(\phi(1), \ldots, \phi(n))$. One has

$$\dim(\mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_n) = \dim(\mathcal{V}_1) + \cdots + \dim(\mathcal{V}_n).$$

If $\mathcal{U}, \mathcal{W}$ are linear subspaces of $\mathcal{V}$ such that $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$, then the *projection* on $\mathcal{U}$ with respect to this decomposition is the map $P : \mathcal{V} \to \mathcal{U}$ defined by

$$P(\phi + \psi) := \phi \qquad (\phi \in \mathcal{U},\ \psi \in \mathcal{W}).$$

Note that this is a good definition since every $\chi \in \mathcal{V}$ can in a unique way be written as $\chi = \phi + \psi$ with $\phi \in \mathcal{U}$ and $\psi \in \mathcal{W}$. Warning: the definition of P depends not only on $\mathcal{U}$ but also on the choice of $\mathcal{W}$!

Exercise 1.3.5 Show that

$$(\mathcal{U} \oplus \mathcal{W})/\mathcal{W} \cong \mathcal{U}.$$

If $\mathcal{V}$ is a linear space and $\mathcal{W} \subset \mathcal{V}$ a linear subspace, are then $\mathcal{V}$ and $\mathcal{V}/\mathcal{W} \oplus \mathcal{W}$ *in a natural way* isomorphic?

If $\mathcal{H}_1, \ldots, \mathcal{H}_n$ are inner product spaces with inner products $\langle\cdot|\cdot\rangle_1, \ldots, \langle\cdot|\cdot\rangle_n$, respectively, then we equip their direct sum $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$ with the inner product

$$\langle(\phi(1), \ldots, \phi(n))|(\psi(1), \ldots, \psi(n))\rangle := \sum_{i=1}^{n} \langle\phi(i)|\psi(i)\rangle.$$

Note that if we view $\mathcal{H}_1, \ldots, \mathcal{H}_n$ as subspaces of $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$, then these subspaces are mutually orthogonal in the inner product on $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$.

Exercise 1.3.6 Let $\mathcal{H}$ be an inner product space and $\mathcal{F}$ a linear subspace. Show that

$$\mathcal{H} \cong \mathcal{F} \oplus \mathcal{F}^{\perp},$$

where $\cong$ means that the two spaces are isomorphic as inner product spaces.

Exercise 1.3.7 Let $\mathcal{H}$ be an inner product space and $\mathcal{F}$ a linear subspace. Show that $\mathcal{H}/\mathcal{F}$ and $\mathcal{F}^{\perp}$ are isomorphic as linear spaces.

The tensor product

Let $\mathcal{U}, \mathcal{V}$, and $\mathcal{W}$ be linear spaces. By definition, a map $b : \mathcal{U} \times \mathcal{V} \to \mathcal{W}$ is *bilinear* if

$$\phi \mapsto b(\phi, \psi) \quad \text{is linear for each fixed } \psi \in \mathcal{V},$$
$$\psi \mapsto b(\phi, \psi) \quad \text{is linear for each fixed } \phi \in \mathcal{U}.$$

Proposition 1.3.8 (Definition of the tensor product) *For any two linear spaces $\mathcal{U}, \mathcal{V}$ there exists a linear space $\mathcal{U} \otimes \mathcal{V}$, called the tensor product of $\mathcal{U}$ and $\mathcal{V}$, and a bilinear map $(\phi, \psi) \mapsto \phi \otimes \psi$ from $\mathcal{U} \times \mathcal{V}$ into $\mathcal{U} \otimes \mathcal{V}$, satisfying the following equivalent properties*

(i) *If* $\{e(1), \ldots, e(n)\}$ *and* $\{f(1), \ldots, f(m)\}$ *are bases of* $\mathcal{U}$ *and* $\mathcal{V}$*, respectively, then*

$$\big\{e(i) \otimes f(j) : i = 1, \ldots, n,\ j = 1, \ldots, m\big\}$$

is a basis for $\mathcal{U} \otimes \mathcal{V}$*.*

(ii) *For any linear space* $\mathcal{W}$ *and for any bilinear map* $b : \mathcal{U} \times \mathcal{V} \to \mathcal{W}$*, there exists a unique linear map* $\overline{b} : \mathcal{U} \otimes \mathcal{V} \to \mathcal{W}$ *such that* $\overline{b}(\phi \otimes \psi) = b(\phi, \psi)$ *for all* $\phi \in \mathcal{U}, \psi \in \mathcal{V}$*.*

We postpone the proof of Proposition 1.3.8 to the end of this section. The next lemma says that the tensor product of two linear spaces is unique up to linear isomorphisms.

Lemma 1.3.9 (Uniqueness of the tensor product) *Let* $\mathcal{U}, \mathcal{V}$ *be linear spaces. Then the tensor product* $\mathcal{U} \otimes \mathcal{V}$ *of* $\mathcal{U}$ *and* $\mathcal{V}$ *is unique in the following sense. If a linear space* $\mathcal{U}\tilde{\otimes}\mathcal{V}$ *together with a bilinear map* $(\phi, \psi) \mapsto \phi\tilde{\otimes}\psi$ *from* $\mathcal{U} \times \mathcal{V}$ *into* $\mathcal{U}\tilde{\otimes}\mathcal{V}$ *satisfy properties (i) and (ii) of Proposition 1.3.8, then there exist a unique linear bijection* $l : \mathcal{U}\tilde{\otimes}\mathcal{V} \to \mathcal{U} \otimes \mathcal{V}$ *such that* $l(\phi\tilde{\otimes}\psi) = \phi \otimes \psi$ *for all* $\phi \in \mathcal{U}, \psi \in \mathcal{V}$*.*

Proof Since $(\phi, \psi) \mapsto \phi \otimes \psi$ is bilinear, by the fact that $\mathcal{U}\tilde{\otimes}\mathcal{V}$ satisfies property (ii), there exists a unique linear map $l : \mathcal{U}\tilde{\otimes}\mathcal{V} \to \mathcal{U} \otimes \mathcal{V}$ such that $l(\phi\tilde{\otimes}\psi) = \phi \otimes \psi$ for all $\phi \in \mathcal{U}, \psi \in \mathcal{V}$. By property (i), l maps some basis of $\mathcal{U}\tilde{\otimes}\mathcal{V}$ into a basis of $\mathcal{U} \otimes \mathcal{V}$, hence l is a linear bijection. ∎

Exercise 1.3.10 Let Ω_1, Ω_2 be finite sets. Show that $\mathbb{C}^{\Omega_1} \otimes \mathbb{C}^{\Omega_2} \cong \mathbb{C}^{\Omega_1 \times \Omega_2}$. Here $\cong$ means that there is a natural linear bijection between the two spaces.

It is obvious from Proposition 1.3.8 that

$$\dim(\mathcal{U} \otimes \mathcal{V}) = \dim(\mathcal{U}) \dim(\mathcal{V}).$$

We warn the reader that the inclusion

$$\{\phi \otimes \psi : \phi \in \mathcal{U},\ \psi \in \mathcal{V}\} \subset \mathcal{U} \otimes \mathcal{V}$$

is strict. In fact, the set on the left-hand side of this equation spans $\mathcal{U} \otimes \mathcal{V}$, but is itself not a linear space. To see this, note that by property (i) of the tensor product, a general vector $\eta \in \mathcal{U} \otimes \mathcal{V}$ can uniquely be written as

$$\eta = \sum_{i,j} \eta_{ij} e(i) \otimes f(j),$$

where the $\eta_{ij} \in \mathbb{C}$ with $i = 1, \ldots, n$ and $j = 1, \ldots, m$ are the coordinates of η with respect to the basis $\{e(i) \otimes f(i) : i = 1, \ldots, n,\ j = 1, \ldots, m\}$. For general $\phi \in \mathcal{U}$ and $\psi \in \mathcal{V}$, we now have

$$\phi \otimes \psi = \Big(\sum_i \phi_1 e(i) \Big) \Big(\sum_j \psi_j f(j) \Big) = \sum_{ij} \phi_i \psi_j e(i) \otimes f(j),$$

which shows that the coordinates of the vector $\phi \otimes \psi$ with respect to the basis $\{e(i) \otimes f(i) : i = 1, \ldots, n,\ j = 1, \ldots, m\}$ are given by

$$(\phi \otimes \psi)_{ij} = \phi_i \psi_j.$$

Exercise 1.3.11 Assume that $\mathcal{U}$ and $\mathcal{V}$ both have dimension at least 2 and let $\eta \in \mathcal{U} \otimes \mathcal{V}$ be given by

$$\eta := e(1) \otimes f(1) + e(1) \otimes f(2) + e(2) \otimes f(1).$$

Show that $\eta \notin \{\phi \otimes \psi : \phi \in \mathcal{U},\ \psi \in \mathcal{V}\}$. Hint: show that the coordinates η_{ij} are not of the form $\eta_{ij} = \phi_i \psi_j$ for any $\phi \in \mathcal{U}$ and $\psi \in \mathcal{V}$.

If $\mathcal{H}_1, \mathcal{H}_2$ are inner product spaces with inner products $\langle \cdot | \cdot \rangle_1$ and $\langle \cdot | \cdot \rangle_1$, respectively, then we equip the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$ with the inner product

$$\langle \phi(1) \otimes \phi(2) | \psi(1) \otimes \psi(2) \rangle := \langle \phi(1) | \psi(1) \rangle_1 \langle \phi(2) | \psi(2) \rangle_2,$$

for any $\phi(1), \psi(1) \in \mathcal{H}_1$ and $\phi(2), \psi(2) \in \mathcal{H}_2$. In this case, if $\{e(1), \ldots, e(n)\}$ and $\{f(1), \ldots, f(m)\}$ are orthonormal bases of $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, then $\{e(i) \otimes f(j) : i = 1, \ldots, n,\ j = 1, \ldots, m\}$ is an orthonormal bases of $\mathcal{H}_1 \otimes \mathcal{H}_2$.

The next Proposition summarizes some useful additional properties of the tensor product.

Proposition 1.3.12 (Properties of the tensor product) *Let $\mathcal{U}, \mathcal{V}$, and $\mathcal{U} \otimes \mathcal{V}$ be linear spaces and let $(\phi, \psi) \mapsto \phi \otimes \psi$ from $\mathcal{U} \times \mathcal{V}$ into $\mathcal{U} \otimes \mathcal{V}$ be bilinear. Then $\mathcal{U} \otimes \mathcal{V}$, equipped with this map, is the tensor product of $\mathcal{U}$ and $\mathcal{V}$ if and only if the following equivalent conditions hold:*

(iii) *There exist bases $\{e(1), \ldots, e(n)\}$ and $\{f(1), \ldots, f(m)\}$ of $\mathcal{U}$ and $\mathcal{V}$, respectively, such that*

$$\big\{ e(i) \otimes f(j) : i = 1, \ldots, n,\ j = 1, \ldots, m \big\}$$

is a basis for $\mathcal{U} \otimes \mathcal{V}$.

(iv) *For any* $k \in \mathcal{U}'$ *and* $l \in \mathcal{V}'$ *there exists a unique* $p \in (\mathcal{U} \otimes \mathcal{V})'$ *such that* $p(\phi \otimes \psi) = k(\phi)l(\psi)$ *for all* $\phi \in \mathcal{U}, \psi \in \mathcal{V}$.

(v) *For any linear space* $\mathcal{W}$ *and for any map* $b : \mathcal{U} \times \mathcal{V} \to \mathcal{W}$ *that is colinear in each of its arguments, there exists a unique colinear map* $\overline{b} : \mathcal{U} \otimes \mathcal{V} \to \mathcal{W}$ *such that* $\overline{b}(\phi \otimes \psi) = b(\phi, \psi)$ *for all* $\phi \in \mathcal{U}, \psi \in \mathcal{V}$.

Proof of Propositions 1.3.8 and 1.3.12 Consider the properties (i)–(v) from Propositions 1.3.8 and 1.3.12. It is easy to see that there exists a linear space $\mathcal{V} \otimes \mathcal{W}$ and a bilinear map $(\phi, \psi) \mapsto \phi \otimes \psi$ from $\mathcal{U} \times \mathcal{V}$ into $\mathcal{U} \otimes \mathcal{V}$ satisfying property (iii): choose any bases $\{e(1), \ldots, e(n)\}$ and $\{f(1), \ldots, f(m)\}$ of $\mathcal{U}$ and $\mathcal{V}$, let $\mathcal{U} \otimes \mathcal{V}$ be any linear space with dimension nm, choose a basis for $\mathcal{U} \otimes \mathcal{V}$, and give the nm basis vectors the names

$$e(i) \otimes f(j) \qquad (i = 1, \ldots, n,\ j = 1, \ldots, m).$$

If we now define a bilinear map $(\phi, \psi) \mapsto \phi \otimes \psi$ from $\mathcal{U} \times \mathcal{V}$ into $\mathcal{U} \otimes \mathcal{V}$ by

$$\Big(\sum_{i=1}^{n} a_i e(i)\Big) \otimes \Big(\sum_{j=1}^{m} b_j f(j)\Big) := \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j\, e(i) \otimes f(j),$$

then property (iii) holds.
To complete the proof, we will show that (iii)⇒(ii)⇒(iv)⇒(i)⇒(iii) and (ii)⇔(v).
To see that (iii)⇒(ii), we define

$$\overline{b}(e(i) \otimes f(j)) := b(e(i), f(j)) \qquad (i = 1, \ldots, n,\ j = 1, \ldots, m).$$

Since the $e(i) \otimes f(j)$ are a basis of $\mathcal{U} \otimes \mathcal{V}$, this definition extends to a unique linear map $\overline{b} : \mathcal{U} \otimes \mathcal{V} \to \mathcal{W}$. Since b is bilinear, it is easy to see that

$$\overline{b}(\phi \otimes \psi) = b(\phi, \psi) \qquad \forall\ \phi \in \mathcal{U},\ \psi \in \mathcal{V}.$$

This proves (ii).
The implication (ii)⇒(iv) is obvious, since $(\phi, \psi) \mapsto k(\phi)l(\psi)$ is bilinear.
To prove (iv)⇒(i), let $\{e(1), \ldots, e(n)\}$ and $\{f(1), \ldots, f(m)\}$ be bases for $\mathcal{U}$ and $\mathcal{V}$, respectively. We claim that $\{e(i) \otimes f(j) : i = 1, \ldots, n,\ j = 1, \ldots, m\}$ is a basis for $\mathcal{U} \otimes \mathcal{V}$. We start by showing that these vectors are linearly independent. Assume that

$$\sum_{ij} a_{ij} e(i) \otimes f(j) = 0.$$

By our assumption, for any $k \in \mathcal{V}'$ and $l \in \mathcal{U}'$, there exists a unique linear form p on $\mathcal{U} \otimes \mathcal{V}$ such that $p(\phi \otimes \psi) = k(\phi)l(\psi)$ for all $\phi \in \mathcal{U}$, $\psi \in \mathcal{V}$, and therefore,

$$\sum_{ij} a_{ij}\, k(e(i))l(f(j)) = p\Big(\sum_{ij} a_{ij}\, e(i) \otimes f(j)\Big) = p(0) = 0,$$

In particular, we may choose

$$k(e(i)) = \delta_{ii'} \quad \text{and} \quad l(f(j)) = \delta_{jj'}.$$

This shows that $a_{i'j'} = 0$ for all i', j', i.e., the vectors $e(i) \otimes f(j)$ are linearly independent. It is easy to see that if these vectors would not span $\mathcal{U} \otimes \mathcal{V}$, then the linear form p would not be unique, hence they must be a basis for $\mathcal{U} \otimes \mathcal{V}$.
The implication (i)⇒(iii) is trivial.
To see that (ii)⇔(v), finally, we use a trick. If $\mathcal{W}$ is a linear space, then we can always find a linear space $\overline{\mathcal{W}}$ together with a conlinear map $l : \mathcal{W} \to \overline{\mathcal{W}}$ such that l is a bijection. (To see this, take $\overline{\mathcal{W}}$ with the same dimension as $\mathcal{W}$, choose bases $\{e(1), \ldots, e(n)\}$ and $\{f(1), \ldots, f(n)\}$ for $\mathcal{W}$ and $\overline{\mathcal{W}}$, respectively, and set $l(\sum_i a_i e(i)) := \sum_i a_i^* f(i)$.) We call $\overline{\mathcal{W}}$ the *complex conjugate* of $\mathcal{W}$. Now if $b : \mathcal{U} \times \mathcal{V} \to \mathcal{W}$ is colinear in each of its arguments, then $l \circ b : \mathcal{U} \times \mathcal{V} \to \overline{\mathcal{W}}$ is bilinear, and vice versa, so it is easy to see that (i) and (v) are equivalent. ∎

1.4 Tensor calculus*

In this section, we give a short introduction to tensor calculus.
If $k \in \mathcal{U}'$ and $l \in \mathcal{V}'$ are linear forms on $\mathcal{U}$ and $\mathcal{V}$, respectively, then we denote the linear form p on $\mathcal{U} \otimes \mathcal{V}$ from property (iv) of Proposition 1.3.12 by $k \otimes l$. I.e.,

$$k \otimes l(\phi \otimes \psi) := k(\phi)l(\phi) \qquad (\phi \in \mathcal{U},\ \psi \in \mathcal{V}).$$

The next lemma says that this is good notation.

Lemma 1.4.1 (Tensor product and dual spaces) *The linear space $(\mathcal{U} \otimes \mathcal{V})'$ together with the bilinear map $(k, l) \mapsto k \otimes l$ is a version of the tensor product $\mathcal{V}_1' \otimes \mathcal{V}_2'$.*

Proof Choose bases $\{e(1), \ldots, e(n)\}$ and $\{f(1), \ldots, f(m)\}$ for $\mathcal{U}$ and $\mathcal{V}$, respectively. Let $\{k(1, \ldots, k(n)\}$ and $\{l(1), \ldots, l(m)\}$ be the corresponding dual bases for $\mathcal{V}_1'$ and $\mathcal{V}_2'$, i.e.,

$$k(i)(e(j)) := \delta_{ij} \quad \text{and} \quad l(q)(f(r)) := \delta_{qr}.$$

Then

$$k(i) \otimes l(q)\,(e(j) \otimes f(r)) = \delta_{ij}\delta_{qr} = \delta_{(i,q),(j,r)},$$

which shows that $\{k(i) \otimes l(q) : i = 1, \ldots, n,\ q = 1, \ldots, m\}$ is the dual basis to $\{e(j) \otimes f(r) : j = 1, \ldots, n,\ r = 1, \ldots, m\}$. Therefore, by property (i) from Proposition 1.3.8, $(\mathcal{U} \otimes \mathcal{V})'$ together with the bilinear map $(k, l) \mapsto k \otimes l$ is a version of the tensor product $\mathcal{U}' \otimes \mathcal{V}'$. ∎

Tensor products of three or more linear spaces are defined in a similar way as the tensor product of two linear spaces.

Exercise 1.4.2 Show that the tensor product is associative: $(\mathcal{V}_1 \otimes \mathcal{V}_2) \otimes \mathcal{V}_3 \cong \mathcal{V}_1 \otimes (\mathcal{V}_2 \otimes \mathcal{V}_3)$. Here $\cong$ means that there is a natural linear bijection between the two spaces.

If $\mathcal{V}_1, \ldots, \mathcal{V}_m$ are linear spaces with dimensions $d_1, \ldots, d_n$ and

$$\{e^1(1), \ldots, e^1(d_1)\}, \ldots, \{e^n(1), \ldots, e^n(d_n)\}$$

are bases for $\mathcal{V}_1, \ldots, \mathcal{V}_n$, respectively, then the collection of all vectors

$$\left\{e^1(i_1) \otimes \cdots \otimes e^n(i_n) : i_1 = 1, \ldots, d_1,\ \ldots\ i_n = 1, \ldots, d_n\right\}$$

is a basis for $\mathcal{V}_1 \otimes \cdots \otimes \mathcal{V}_n$. A vector $T \in \mathcal{V}_1 \otimes \cdots \otimes \mathcal{V}_n$ is called a *tensor*. It can be expressed in the basis just mentioned as

$$T = \sum_{i_1=1}^{d_1} \cdots \sum_{i_n=1}^{d_n} T_{i_1 \ldots i_n} e^1(i_1) \otimes \cdots e^n(i_n).$$

If $T \in \mathcal{V}_1 \otimes \cdots \otimes \mathcal{V}_m$ and $S \in \mathcal{V}_{m+1} \otimes \cdots \otimes \mathcal{V}_n$ then

$$T \otimes S = \sum_{i_1=1}^{d_1} \cdots \sum_{i_n=1}^{d_n} T_{i_1 \ldots i_m} S_{i_{m+1} \ldots i_n} e^1(i_1) \otimes \cdots e^n(i_n),$$

so

$$(T \otimes S)_{i_1 \ldots i_n} = T_{i_1 \ldots i_m} S_{i_{m+1} \ldots i_n}.$$

Tensor spaces $\mathcal{V}_1 \otimes \cdots \otimes \mathcal{V}_n$ get only really interesting if a linear space and its dual both occur somewhere in the product. For example, if our tensor space has the form $\mathcal{W} \otimes \mathcal{V}' \otimes \mathcal{V}$, then we define the *contraction* of the 2nd and the 3rd coordinate by

$$c_{23}(\psi \otimes l \otimes \phi) := l(\phi)\psi.$$

Note that this formula is linear in each component (i.e., 'trilinear'), so by the analogue of property (i) of Proposition 1.3.8, c_{23} extends to a unique linear map

$$c_{23} : \mathcal{W} \otimes \mathcal{V}' \otimes \mathcal{V} \to \mathcal{W}.$$

If $\{e(1), \ldots, e(n)\}$ is a basis for $\mathcal{V}$, $\{l(1), \ldots, l(n)\}$ is the corresponding dual basis for $\mathcal{V}'$, and $\{f(1), \ldots, f(m)\}$ is a basis for $\mathcal{W}$, then

$$c_{23}\Big(\sum_{ijk} T_{ijk} f(i) \otimes l(j) \otimes e(k)\Big) = \sum_{ijk} T_{ijk} l(j)\big(e(k)\big) f(i) = \sum_{ij} T_{ijj} f(i),$$

so with respect to these bases

$$\big(c_{23}(T)\big)_i = \sum_j T_{ijj}.$$

Without knowing it, we have already seen a number of contractions.

Lemma 1.4.3 (Examples of contractions) *For any linear spaces $\mathcal{V}, \mathcal{W}$ there exists a natural linear isomorphism*

$$\mathcal{L}(\mathcal{V}, \mathcal{W}) \cong \mathcal{W} \otimes \mathcal{V}'.$$

If $\{e(1), \ldots, e(n)\}$ and $\{f(1), \ldots, f(m)\}$ are bases for $\mathcal{V}$ and $\mathcal{W}$, respectively, and $\{l(1), \ldots, l(n)\}$ is the dual basis of $\{e(1), \ldots, e(n)\}$, then in this isomorphism,

$$A = \sum_{ij} A_{ij}\, f_i \otimes l_j,$$

where A_{ij} is the matrix of A written with respect to the bases $\{e(1), \ldots, e(n)\}$ and $\{f(1), \ldots, f(m)\}$. Moreover,

$$\begin{array}{lll}
\text{(i)} & \operatorname{tr}(A) = c_{12}(A) & (A \in \mathcal{V} \otimes \mathcal{V}'), \\
\text{(ii)} & A\phi = c_{23}(A \otimes \phi) & (A \in \mathcal{W} \otimes \mathcal{V}',\ \phi \in \mathcal{V}), \\
\text{(iii)} & A'l = c_{12}(l \otimes A) & (A \in \mathcal{W} \otimes \mathcal{V}',\ l \in \mathcal{W}'), \\
\text{(iv)} & AB = c_{23}(A \otimes B) & (A \in \mathcal{W} \otimes \mathcal{V}',\ B \in \mathcal{V} \otimes \mathcal{U}').
\end{array}$$

Here, in (i), $c_{12} : \mathcal{V} \otimes \mathcal{V}' \to \mathbb{C}$ is defined as $c_{12}(\phi \otimes l) := l(\phi)$. In (iii), $A' \in \mathcal{L}(\mathcal{W}', \mathcal{V}')$ denotes the adjoint of A (see page 17).

Proof of Lemma 1.4.3 Every $A \in \mathcal{W} \otimes \mathcal{V}'$ defines a linear operator $\tilde{A} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ by

$$\tilde{A}\phi := c_{23}(A \otimes \phi)$$

Conversely, we will show that for every $\tilde{A} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ there exists an $A \in \mathcal{W} \otimes \mathcal{V}'$ such that this formula holds. Indeed, if $\tilde{A}_{ij}$ denotes the matrix of $\tilde{A}$, then define $A \in \mathcal{W} \otimes \mathcal{V}'$ by

$$A = \sum_{ij} \tilde{A}_{ij}\, e_i \otimes l_j.$$

Consistent with our notation for tensors, write a vector $\phi \in \mathcal{V}$ as $\phi = \sum_i \phi_i e(i)$. Then

$$c_{23}(A \otimes \phi)_i = \sum_j \tilde{A}_{ij}\phi_j = (\tilde{A}\phi)_i,$$

which shows that $\tilde{A}\phi := c_{23}(A \otimes \phi)$ and therefore $\mathcal{L}(\mathcal{V}, \mathcal{W}) \cong \mathcal{W} \otimes \mathcal{V}'$. Written in coordinates, formulas (i)–(iv) say that

$$\begin{array}{lll}
\text{(i)} & \operatorname{tr}(A) = \sum_i A_{ii} & (A \in \mathcal{V} \otimes \mathcal{V}'), \\
\text{(ii)} & (A\phi)_i = \sum_j A_{ij}\phi_j & (A \in \mathcal{W} \otimes \mathcal{V}',\ \phi \in \mathcal{V}), \\
\text{(iii)} & (A'l)_j = \sum_i l_i A_{ij} & (A \in \mathcal{W} \otimes \mathcal{V}',\ l \in \mathcal{W}'), \\
\text{(iv)} & (AB)_{ij} = \sum_k A_{ik}B_{kj} & (A \in \mathcal{W} \otimes \mathcal{V}',\ B \in \mathcal{V} \otimes \mathcal{U}'),
\end{array}$$

which are all well-known facts. ∎

Chapter 2

Two kinds of probability

2.1 Q-algebras

By definition, an *algebra* is a linear space $\mathcal{A} \neq \{0\}$ over $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, that is equipped with a multiplication $(A, B) \mapsto AB$ from $\mathcal{A} \times \mathcal{A}$ into $\mathcal{A}$ that is associative, bilinear, and has a unit element $1 \in \mathcal{A}$, i.e.,[1]

$$\begin{array}{lll} \text{(i)} & (AB)C = A(BC) & (A, B, C \in \mathcal{A}), \\ \text{(ii)} & A(bB + cC) = bAB + cAC & (A, B, C \in \mathcal{A},\ b, c \in \mathbb{K}), \\ \text{(iii)} & (aA + bB)C = aAC + bBC & (A, B, C \in \mathcal{A},\ a, b \in \mathbb{K}) \\ \text{(iv)} & 1A = A = A1 & (A \in \mathcal{A}). \end{array}$$

Another word for the *unit element* is *identity*. We say that an algebra $\mathcal{A}$ is *abelian* if the multiplication is commutative, i.e.,

$$AB = BA \qquad (A, B \in \mathcal{A}) \qquad .$$

By definition, an *adjoint operation* (also called *involution*) on $\mathcal{A}$ is a map $A \mapsto A^*$ from $\mathcal{A}$ into $\mathcal{A}$ that has the following properties:

$$\begin{array}{lll} \text{(v)} & (A^*)^* = A & (A \in \mathcal{A}), \\ \text{(vi)} & (aA + bB)^* = a^*A^* + b^*B^* & (A, B \in \mathcal{A},\ a, b \in \mathbb{C}), \\ \text{(vii)} & (AB)^* = B^*A^* & (A, B \in \mathcal{A}). \end{array}$$

Here a^* denotes the complex conjugate of a complex number a. Let us say that an adjoint operation is *positive* if

$$\text{(viii)} \quad A^*A = 0 \ \Rightarrow A = 0 \qquad (A \in \mathcal{A}).$$

[1] The existence of a unit element is not always included in the definition of an algebra. Actually, depending on the mathematical context, the word algebra can mean many things.

By definition, a $*$-*algebra* (pronounce: star-algebra) is an algebra $\mathcal{A}$ that is equipped with an adjoint operation. Let us say that $\mathcal{A}$ is a *Q-algebra* if $\mathcal{A}$ is a finite-dimensional $*$-algebra over the complex numbers and the adjoint operation is positive. The term Q-algebra (Q stands for Quantum) is not standard. In fact, Q-algebras, as we have just defined them, are finite dimensional C$*$-algebras; see Section 3.4.

Exercise 2.1.1 Let $\mathcal{H}$ be an inner product space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and let $\mathcal{L}(\mathcal{H})$ be the space of linear operators on $\mathcal{H}$, equipped with operator multiplication and adjugation. Then, obviously, $\mathcal{L}(\mathcal{H})$ is a $*$-algebra. Show that the adjoint operation is positive, i.e., $\mathcal{L}(\mathcal{H})$ satisfies property (viii).

Exercise 2.1.2 Let $\mathcal{A}$ be a $*$-algebra. Show that the space of self-adjoint elements $\mathcal{A}_r := \{A \in \mathcal{A} : A^* = A\}$ is a real linear subspace of $\mathcal{A}$. Show that each $A \in \mathcal{A}$ can in a unique way be written as $A = \mathrm{Re}(A) + i\mathrm{Im}(A)$ with $\mathrm{Re}(A), \mathrm{Im}(A) \in \mathcal{A}_r$.

Exercise 2.1.3 Let $\mathcal{H}$ be an inner product space over $\mathbb{C}$ and let $A \in \mathcal{L}(\mathcal{H})$. Show that $A^*A = \mathrm{Re}(A)^2 + \mathrm{Im}(A)^2$ if and only if A is normal.

Let $\mathcal{A}, \mathcal{B}$ be algebras. We say that that a map $l : \mathcal{A} \to \mathcal{B}$ is an *algebra homomorphism* if

$$\begin{array}{lll} \text{(a)} & l(aA + bB) = al(A) + bl(B) & (A, B \in \mathcal{A},\ a, b \in \mathbb{C}), \\ \text{(b)} & l(AB) = l(A)l(B) & (A, B \in \mathcal{A}), \\ \text{(c)} & l(1) = 1. & \end{array}$$

If $\mathcal{A}, \mathcal{B}$ are $*$-algebras, then l is called a $*$-algebra homomorphism if moreover

$$\text{(d)} \quad l(A^*) = l(A)^* \qquad (A \in \mathcal{A}).$$

If an algebra homomorphism (resp. $*$-algebra homomorphism) l is a bijection then one can check that also l^{-1} is also an algebra homomorphism (resp. $*$-algebra homomorphism). In this case we call l an *algebra isomorphism* (resp. $*$-algebra isomorphism) and we say that $\mathcal{A}$ and $\mathcal{B}$ are *isomorphic* as algebras (resp. as $*$-algebras).

By definition, a *subalgebra* of an algebra $\mathcal{A}$ is a linear subspace $\mathcal{A}' \subset \mathcal{A}$ such that $1 \in \mathcal{A}'$ and $\mathcal{A}'$ is closed under multiplication. If $\mathcal{A}$ is a $*$-algebra then we call $\mathcal{A}'$ a sub-$*$-algebra if moreover $\mathcal{A}'$ is closed under adjugation. If $\mathcal{A}'$ is a subalgebra (resp. sub-$*$-algebra) of $\mathcal{A}$, then $\mathcal{A}'$, equipped with the multiplication and adjoint operation from $\mathcal{A}$, is itself an algebra (resp. $*$-algebra).

Exercise 2.1.4 Let $\mathcal{A}, \mathcal{B}$ be $*$-algebras and let $l : \mathcal{A} \to \mathcal{B}$ be a $*$-algebra homomorphism. Show that the range $\mathrm{Ran}(l) := \{l(A) : A \in \mathcal{A}\}$ of l is a sub-$*$-algebra of $\mathcal{B}$.

A *representation* of an algebra $\mathcal{A}$ over $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$ is a linear space $\mathcal{H}$ over $\mathbb{K}$ together with an algebra homomorphism $l : \mathcal{A} \to \mathcal{L}(\mathcal{H})$. If $\mathcal{A}$ is a $*$-algebra, then we also require that $\mathcal{H}$ is equipped with an inner product such that $l : \mathcal{A} \to \mathcal{L}(\mathcal{H})$ is a $*$-algebra homomorphism. (Otherwise, we speak of a representation of $\mathcal{A}$ *as an algebra*.) A representation is *faithful* if l is one-to-one. Note that in this case, l is an algebra isomorphism (resp. $*$-algebra isomorphism) between $\mathcal{A}$ and the subalgebra (resp. sub-$*$-algebra) $\mathrm{Ran}(l) \subset \mathcal{L}(\mathcal{H})$.

A basic result about Q-algebras is:

Theorem 2.1.5 (Representation of positive $*$-algebras) *Every Q-algebra has a faithful representation.*

Unfortunately, the proof of Theorem 2.1.5 is mildly complicated. For a proof, we refer the reader to [GHJ89, Appendix II.a] or [Swa04]. A rough sketch of the proof will be given in Section 5.8. Those who are not satisfied with this may find some consolation in hearing that, actually, we will not use Theorem 2.1.5 at all. Replace 'Q-algebra' by 'representable Q-algebra' in what follows, and all proofs remain valid. While it is certainly nice to know that these notions coincide, we will never really need this.

Theorem 2.1.5 says that every Q-algebra $\mathcal{A}$ is isomorphic to some sub-$*$-algebra $\mathcal{A}' \subset \mathcal{L}(\mathcal{H})$, for a suitable inner product space $\mathcal{H}$. Thus, we may think of the elements of $\mathcal{A}$ as linear operators on an inner product space $\mathcal{H}$. We must be careful, however, since some properties of these operators may depend on the (faithful) representation. A lot, however, turns out to be representation independent.

We start by noting that being a normal operator is, obviously, representation independent. The same is true for being a hermitian operator, being a projection, or being a partition of the identity. (Note that $\{P_1, \ldots, P_n\}$ is a partition of the identity iff $P_i = P_i^*$, $P_i P_j = \delta_{ij} P_i$ and $\sum_i P_i = 1$.) In fact, the whole spectral decomposition of normal operators is representation independent:

Lemma 2.1.6 (Spectral decomposition is representation independent) *Let $\mathcal{H}$ be a complex inner product space and let $\mathcal{A}$ be a sub-algebra of $\mathcal{L}(\mathcal{H})$. Assume that $A \in \mathcal{A}$ is normal. Then A can uniquely be written as*

$$A = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda$$

where $\sigma(A)$ is a finite subset of $\mathbb{C}$ and $\{P_\lambda : \lambda \in \sigma(A)\}$ is a partition of the identity. Moreover, $P_\lambda \in \mathcal{A}$ for all $\lambda \in \sigma(A)$.

Proof By Theorem 1.2.10, the operator A can uniquely be written as $A = \sum_{\lambda\in\sigma(A)} \lambda P_\lambda$, where $\sigma(A)$ is a finite subset of $\mathbb{C}$ and $\{P_\lambda : \lambda \in \sigma(A)\}$ is a partition of the identity. Fix $\lambda \in \sigma(A)$. We claim that $P_\lambda \in \mathcal{A}$. To prove this, choose a polynomial p such that $p(\lambda) = 1$ and $p(\lambda') = 0$ for all $\lambda' \in \sigma(A)$, $\lambda' \neq \lambda$. Then $P_\lambda = p(A)$, where $p(A)$ is defined using the functional calculus for normal operators. By Excercise 1.2.11, $p(A) \in \mathcal{A}$. ∎

By Lemma 2.1.6, the spectrum of a normal operator is representation independent. It follows that being a unitary operator, or a positive operator, is also representation independent. The functional calculus for normal operators is also representation independent.

Lemma 2.1.7 (Functional calculus is representation independent) *Let $\mathcal{A}$ and $\tilde{\mathcal{A}}$ be Q-algebras, let $A \in \mathcal{A}$ be normal, let $f : \mathbb{C} \to \mathbb{C}$ be a function, and let $l : \mathcal{A} \to \tilde{\mathcal{A}}$ be an algebra homomorphism. Then $f(A) \in \mathcal{A}$ and $f(l(A)) = l(f(A))$.*

Proof Immediate from Lemma 2.1.6 ∎

Just when we start to believe that almost everything we can think of is representation independent, a little warning is in place:

Exercise 2.1.8 Show that the trace of an operator is *not* a representation independent quantity. Hint: observe that the Q-algebra consisting of all operators of the form

$$\begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{pmatrix} \qquad (a, b, c, d \in \mathbb{C})$$

is isomorphic with $\mathcal{L}(\mathbb{C}^2)$.

Exercise 2.1.9 Let $\mathcal{A}$ be the space of all matrices of the form

$$\begin{pmatrix} a & -c & -b \\ b & a & -c \\ c & b & a \end{pmatrix} \qquad \text{with } a, b, c \in \mathbb{C}.$$

Equip $\mathcal{A}$ with the usual matrix multiplication and define an adjoint operation on $\mathcal{A}$ by

$$\begin{pmatrix} a & -c & -b \\ b & a & -c \\ c & b & a \end{pmatrix}^* := \begin{pmatrix} a^* & -c^* & -b^* \\ b^* & a^* & -c^* \\ c^* & b^* & a^* \end{pmatrix}.$$

Show that $\mathcal{A}$ is a $*$-algebra. Is $\mathcal{A}$ abelian? Is the adjoint operation positive? (Hint: consider the operator

$$X := \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Show that a general element of $\mathcal{A}$ is of the form $a1 + bX + cX^2$.)

2.2 Probability spaces

For any set Ω, we write $\mathcal{P}(\Omega) := \{A : A \subset \Omega\}$ to denote the set of all subsets of Ω. On $\mathcal{P}(\Omega)$ are defined set operations such as $A \cap B$, $A \cup B$, and

$$\begin{aligned} A\backslash B &:= \{\omega \in A : \omega \notin B\}, \\ A^{\mathrm{c}} &:= \Omega\backslash A. \end{aligned}$$

By definition, a finite *probability space* is a triple $(\Omega, \mathcal{P}(\Omega), \mu)$, where Ω is a finite set, $\mathcal{P}(\Omega)$ is the set of all subsets of Ω, and $\mu : \mathcal{P}(\Omega) \to [0, 1]$ is a function with the following properties:

(a) $\mu(\Omega) = 1$,
(b) $A, B \subset \Omega,\ A \cap B = \emptyset \Rightarrow \mu(A \cup B) = \mu(A) + \mu(B)$.

We call Ω the *state space*, $\mathcal{P}(\Omega)$ the *space of events* and μ a *probability law*.

Exercise 2.2.1 Show that every probability law on a finite set Ω is of the form

$$\mu(A) = \sum_{\omega \in A} m(\omega),$$

where $m : \Omega \to [0, 1]$ is a function satisfying $\sum_{\omega \in \Omega} m(\omega) = 1$.

We interpret a finite probability space $(\Omega, \mathcal{P}(\Omega), \mu)$ as follows.

1° A finite probability space $(\Omega, \mathcal{P}(\Omega), \mu)$ describes incomplete knowledge about a system in the physical reality.

2° The state space Ω contains elements ω, called states. Each state gives an exhausting description of all properties of the physical system that are of interest to us.

3° A subset $A \subset \Omega$ is interpreted as the event that the actual state of the physical system lies in A. In this interpretation, A^c is the event 'not A', $A \cap B$ is the event 'A and B', $A \cup B$ is the event 'A and/or B', $A \backslash B$ is the event 'A and not B', and so on.

4° The probability law μ assigns to each event $A \in \mathcal{P}(\Omega)$ a number $\mu(A) \in [0, 1]$, called the probability of A. The probability law $\mu(A)$ measures how likely we judge the event A to be true on the basis of our incomplete knowledge. The larger $\mu(A)$ is, the more likely is A. If $\mu(A) = 1$ then A is sure.

5° If we observe that an event B is true, then our knowledge about the physical system changes. We express our changed knowledge with a new probability law $\tilde{\mu}$ on $\mathcal{P}(\Omega)$, defined as $\tilde{\mu}(A) := \mu(A \cap B)/\mu(B)$. This formula is not defined if $\mu(B) = 0$ but in that case we were sure that the event B was not true before we performed our observation, so in this situation there was something wrong with the way we described our knowledge before the observation.

In point 5°, we call $\tilde{\mu}(A) := \mu(A \cap B)/\mu(B)$ the conditional probability of the event A given B, and we call $\tilde{\mu}$ the conditioned probability law. We also use the notation

$$\mu(A|B) := \mu(A \cap B)/\mu(B) \qquad (A, B \in \mathcal{P}(\Omega),\ \mu(B) > 0).$$

The interpretation of finite probability spaces we have just given is not undisputed. Many authors insist that an interpretation of probability spaces must link probabilities in some way to relative frequencies, either by saying that the probability of an event is likely to be the relative frequency of that event in a long sequence of independent trials, or by saying that the probability of an event is the relative frequency of that event in an infinite sequence of independent trials. The appeal of these interpretations lies in the fact that they refer directly to the way probabilities are experimentally measured.

The difficulty with the first definition is that 'likely to be' seems to involve the concept of probability again, while the difficulty with the second definition is that infinite sequences of independent trials do not occur in reality. Both definitions have the difficulty that they lean heavily on the concept of independence, the definition of which also seems to involve probabilities. The disadvantage of the interpretation we have just given is that the additive property (b) of probability laws has no justification, but the point of view taken here is that nature is as it is and does not need justification.

By definition, a real-valued *random variable*, defined on a finite probability space $(\Omega, \mathcal{P}(\Omega), \mu)$, is a function $X : \Omega \to \mathbb{R}$. We interpret the event

$$\{X = x\} := \{\omega \in \Omega : X(\omega) = x\}$$

as the event that the random variable X takes on the value x. Similarly, we write $\{X < x\} := \{\omega \in \Omega : X(\omega) < x\}$ to denote the event that X takes on a value smaller than x, and so on. Note that since Ω is finite, the range $\mathcal{R}(X) = \{X(\omega) : \omega \in \Omega\}$ is finite. We call

$$\int X \, \mathrm{d}\mu := \sum_{\omega \in \Omega} X(\omega)\mu(\omega) = \sum_{x \in \mathcal{R}(X)} x \, \mu(\{X = x\})$$

the *expected value* of X.

Example Consider a shuffled deck of cards from which the jacks, queens, kings, and aces have been removed. Let $V := \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$ be the set of values and $C := \{\text{heart,spade,diamond,clover}\}$ the set of colors. Then $C \times V = \{(c, v) : c \in C, v \in V\}$ is the set of all cards in our deck and

$$\Omega := \big\{\big((c_1, v_1), \ldots, (c_{36}, v_{36})\big) : (c_i, v_i) \neq (c_j, v_j) \; \forall i \neq j, \; (c_i, v_i) \in C \times D \; \forall i\big\}$$

is the set of all permutations of $C \times V$. We choose Ω as our state space. A state $\omega = \big((c_1, v_1) \ldots, (c_{36}, v_{36})\big) \in \Omega$ describes the cards in our reduced deck, ordered from top to bottom. Since we believe that every order of the cards has the same probability, we choose as our probability law

$$\mu(A) := \frac{|A|}{|\Omega|} \qquad (A \in \mathcal{P}(\Omega)),$$

where $|A|$ denotes the number of elements in a set A. For example, the set

$$A := \{\big((c_1, v_1) \ldots, (c_{36}, v_{36})\big) \in \Omega : c_1 = c_2\}$$

describes the event that the first two cards have the same color. The probability of this event is

$$\mu(A) = \frac{|A|}{|\Omega|} = \frac{36 \cdot 8 \cdot 34!}{36!} = \frac{8}{35}.$$

The random variable

$$X\big((c_1, v_1) \ldots, (c_{36}, v_{36})\big) := v_1$$

describes the value of the first card. The expected value of X is

$$\int X \, \mathrm{d}\mu = \sum_{x=2}^{10} x\mu(\{X = x\}) = \frac{1}{9} \sum_{x=2}^{10} x = \frac{55}{9} = 6\tfrac{1}{9}.$$

2.3 Quantum probability spaces

By definition, a (finite dimensional) *quantum probability space* is a pair $(\mathcal{A}, \rho)$ where $\mathcal{A}$ is a Q-algebra and $\rho : \mathcal{A} \to \mathbb{C}$ is a function with the following properties:

(a)	$\rho(aA + bB) = a\rho(A) + b\rho(B)$	$(A, B \in \mathcal{A},\ a, b \in \mathbb{C})$,
(b)	$\rho(A^*) = \rho(A)^*$	$(A \in \mathcal{A})$,
(c)	$\rho(A^*A) \geq 0$	$(A \in \mathcal{A})$,
(d)	$\rho(1) = 1$.	

We call ρ a *probability law* on $\mathcal{A}$. Note that by property (b), $\rho(A^*A)$ is a real number for all $A \in \mathcal{A}$. By Exercise 1.2.16, property (c) is equivalent to saying that $\rho(A) \geq 0$ whenever A is a positive operator. Note that by linearity this implies that $\rho(A) \leq \rho(B)$ whenever $A \leq B$.

We interpret a quantum probability space $(\mathcal{A}, \rho)$ as follows.

1° A quantum probability space $(\mathcal{A}, \rho)$ describes incomplete knowledge about a system in the physical reality.

2° We interpret a projection $P \in \mathcal{A}$ as a possible observation on the system. We interpret a partition of the identity $\{P_1, \ldots, P_n\}$ as an ideal measurement on the system, that can yield the observations $P_1, \ldots, P_n$.

3° The probability law ρ assigns to each observation $P \in \mathcal{A}$ a probability $\rho(P)$. The probability $\rho(P)$ measures how likely we judge it to be that an ideal measurement $\{P_1, P_2, \ldots, P_n\}$ with $P = P_i$ for some i, will yield the observation P, if we perform the measurement. The larger $\rho(P)$ is, the more likely is P. If $\rho(P) = 1$, then any measurement that can yield P will surely yield it, if we perform the measurement.

4° If someone performs the ideal measurement $\{P_1, \ldots, P_n\}$ on the system, then this in general influences the system, with the result that we must describe our knowledge about the system with a new probability law ρ' on $\mathcal{A}$, defined as $\rho'(A) := \sum_{i=1}^n \rho(P_i A P_i)$.

5° If an ideal measurement is performed on the system and we learn that this measurement has yielded the observation P, then our knowledge about the system changes. We must describe our changed knowledge with a new probability law $\tilde{\rho}$ on $\mathcal{A}$, defined as $\tilde{\rho}(A) := \rho(PAP)/\rho(P)$. This formula is not defined if $\rho(P) = 0$ but in that case we were sure that the ideal measurement would not yield P, so that in this situation there was something wrong with the way we described our knowledge before the observation.

Exercise 2.3.1 If ρ is a probability law on $\mathcal{A}$ and $\{P_1, \ldots, P_n\}$ is a partition of the identity, then show that $\rho(P_1), \ldots, \rho(P_n)$ are nonnegative real numbers, summing up to one. Show that the functions ρ' and $\tilde{\rho}$ defined in point 4° and 5°, respectively, are probability laws on $\mathcal{A}$.

A characteristic property of the interpretation of quantum probability we have just given is the central role played by ideal measurements. While not every measurement is 'ideal', for the interpretation given above it is essential that we have a collection of measurements at our disposal that for all practical purposes may be regarded as ideal. Typically, observations in our everyday macroscopic world that do not disturb the subject we are measuring are ideal. For example, seeing a subject with our eyes of hearing it make a sound may typically be regarded as an ideal observation on that subject.

Although the rules of quantum mechanics presumably govern everything around us, the typical quantum mechanical effects can usually only be observed on particles that are extremely small, like electrons, protons, or photons. Therefore, we typically need some delicate measuring equipment to observe these objects. While the observations we perform on the measuring equipment (e.g. reading off a display) may for all practical purposes be regarded as an ideal measurement on the equipment, it is not always true that the resulting effect on our objects of interests (such as electrons, protons, or photons) is that of an ideal measurement. In order to determine this, we need to study the complex physical (quantum mechanical) laws governing the interaction of the measuring equipment with our objects of interest. Since this falls outside the scope of the present lecture notes, we will usually take the possibility of performing ideal measurements for granted.

Apart from the central role played by ideal measurements, two awkward differences between quantum probability and classical probability strike us immediately. First of all, the states ω that play such an important role in classical probability have completely disappeared from the picture. Second, the bare fact that someone performs a measurement on a system, even when we don't know the outcome, changes the system in such a way that we must describe our knowledge with a new probability law ρ'. In the next section we will see that if the algebra $\mathcal{A}$ is abelian, then these differences are only seemingly there, and in fact we are back at classical probability. On the other hand, if $\mathcal{A}$ is not abelian, quantum probabilities are really different, and pose a serious challenge to our imagination.

The interpretation of quantum mechanics is notoriously difficult, and the interpretation we have just given is not undisputed. There is an extensive literature on the subject in which innumerably many different interpretations have been suggested,

with the result that almost everything one can say on this subject has at some point been fiercely denied by someone. As an introduction to some of the different points of view, the book by Redhead [Red87] is very readable.

Not only the interpretation of quantum mechanics, but also the presentation of the mathematical formalism shows a broad variation in the literature. Apart from the approach taken here, one finds introductions to quantum mechanics based on wave functions, Hilbert spaces, or projection lattices. To add to the confusion, it is tradition to call the probability law ρ a 'mixed state', even though it is conceptually something very different from the states ω of classical probability.

In quantum probability, hermitian operators are called *observables*. They correspond to real-valued physical quantities and may be regarded as the equivalent of the real random variables from classical probability. Let

$$A = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda$$

be the spectral decomposition of a hermitian operator A in some Q-algebra. We interpret

$$\{P_\lambda : \lambda \in \sigma(A)\}$$

as an ideal measurement of the observable A. We interpret P_λ as the observation that A takes on the value λ. We call

$$\rho(A) = \sum_{\lambda \in \sigma(A)} \lambda \rho(P_\lambda)$$

the *expected value* of A.

Example (Polarization) It is well-known that light can be decomposed into two polarization directions, perpendicular to the direction in which it travels. For example, polaroid sunglasses usually filter the vertically polarized component of light away, leaving only the horizontally polarized component. Using prismas, it is possible to split a light beam into two orthogonally polarized beams. Apart from the well-known *linear polarization*, there is also the often-neglected *circular polarization*, which has recently come into the limelight because of its use in 3D cinemas. The most general form of polarization is *elliptic polarization*, which interpolates between linear and circular polarization.

On the level of the individual *photons* (light particles), our knowledge of the polarization of a single photon can be described by a probability law on a Q-algebra of the form $\mathcal{L}(\mathcal{H})$, where $\mathcal{H}$ is a two-dimensional inner product space. An ideal

measurement of the polarization is described by a partition of the identity $\{P, Q\}$, where P and Q project on orthogonal one-dimensional subspaces of $\mathcal{H}$. Each one-dimensional subspace of $\mathcal{H}$ corresponds to a certain way in which the photon can be polarized.
For concreteness, let us assume that the photon moves horizontally and in the direction of the observer. We may choose an orthonormal basis $\{e(1), e(2)\}$ of $\mathcal{H}$ such that the linear subspaces spanned by $e(1)$ and $e(2)$ correspond to horizontal and vertical linear polarization, respectively. Let $\mathcal{F} \subset \mathcal{H}$ be a general one-dimensional subspace of $\mathcal{H}$. To determine what kind of polarization $\mathcal{F}$ corresponds to, we choose a vector $\phi \in \mathcal{F}$ with norm $\|\phi\| = 1$, and consider the function $\phi(t) := e^{it}\phi$ ($t \in \mathbb{R}$). With respect to the basis $\{e(1), e(2)\}$, we can write ϕ and more generally $\phi(t)$ in coordinates as

$$\phi = \phi_1 e(1) + \phi_2 e(2) \quad \text{and} \quad \phi(t) = \phi_1(t)e(1) + \phi_2(t)e(2).$$

Now let us look at the function

$$t \mapsto \big(\mathrm{Re}(\phi_1(t)), \mathrm{Re}(\phi_2(t))\big). \tag{2.1}$$

This function is obviously periodic, with period 2π. If we had chosen another vector $\phi' \in \mathcal{F}$ of norm one, then $\phi' = e^{is}\phi$ for some $s \in [0, 2\pi)$, so except for a time-shift, we would have obtained the same function. In general, the function in (2.1) moves in an ellips around the origin.
In the special case that $\phi = e(1)$, this ellips reduces to a horizontal line, in line with the fact that the subspace spanned by $e(1)$ corresponds to a horizontally polarized photon. Likewise, if $\phi = e(2)$, then the function in (2.1) moves in a vertical line, which corresponds to vertical polarization. More generally, if ϕ is of the form

$$\phi = \cos(\alpha)e(1) + \sin(\alpha)e(2) =: \eta(\alpha), \tag{2.2}$$

with $\alpha \in [0, 2\pi)$, then the function in (2.1) becomes

$$t \mapsto \big(\cos(\alpha), \sin(\alpha)\big)\mathrm{Re}(e^{it}),$$

which moves in a line that makes an angle α with the horizontal line. This corresponds to linearly polarized light in the direction α. Note that $\eta(\alpha) = -\eta(\alpha + \pi)$, so $\eta(\alpha)$ and $\eta(\alpha + \pi)$ span the same subspace of $\mathcal{H}$.
If we take

$$\phi = \tfrac{1}{\sqrt{2}}\big(e(1) \pm ie(2)\big),$$

then the function in (2.1) becomes

$$t \mapsto \tfrac{1}{\sqrt{2}}\big(\mathrm{Re}(e^{it}), \mathrm{Re}(e^{i(t\pm\pi/2)})\big) = \tfrac{1}{\sqrt{2}}\big(\cos(t), \mp\sin(t)\big),$$

which moves in a circle of radius $\frac{1}{\sqrt{2}}$ around the origin. Depending on whether the function moves in a clockwise or anticlockwise direction, this corresponds to the two possible forms of circular polarization.
Polarization of photons can be measured. The easiest way is to place a filter in a beam of photons that lets pass only photons of a certain polarization (linear in a given direction or circular with a given orientation). For a good filter, if P denotes the projection on the subspace of the give type of polarization, and ρ is a probability law describing our knowledge about the polarization of the photon before it reaches the filter, then the probability that the photon passes the filter is $\rho(P)$, and all photons that pass the filter have to be described by the new probability law $\tilde{\rho}(A) := \rho(PAP)/\rho(P)$. Thus, this is like performing the ideal measurement $\{P, 1-P\}$ on all incoming photons, and then throwing away those that did not yield the desired outcome of the measurement.
For linear polarization, there is a better way of measuring polarization, that does not destroy any photons. Using a prisma, we can split an incoming beam of photons into two beams, going in different directions, of which one beam is polarized in a direction α and the other in the perpendicular direction $\alpha + \pi/2$. On each beam, we can then perform further experiments; in particular, we can split each outgoing beam into a part that is polarized in a direction α' and a direction $\alpha' + \pi/2$. If the original beam has polarization α, then, as we will see in Excercise 2.3.2 below, the fraction of photons that will be found to have polarization direction α' in this second experiment is $\cos(\alpha' - \alpha)^2$. Although this comes closer to being an ideal measurement, in practice, we do not know when a photon passes through a certain beam. In the end, we still have to detect the photon. Although it is possible to detect a single photon, this usually means destroying it.

Exercise 2.3.2 Let $\eta(\alpha)$ be defined as in (2.2). For each $\alpha \in \mathbb{R}$, let P_α denote the projection operator $P_\alpha := |\eta(\alpha)\rangle\langle\eta(\alpha)|$. Let ρ be any probability law on $\mathcal{L}(\mathcal{H})$ such that $\rho(P_\alpha) > 0$, and define $\rho_\alpha(A) := \rho(P_\alpha A P_\alpha)/\rho(P_\alpha)$. Prove that

$$\rho_\alpha(A) = \langle\eta(\alpha)|A|\eta(\alpha)\rangle.$$

Prove in particular that

$$\rho_\alpha(P_{\alpha'}) = \cos(\alpha' - \alpha)^2.$$

Show that the projections P_α and P_β in different directions α and β in general do not commute.

Example (Spin) Electrons have a property called *spin*, which is a form of angular momentum. Let $\mathcal{H}$ be a two-dimensional inner product space with orthonormal

basis $\{e(1), e(2)\}$. Define hermitian operators $S_{\mathrm{x}}, S_{\mathrm{y}}, S_{\mathrm{z}} \in \mathcal{L}(\mathcal{H})$ by their matrices with respect to $\{e(1), e(2)\}$ as:

$$S_{\mathrm{x}} := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$
$$S_{\mathrm{y}} := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$
$$S_{\mathrm{z}} := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Chosing an appropriate basis, we can describe the three-dimensional space that we live in by $\mathbb{R}^3$. Let $\theta = (\theta_{\mathrm{x}}, \theta_{\mathrm{y}}, \theta_{\mathrm{z}}) \in \mathbb{R}^3$ be a vector such that $\|\theta\| = \theta_{\mathrm{x}}^2 + \theta_{\mathrm{y}}^2 + \theta_{\mathrm{z}}^2 = 1$. Then the spin of an electron in the direction θ is a physical quantity, described by the observable

$$S_\theta := \theta_{\mathrm{x}} S_{\mathrm{x}} + \theta_{\mathrm{y}} S_{\mathrm{y}} + \theta_{\mathrm{z}} S_{\mathrm{z}}.$$

One can check that its spectrum is

$$\sigma(S_\theta) = \{-1, +1\}.$$

Thus, no matter in which direction θ we measure the spin of an electron, we can always find only two values: -1 ('spin down') or $+1$ ('spin up'). The matrices $S_{\mathrm{x}}, S_{\mathrm{y}}$, and S_{z} are called the *Pauli matrices*. Ideal measurements of the spin of an electron are possible, using magnetic fields that deflect electrons in a beam in different directions depending on their spin.

2.4 (Non)commutative probability

Although the quantum probability spaces and their interpretation from Section 2.3 seem rather different from the 'classical' probability spaces from Section 2.2, we will see here that the latter are actually a special case of the former. More precisely, we will show that a quantum probability space $(\mathcal{A}, \rho)$ is equivalent to a 'classical' probability space $(\Omega, \mathcal{P}(\Omega), \mu)$ if and only if the algebra $\mathcal{A}$ is abelian.

If Ω is a finite set, we write

$$\mathbb{C}^\Omega := \{f : \Omega \to \mathbb{C}\}$$

to denote the space of all functions from Ω into $\mathbb{C}$. We equip $\mathbb{C}^\Omega$ with the structure of a $*$-algebra in the obvious way, i.e.,

$$\begin{array}{ll} (af + bg)(\omega) := af(\omega) + bg(\omega) & (f, g \in \mathbb{C}^\Omega,\ a, b \in \mathbb{C},\ \omega \in \Omega), \\ (fg)(\omega) := f(\omega)g(\omega) & (f, g \in \mathbb{C}^\Omega,\ \omega \in \Omega), \\ f^*(\omega) := f(\omega)^* & (f \in \mathbb{C}^\Omega,\ \omega \in \Omega). \end{array}$$

It is clear from the second relation that $\mathbb{C}^\Omega$ is abelian. Note that $\mathbb{C}^\Omega$ satisfies property (viii) from the Section 2.1, i.e., $\mathbb{C}^\Omega$ is a Q-algebra. The next theorem shows that there is a one-to-one correspondence between abelian quantum probability spaces and classical probability spaces.

Theorem 2.4.1 (Abelian Q-algebras) *Let $\mathcal{A}$ be a Q-algebra. Then $\mathcal{A}$ is abelian if and only if $\mathcal{A}$ is isomorphic to a Q-algebra of the form $\mathbb{C}^\Omega$, where Ω is a finite set. If $\mu : \mathcal{P}(\Omega) \to \mathbb{R}$ is a probability law, then*

$$\rho(f) := \int f \, \mathrm{d}\mu \tag{2.3}$$

defines a probability law on $\mathbb{C}^\Omega$, and conversely, every probability law ρ on $\mathbb{C}^\Omega$ arises in this way.

We defer the proof of Theorem 2.4.1 to Section 5.7.

Remark We may represent the algebra $\mathbb{C}^\Omega$ as an algebra of linear operators on an inner product space $\mathcal{H}$ as follows. Enumerate the elements of Ω in some way, say $\Omega = \{1, \ldots, n\}$, and let $\mathcal{A}$ be the algebra of all diagonal $n \times n$ matrices of the form

$$A_{ij} = \delta_{ij} f(i) \qquad (i = 1, \ldots, n),$$

where $f \in \mathbb{C}^\Omega$. Then obviously $\mathcal{A} \cong \mathbb{C}^\Omega$.

It is not hard to see that an element f of the abelian Q-algebra $\mathbb{C}^\Omega$ is a projection if and only if $f = 1_A$ for some $A \subset \Omega$, where for any subset $A \subset \Omega$ the *indicator function* $1_A \in \mathbb{C}^\Omega$ is defined as

$$1_A(\omega) := \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{if } \omega \notin A. \end{cases}$$

An ideal measurement on $\mathbb{C}^\Omega$ is a collection of indicator functions $\{1_{A_1}, \ldots, 1_{A_n}\}$ where $\{A_1, \ldots, A_n\}$ is a partition of Ω, i.e., $A_i \cap A_j = \emptyset$ for all $i \neq j$ and $A_1 \cup \cdots \cup A_n = \Omega$. Thus, ideal measurements on $\mathbb{C}^\Omega$ determine which of the mutually exclusive events $A_1, \ldots, A_n$ takes place. We can list the corresponding notions in classical and quantum probability in the following table:

Classical probability	**Quantum probability**
Event A	Observation P
Partition $\{A_1, \ldots, A_n\}$ of Ω	Ideal measurement $\{P_1, \ldots, P_n\}$
Probability law μ	Probability law ρ
Conditioned probability law $\tilde{\mu}$	Conditioned probability law $\tilde{\rho}$
Real random variable X	Hermitian operator A

In the abelian case, there is a one-to-one correspondence between the objects on the left-hand and right-hand side. In general, the objects on the right-hand side may be seen as a sort of generalization of those on the left-hand side.

The law ρ' from point 4° of our interpretation of quantum probability spaces does not have a classical counterpart. Indeed, if $\mathcal{A}$ is abelian and $\{P_1, \ldots, P_n\}$ is an ideal measurement, then $\rho'(A) := \sum_{i=1}^{n} \rho(P_i A P_i) = \rho(A)$. Thus, in classical probability, ideal measurements do not perturb the system they are measuring.

The states $\omega \in \Omega$ from classical probability do not have a quantum mechanical counterpart. Let us say that a probability law ρ on a Q-algebra $\mathcal{A}$ is a *precise state* if

$$\rho(P) \in \{0, 1\} \qquad \forall P \in \mathcal{A} \text{ such that } P \text{ is a projection.}$$

On an abelian Q-algebra $\mathbb{C}^\Omega$, it is easy to see that the precise states are exactly the probability laws of the form $\rho = \delta_\omega$, where

$$\delta_\omega(f) := f(\omega) \qquad (\omega \in \Omega),$$

and that every probability law on $\mathbb{C}^\Omega$ can in a unique way be written as a convex combination of these precise states. Thus, 'precise states' on an abelian Q-algebra correspond to the states ω from classical probability. We will later see that on a nonabelian Q-algebra, not every probability can be written as a convex combination of precise states. In fact, if $\mathcal{A} = \mathcal{L}(\mathcal{H})$ with $\dim(\mathcal{H}) \geq 2$, then there do not exist any precise states on $\mathcal{A}$ at all.

Chapter 3

Infinite dimensional spaces*

3.1 Measure theory*

In measure theory, it is custom to extend the real numbers by adding the points ∞ and $-\infty$, with which one calculates according to the rules

$$a \cdot \infty := \begin{cases} -\infty & \text{if } a < 0, \\ 0 & \text{if } a = 0, \\ \infty & \text{if } a > 0, \end{cases}$$

while $a + \infty := \infty$ if $a \neq -\infty$, and $\infty - \infty$ is not defined.

By definition, *measure space* is a triple $(\Omega, \mathcal{F}, \mu)$ with the following properties. 1° Ω is a set (possibly infinite). 2° $\mathcal{F} \subset \mathcal{P}(\Omega)$ is a subset of the set of all subsets of Ω with the following properties:

(a) $A_1, A_2, \ldots \in \mathcal{F} \;\Rightarrow\; \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$,
(b) $A \in \mathcal{F} \;\Rightarrow\; A^{\mathrm{c}} \in \mathcal{F}$,
(c) $\Omega \in \mathcal{F}$.

Such a $\mathcal{F}$ is called a *σ-algebra* or *σ-field*. 3° $\mu : \mathcal{F} \to [0, \infty]$ is a function such that

(a) $A_1, A_2, \ldots \in \mathcal{F},\ A_i \cap A_j = \emptyset\ \forall i \neq j \;\Rightarrow\; \mu\big(\bigcup_{i=1}^{\infty} A_i\big) = \sum_{i=1}^{\infty} \mu(A_i)$.

Such a function is called a *measure*. If

(b) $\mu(\Omega) = 1$,

then μ is called a *probability measure*. In this case $(\Omega, \mathcal{F}, \mu)$ is called a *probability space*. It is not hard to see that if Ω is a finite set and $\mathcal{F} = \mathcal{P}(\Omega)$, then we are back at our previous definition of a probability space.

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. By definition, a function $X : \Omega \to [-\infty, \infty]$ is *measurable* if

$$\{\omega : X(\omega) \leq a\} \in \mathcal{F} \qquad \forall a \in \mathbb{R}.$$

If X is nonnegative, then this is equivalent to the fact that X can be written as

$$X = \sum_{i=1}^{\infty} a_i 1_{A_i} \qquad (a_i \geq 0,\ A_i \in \mathcal{F}).$$

For such functions, one defines the integral as

$$\int X \mathrm{d}\mu := \sum_{i=1}^{\infty} a_i \mu(A_i).$$

One can show that this definition is unambiguous, i.e., does not depend on the choice of the a_i and A_i. If X is not nonnegative, then one puts $X = X^+ + X^-$ where X^+, X^- are nonnegative measurable functions and defines $\int X \mathrm{d}\mu := \int X^+ \mathrm{d}\mu - \int X^- \mathrm{d}\mu$. The integral of X is not defined if $\int X^+ \mathrm{d}\mu - \int X^- \mathrm{d}\mu$ happens to be $\infty - \infty$.

3.2 Metric and normed spaces*

Let E be a set. By definition, a *metric* on E is a function $d : E \times E \to [0, \infty)$ such that

(a)	$d(x, y) = d(y, x)$	$(x, y \in E)$,
(b)	$d(x, z) \leq d(x, y) + d(y, z)$	$(x, y, z \in E)$,
(c)	$d(x, y) = 0$ if and only if $x = y$	$(x, y \in E)$.

A *metric space* is a pair (E, d) where E is a set and d is a metric on E.

We say that sequence $x_n \in E$ converges to a limit x in the metric d, and write $x_n \to x$, if

$$\forall \varepsilon > 0\ \exists n \text{ s.t. } \forall m \geq n :\ d(x_n, x) \leq \varepsilon.$$

For any $D \subset E$, we call

$$\overline{D} := \{x \in E : \exists x_n \in D \text{ s.t. } x_n \to x\}$$

the *closure* of D. A subset $D \subset E$ is closed if $D = \overline{D}$. A subset $D \subset E$ is open if its complement D^{c} is closed. A subset $D \subset E$ is *dense* if $\overline{D} = E$. A metrix space is *separable* if there exists a countable dense set $D \subset E$. If E, F are metric spaces, then a function $f : E \to F$ is *continuous* if $f(x_n) \to f(x)$ whenever $x_n \to x$.

A *Cauchy sequence* is a sequence x_n such that

$$\forall \varepsilon > 0 \ \exists n \text{ s.t. } d(x_k, x_m) \leq \varepsilon \ \forall k, m \geq n.$$

A metric space is *complete* if every Cauchy sequence has a limit.

A metric space is *compact* if every sequence $x_n \in E$ has a convergent subsequence, i.e., there exist $m(n) \to \infty$ and $x \in E$ such that $x_{m(n)} \to x$.

Let $\mathcal{V}$ be a linear space (possibly infinite dimensional) over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. By definition, a norm on $\mathcal{V}$ is a function $\mathcal{V} \ni \phi \mapsto \|\phi\|$ from $\mathcal{V}$ into $[0, \infty)$ such that

$$\begin{array}{lll} \text{(a)} & \|a\phi\| = |a|\|\phi\| & (a \in \mathbb{K},\ \phi \in \mathcal{V}), \\ \text{(b)} & \|\phi + \psi\| \leq \|\phi\| + \|\psi\| & (\phi, \psi \in \mathcal{V}), \\ \text{(c)} & \|\phi\| = 0 \text{ implies } \phi = 0 & (\phi \in \mathcal{V}). \end{array}$$

A *normed space* is a pair $(\mathcal{V}, \|\cdot\|)$ where $\mathcal{V}$ is a linear space and $\|\cdot\|$ is a norm on $\mathcal{V}$. If $\|\cdot\|$ is a norm on $\mathcal{V}$, then

$$d(\phi, \psi) := \|\phi - \psi\|$$

defines a metric on $\mathcal{V}$, which is called the metric associated with $\|\cdot\|$. Two norms $\|\cdot\|$ and $\|\cdot\|'$ are called *equivalent* if there exists constants $0 < c < C$ such that

$$c\|\phi\| \leq \|\phi\|' \leq C\|\phi\| \qquad (\phi \in \mathcal{V}).$$

If $\|\cdot\|$ and $\|\cdot\|'$ are equivalent norms, then a sequence x_n converges in $\|\cdot\|$, or is a Cauchy sequence in $\|\cdot\|$, if and only if the corresponding property holds for $\|\cdot\|'$. Thus, concepts such as open, closed, complete, and compact do not depend on the choice of an equivalent metric.

If $\mathcal{H}$ is a linear space (possibly infinite dimensional) equipped with an inner product $\langle\cdot|\cdot\rangle$, then

$$\|\phi\| := \sqrt{\langle\phi|\phi\rangle} \qquad (\phi \in \mathcal{H})$$

defines a norm on $\mathcal{H}$, called the norm associated with the inner product.

Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Then the space $\mathbb{K}^n$ equipped with the inner product

$$\langle(\phi_1, \ldots, \phi_n)|(\phi_1, \ldots, \phi_n)\rangle := \sum_{i=1}^{n} \phi_i^* \psi_i$$

and the associated norm and metric, is complete and separable. In fact, all norms on $\mathbb{K}^n$ are equivalent and therefore $\mathbb{K}^n$ is complete and separable in any norm. A

subset D of $\mathbb{K}^n$ is compact if and only if it is closed and *bounded*, i.e., $\sup_{\phi\in D}\|\phi\| < \infty$.

In the infinite dimensional case, not all normed spaces are complete. A complete normed space is called a *Banach space*. A complete inner product space is called a *Hilbert space*.

Example I Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Let E be a compact metric space and let

$$\mathcal{C}(E) := \{f : E \to \mathbb{K} \,:\, f \text{ is continuous}\},$$

equipped with the *supremum norm*

$$\|f\| := \sup_{x\in E} |f(x)|.$$

Then $\mathcal{C}(E)$ is a Banach space.

Example II Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and

$$\mathcal{L}^2(\mu) := \{\phi : \Omega \to \mathbb{K} \,:\, \phi \text{ is measurable and } \int |\phi|^2 \,\mathrm{d}\mu < \infty\}.$$

Let $L^2(\mu)$ be the quotient space

$$L^2(\mu) := \mathcal{L}^2(\mu)/\mathcal{N}(\mu),$$

where $\mathcal{N}(\mu) := \{\phi \in \mathcal{L}^2(\mu) : \int |\phi|^2 \,\mathrm{d}\mu = 0\}$. Then $L^2(\mu)$, equipped with the inner product

$$\langle\phi|\psi\rangle := \int (\phi^*\psi)\mathrm{d}\mu$$

is a Hilbert space.

3.3 Hilbert spaces*

Recall that a Hilbert space is a complete inner product space. For any two Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$, a linear operator $A : \mathcal{H}_1 \to \mathcal{H}_2$ is continuous if and only if it is *bounded*, i.e.,

$$\|A\| := \sup_{\|\phi\|\le 1} \|A\phi\| < \infty.$$

We let $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ denote the Banach space of all *bounded linear operators* $A : \mathcal{H}_1 \to \mathcal{H}_2$, equipped with the *operator norm* $\|A\|$. Generalizing our earlier definition, we call the space of all bounded linear forms $\mathcal{H}' := \mathcal{L}(\mathcal{H}, \mathbb{K})$ the *dual* of $\mathcal{H}$. The *Riesz*

lemma says that the map $\phi \mapsto \langle\phi|$ is a colinear bijection from $\mathcal{H}$ to $\mathcal{H}'$, which preserves the norm. In particular

$$\mathcal{H}' = \{\langle\phi| : \phi \in \mathcal{H}\}.$$

If $\mathcal{H}_1, \mathcal{H}_2$ are Hilbert spaces with inner products $\langle\cdot|\cdot\rangle_1$ and $\langle\cdot|\cdot\rangle_2$, respectively, and $A \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, then there exists a unique adjoint $A^* \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ such that

$$\langle\phi|A\psi\rangle_2 = \langle A^*\phi|\psi\rangle_1 \qquad (\phi \in \mathcal{H}_2,\ \psi \in \mathcal{H}_1).$$

If $\mathcal{F} \subset \mathcal{H}$ is a *closed* linear subspace of $\mathcal{H}$, then each vector $\phi \in \mathcal{H}$ can in a unique way be written as

$$\phi = \phi' + \phi'' \qquad (\phi' \in \mathcal{F},\ \phi'' \in \mathcal{F}^\perp).$$

We call ϕ' the *orthogonal projection* of ϕ on the subspace $\mathcal{F}$, and write

$$\phi' =: P_\mathcal{F}\phi.$$

One can check that $P_\mathcal{F} \in \mathcal{L}(\mathcal{H}) := \mathcal{L}(\mathcal{H}, \mathcal{H})$ satisfies $P_\mathcal{F}^* = P_\mathcal{F} = P_\mathcal{F}^2$. Conversely, every $P \in \mathcal{L}(\mathcal{H}) := \mathcal{L}(\mathcal{H}, \mathcal{H})$ such that $P^* = P = P^2$ is of the form $P = P_\mathcal{F}$ for some closed subspace $\mathcal{F} \subset \mathcal{H}$.

The *spectrum* of a bounded linear operator $A \in \mathcal{L}(\mathcal{H})$ is defined as

$$\sigma(A) := \{\lambda \in \mathbb{K} : (\lambda - A) \text{ is not invertible}\}.$$

(Compare Exercise 1.1.2.) Warning: the spectrum is in general larger than the set of eigenvalues of A! One can show that $\sigma(A)$ is a compact[1] subset of $\mathbb{K}$. If $\mathbb{K} = \mathbb{C}$, then $\sigma(A)$ is nonempty.

There is also an analogue of Theorem 1.2.10. Indeed, if $A \in \mathcal{L}(\mathcal{H})$ is normal, i.e., $AA^* = A^*A$, then one can define a *spectral measure* $\mathcal{P}$ that assigns to each measurable subset $D \subset \mathbb{C}$ a projection operator $\mathcal{P}(D) \in \mathcal{L}(\mathcal{H})$. One can define integration with respect to the spectral measure, and give sense to the formula

$$A = \int_{\sigma(A)} \lambda \mathcal{P}(\mathrm{d}\lambda).$$

In fact, $\mathcal{P}$ is concentrated on $\sigma(A)$, so it makes no difference whether we integrate over $\sigma(A)$ or over $\mathbb{C}$. If $f : \mathbb{C} \to \mathbb{C}$ is a continuous function and $A \in \mathcal{L}(\mathcal{H})$ is a normal operator, then one defines a normal operator $f(A)$ by

$$f(A) := \int_{\sigma(A)} f(\lambda) \mathcal{P}(\mathrm{d}\lambda).$$

[1]There also exists mathematical theory for self-adjoint operators with a non-compact spectrum, but such operators are unbounded and only defined on a dense subset of $\mathcal{H}$. Many important observables in quantum mechanics, such as those for the position and momentum of a particle, are described by unbounded self-adjoint operators.

3.4 C∗-algebras*

By definition, a *C∗-algebra* is a (possibly infinite dimensional) complex ∗-algebra $\mathcal{A}$ equipped with a norm $\|\cdot\|$ such that, in addition to the properties (i)–(vii) from Section 2.1,[2]

$$\begin{array}{rl} \text{(viii)}' & \mathcal{A} \text{ is complete in the norm } \|\cdot\| \\ \text{(ix)}' & \|AB\| \leq \|A\|\|B\| \qquad (A, B \in \mathcal{A}) \\ \text{(x)}' & \|A^*A\| = \|A\|^2 \end{array}$$

Note that property (x)' implies property (viii) from Section 2.1, so finite dimensional C∗-algebras are Q-algebras. Conversely, every Q-algebra can in a unique way be equipped with a norm $\|\cdot\|$ such that (viii)'–(x)' hold.[3]

If $\mathcal{H}$ is a Hilbert space, then the space $\mathcal{L}(\mathcal{H})$ of bounded linear operators on $\mathcal{H}$, equipped with the operator product, adjoint, and norm, is a C∗-algebra. In analogy with Theorem 2.1.5 one has the following theorem about representations of C∗-algebras.

Theorem 3.4.1 (Gelfand-Naimark) *Let $\mathcal{A}$ be a C∗-algebra. Then there exists a Hilbert space $\mathcal{H}$ and a sub-∗-algebra $\mathcal{A}'$ of $\mathcal{L}(\mathcal{H})$ such that $\mathcal{A}$ is isomorphic to $\mathcal{A}'$. If $\mathcal{A}$ is separable then we may take $\mathcal{H}$ separable.*

Probability laws on C∗-algebras are defined exactly as in the finite dimensional case. We can therefore define an infinite dimensional quantum probability space as a pair $(\mathcal{A}, \rho)$ where $\mathcal{A}$ is a C∗-algebra and ρ is a probability on $\mathcal{A}$.

Let E be a compact metric space and let $\mathcal{C}(E) := \{f : E \to \mathbb{C} \text{ continuous}\}$, equipped with the supremum norm. We equip $\mathcal{C}(E)$ with the structure of a ∗-algebra by putting $fg(x) := f(x)g(x)$ and $f^*(x) := f(x)^*$. Then $\mathcal{C}(E)$ is a separable abelian C∗-algebra. The following infinite dimensional analogue of Theorem 2.4.1 says that conversely, every separable abelian C∗-algebra arises in this way.

Theorem 3.4.2 (Abelian C∗-algebras) *Let $\mathcal{A}$ be a separable abelian C∗-algebra. Then there exists a compact metric space E such that $\mathcal{A}$ is isomorphic to $\mathcal{C}(E)$.*

[2] Here, we only consider C∗-algebras which contain a unit element.

[3] This can be proved using Theorem 2.1.5. I do not know of any way to prove the statement without making use of Theorem 2.1.5. If one could find such a proof, then Theorem 2.1.5 would follow from Theorem 3.4.1 by specializing to the finite-dimensional case. The only proof of Theorem 2.1.5 known to me is quite different from the proof of Theorem 3.4.1.

It can moreover be proved that if μ is a probability measure on E, equipped with the σ-field generated by the open sets, then

$$\rho(f) := \int f \, \mathrm{d}\mu$$

defines a probability law ρ on the C∗-algebra $\mathcal{C}(E)$, and conversely, every probability law on $\mathcal{C}(E)$ arises in this way. Thus, abelian quantum probability spaces correspond to classical probability spaces. (The facts that $\mathcal{A}$ is separable and E is a compact metric space are not really restrictions. In fact, in quantum probability, it is standard to assume that the C∗-algebra is separable, while all interesting models of classical probability can be constructed with probabilities defined on compact metric spaces.)

Chapter 4

Some quantum mechanics

4.1 States

So far, probability laws on Q-algebras have been defined abstractly, as functions $\rho : \mathcal{A} \to \mathbb{C}$ having certain properties. In practice, we usually work with a concrete representation of $\mathcal{A}$ as a sub-$*$-algebra of the algebra $\mathcal{L}(\mathcal{H})$ of all linear operators on some (complex) inner product space $\mathcal{H}$. We need a practical way of constructing probability laws on such a Q-algebra. To prepare for this, we need some definitions.

Let $\mathcal{A}$ be any (abstract) Q-algebra. By definition, a *positive linear form* is a map $\rho : \mathcal{A} \to \mathbb{C}$ that is (a) *linear*, (b) *real*, and (c) *positive*, i.e.,

$$\begin{aligned} &\text{(a)} \quad \rho(aA + bB) = a\rho(A) + b\rho(B) && (a, b \in \mathbb{C},\ A, B \in \mathcal{A}), \\ &\text{(b)} \quad \rho(A^*) = \rho(A)^* && (A \in \mathcal{A}), \\ &\text{(c)} \quad \rho(A^*A) \geq 0 && (A \in \mathcal{A}). \end{aligned}$$

Note that probability laws are normalized positive linear forms. A positive linear form is called *faithful* if in addition

$$\text{(d)} \quad \rho(A^*A) = 0 \text{ implies } A = 0.$$

If ρ is a faithful positive linear form on $\mathcal{A}$, then

$$\langle A|B\rangle_\rho := \rho(A^*B) \qquad (A, B \in \mathcal{A}) \tag{4.1}$$

defines an inner product on $\mathcal{A}$. A positive linear form τ is called a *pseudotrace* if

$$\tau(AB) = \tau(BA) \qquad (A, B \in \mathcal{A}).$$

If $\mathcal{L}(\mathcal{H})$ is the algebra of all linear operators on some (complex) inner product space $\mathcal{H}$, and $\mathcal{A}$ is a sub-$*$-algebra of $\mathcal{L}(\mathcal{H})$, then the usual trace $\mathrm{tr}(A)$ is a faithful pseudotrace on $\mathcal{A}$. Thus, by Theorem 2.1.5, on each Q-algebra there exists at least one faithful pseudotrace.[1] In Exercise 5.5.3 below, we will see that on algebras of the form $\mathcal{L}(\mathcal{H})$, all pseudotraces are constant multiples of the usual trace.

The next proposition says that if $\mathcal{A}$ is a Q-algebra with a concrete representation on an inner product space $\mathcal{H}$, then every probability law ρ has a *density* (or *density operator*) R with respect to the trace.[2]

Proposition 4.1.1 (Density operator) *Let $\mathcal{L}(\mathcal{H})$ be the algebra of all linear operators on an inner product space $\mathcal{H}$ and let $\mathcal{A}$ be a sub-$*$-algebra of $\mathcal{L}(\mathcal{H})$. Let $R \in \mathcal{A}$ be positive hermitian such that $\mathrm{tr}(R) = 1$. Then the formula*

$$\rho(A) := \mathrm{tr}(RA) \qquad (A \in \mathcal{A})$$

defines a probability law on $\mathcal{A}$. Conversely, every probability law on $\mathcal{A}$ arises in this way and R is uniquely determined by ρ.

Proof It is easy to check that the formula $\rho(A) := \mathrm{tr}(RA)$ defines a probability law.[3] To prove that every probability law arises in this way, we use that $\langle A|B\rangle := \mathrm{tr}(A^*B)$ defines an inner product on $\mathcal{A}$. Therefore, since a probability law ρ is a linear form on $\mathcal{A}$, there exists a unique $R \in \mathcal{A}$ such that

$$\rho(A) = \langle R|A\rangle = \mathrm{tr}(R^*A) \qquad (A \in \mathcal{A}).$$

Since ρ is real,

$$\mathrm{tr}(R^*A^*) = \rho(A^*) = \rho(A)^* = \mathrm{tr}(R^*A)^* = \mathrm{tr}(A^*R) = \mathrm{tr}(RA^*).$$

Since this holds for all $A \in \mathcal{A}$, we must have $R^* = R$, i.e., R is hermitian. By Lemma 2.1.6 we can write $R = \sum_{i=1}^n \lambda_i P_i$ where $\{P_1, \ldots, P_n\}$ is a partition of the identity with $P_i \in \mathcal{A}$ for each i; assume that one of the eigenvalues λ_i is

[1] In fact, it can be proved that property (viii) in the definition of a Q-algebra is equivalent to the existence of a faithful pseudotrace; see Exercise 5.3.1.

[2] More generally, if $\mathcal{A}$ is an abstract Q-algebra and τ is a faithful pseudotrace on $\mathcal{A}$, then every probability law ρ on $\mathcal{A}$ has a density R with respect to τ. Since there exists, in general, more than one pseudotrace on a given Q-algebra, the density R will depend on the choice of the pseudotrace.

[3] To check property (c), we use the functional calculus for normal operators to define $\sqrt{R}$ and write $\rho(A^*A) = \mathrm{tr}(RA^*A) = \mathrm{tr}(\sqrt{R}A^*A\sqrt{R}) = \mathrm{tr}((A\sqrt{R})^*A\sqrt{R}) \geq 0$.

strictly negative. Then $\rho(P_i) = \mathrm{tr}((\sum_j \lambda_j P_j)P_i) = \lambda_j \mathrm{tr}(P_i^2) < 0)$, which gives a contradiction. Thus R must be positive. ∎

Let $\mathcal{A}$ be a Q-algebra. In quantum mechanics, it is a (bad) tradition to call a probability law ρ on $\mathcal{A}$ a *state*. Note that the set of all probability laws is a convex subset of the space of all real linear forms, i.e., if $\rho_1, \ldots, \rho_n$ are probabilities and $p_1, \ldots, p_n \geq 0$ with $\sum_i p_i = 1$, then

$$\rho := \sum_i p_i \rho_i$$

is a probability law on $\mathcal{A}$. By definition, a *pure state* is an *extremal element* of the convex set of all probability laws, i.e., a pure state is a probability law ρ that is not a nontrivial convex combination of other states. It suffices to check that ρ cannot be written as a nontrivial convex combination of two states, i.e., ρ is pure if it is not possible to write $\rho = p\rho_1 + (1-p)\rho_2$ with $0 < p < 1$ and $\rho_1 \neq \rho_2$. A probability law that is not a pure state is called a *mixed state*. Note that if a probability law ρ can be written as a nontrivial convex combination of two or more other states as $\rho = \sum_i p_i \rho_i$, then we can interpret this by saying that with probability p_i, the "true" state is ρ_i. Thus, pure states represent in some sense maximal knowledge about the system.

We will investigate the shape of the convex set of all probability laws on a Q-algebra and its extremal elements first in two special cases: for abelian algebras (which correspond to classical probability theory) and for algebras of the form $\mathcal{L}(\mathcal{H})$ (which is typically what one has in quantum mechanics). Finally, we will look at general Q-algebras.

By Theorem 2.4.1, every abelian Q-algebra $\mathcal{A}$ is of the form $\mathcal{A} \cong \mathbb{C}^\Omega$ for some finite set Ω, and the space of all probability laws on $\mathcal{A}$ corresponds to the space of all probability measures on Ω. We claim that:

(i) The pure states on $\mathbb{C}^\Omega$ are the probability laws of the form δ_ω with $\omega \in \Omega$.

(ii) Each probability law ρ on $\mathbb{C}^\Omega$ can in a unique way be written as a convex combination of pure states.

Here we recall from Section 2.4 that $\delta_\omega(A) := 1_{\{\omega \in A\}}$ $(A \subset \Omega)$ denotes the delta-measure at Ω. The proof of part (ii) is easy, since any probability measure ρ on Ω can uniquely be written as

$$\rho = \sum_{\omega \in \Omega} \rho\big(\{\Omega\}\big)\delta_\omega.$$

Also, if $\delta_\omega = p\rho_1 + (1-p)\rho_2$ with $0 < p < 1$, then ρ_1 and ρ_2 must both give the event $\{\Omega\}$ probability one and hence $\rho_1 = \rho_2 = \delta_\omega$, proving that the delta-measures are pure states.
Now let $\mathcal{H}$ be a complex inner product space and let $\mathcal{A} = \mathcal{L}(\mathcal{H})$. We claim that:

(i) The pure states on $\mathcal{L}(\mathcal{H})$ are the probability laws of the form $\rho_\psi(A) := \langle\psi|A|\psi\rangle$ where $\psi \in \mathcal{H}$ satisfies $\|\psi\| = 1$.

(ii) Each probability law ρ on $\mathcal{L}(\mathcal{H})$ can be written as a convex combination of pure states, but this decomposition is in general not unique.

We note that if $\psi \in \mathcal{H}$ is a vector of norm one, then $\rho_\psi(A) = \mathrm{tr}(|\psi\rangle\langle\psi|A)$ so the density operator of ρ_ψ is $R = |\psi\rangle\langle\psi|$. By Proposition 4.1.1, each state ρ is of the form $\rho(A) = \mathrm{tr}(RA)$ where the density operator R is a positive hermitian operator with trace one. Each such density operator can be written as $R = \sum_i p_i |e(i)\rangle\langle e(i)|$ where $\{e(1), \ldots, e(n)\}$ is an orthonormal basis of $\mathcal{H}$ and $p_i \geq 0$ sum up to one. It follows that $\rho = \sum_i p_i \rho_{e(i)}$, proving part (ii). In general, such a decomposition is not unique. For example, if $R = n^{-1}1$, where 1 denotes the identity operator, then $\rho = n^{-1}\sum_i \rho_{e(i)}$ for *any* orthonormal basis. Also, if $\phi(1), \ldots, \phi(n)$ are vectors of norm one, but not necessarily orthogonal, and $p_1, \ldots, p_n \geq 0$ sum up to one, then $\rho := \sum_i p_i \rho_{\phi(i)}$ is a state, so not every decomposition into states of the form ρ_ϕ corresponds to an orthonormal basis.
To see that states of the form ρ_ψ are pure, assume that $\rho_\psi = p\rho_1 + (1-p)\rho_2$ with $0 < p < 1$. Since $\rho_\psi\big(|\psi\rangle\langle\psi|\big) = 1$, the same must be true for ρ_1. By what we have already proved, we can decompose ρ_1 into pure states, i.e., we can write $\rho_1 := \sum_i p_i \rho_{\phi(i)}$ with $p_i > 0$ and $\phi(i)$ vectors of norm one. Then $\rho_{\phi(i)}\big(|\psi\rangle\langle\psi|\big) = 1$ implies $|\langle\psi|\phi(i)\rangle|^2 = 1$ and hence each $\phi(i)$ must be of the form $\phi(i) = \lambda\psi$ with $|\lambda| = 1$. This implies that $\rho_1 = \rho$. By the same argument, $\rho_2 = \rho$, so ρ is pure.

We now treat the case that $\mathcal{A}$ is a general Q-algebra, which contains the previous two situations as special cases. By definition, a *minimal projection* in some Q-algebra $\mathcal{A}$ is a projection $P \in \mathcal{A}$ such that $P \neq 0$ and the only projections Q with $Q \leq P$ are $Q = 0$ and $Q = P$. By definition, a *maximally fine partition of the identity* is a partition of the identity that consists of minimal projections.

The following proposition identifies minimal projections with pure states and says, among other things, that every state can be written as a convex combination of pure states. This decomposition is in general not unique!

Proposition 4.1.2 (Pure states and minimal projections) *If P is a minimal projection in a Q-algebra $\mathcal{A}$ then there exists a pure state ρ_P on $\mathcal{A}$ such that*

$$PAP = \rho_P(A)P \qquad (A \in \mathcal{A}).$$

Conversely, every pure state is of this form. Every state ρ on $\mathcal{A}$ can be written as

$$\rho(A) = \sum_{j=1}^{n} p_j \rho_{P_j}$$

where $\{P_1, \ldots, P_n\}$ is a maximally fine partition of the identity and the p_j are nonnegative numbers, summing up to one.

To prepare for the proof of Proposition 4.1.2, we need one preparatory lemma.

Lemma 4.1.3 (Existence of nontrivial projections) *Let $\mathcal{H}$ be an inner product space and let $\mathcal{A} \subset \mathcal{L}(\mathcal{H})$ be a sub-$*$-algebra of $\mathcal{L}(\mathcal{H})$. Assume that 1 and 0 are the only projections in $\mathcal{A}$. Then $\mathcal{A} = \{\lambda 1 : \lambda \in \mathbb{C}\}$.*

Proof Imagine that $\mathcal{A}$ contains some element A that is not a constant multiple of the identity operator. By Exercise 2.1.2, we may write $A = \mathrm{Re}(A) + i\mathrm{Im}(A)$, where $\mathrm{Re}(A)$ and $\mathrm{Im}(A)$ are hermitian operators. Since A is not a constant multiple of the identity operator, either $\mathrm{Re}(A)$ or $\mathrm{Im}(A)$ must be not a constant multiple of the identity operator, so $\mathcal{A}$ contains some hermitian operator, say B, that is not a constant multiple of the identity operator. By Lemma 2.1.6, we can write $B = \sum_{\lambda \in \sigma(B)} \lambda P_\lambda$ where $\{P_\lambda : \lambda \in \sigma(B)\}$ is a partition of the identity and $P_\lambda \in \mathcal{A}$ for each $\lambda \in \sigma(B)$. Since B is not a constant multiple of the identity operator, $\sigma(B)$ must have more than one element, so $\mathcal{A}$ must contain a projection that is different from 1 or 0. ∎

Proof of Proposition 4.1.2 By Theorem 2.1.5, we may without loss of generality assume that there exists an inner product space $\mathcal{H}$ such that $\mathcal{A}$ is a sub-$*$-algebra of $\mathcal{L}(\mathcal{H})$. Let $P \in \mathcal{A}$ be a minimal projection and let $\mathcal{F} \subset \mathcal{H}$ be the space that P projects on. Consider the space

$$\mathcal{B} := \{PAP : A \in \mathcal{A}\}.$$

An element PAP of $\mathcal{B}$ maps the space $\mathcal{F}$ into itself, so we may view such an operator as an element of $\mathcal{L}(\mathcal{F})$. If we do this, then we claim that $\mathcal{B}$ is a sub-$*$-algebra of $\mathcal{L}(\mathcal{F})$. Indeed, $(PAP)(PBP) = P(APB)P \in \mathcal{B}$ and $(PAP)^* = PA^*P \in \mathcal{B}$ for all $PAP, PBP \in \mathcal{B}$, and $P1P = P$ is the identity operator in $\mathcal{L}(\mathcal{F})$. We claim that each element of $\mathcal{B}$ is a constant multiple of P. Indeed, if this is not the case, then by Lemma 4.1.3, there exists a projection in $\mathcal{B}$ that is neither 0 nor the identity on $\mathcal{L}(\mathcal{F})$, which is P. Thus, there exists some $A \in \mathcal{A}$ such that $Q \in PAP$ is a projection operator in $\mathcal{L}(\mathcal{F})$ with $Q \neq 0, P$. But then Q is a projection operator in $\mathcal{A}$ with $Q \leq P$, contradicting the minimality of P.

It follows that for each minimal projection $P \in \mathcal{A}$, there exists a function $\rho_P : \mathcal{A} \to \mathbb{C}$ such that

$$PAP = \rho_P(A)P \qquad (A \in \mathcal{A}).$$

We claim that ρ_P is a state. The facts that ρ_P is linear, real, and normalized such that $\rho_P(1) = 1$ are straightforward. To see that ρ_P is positive, we write $(AP)^*(AP) = P(A^*A)P = \rho_P(A^*A)P$. By Exercise 1.2.16, $(AP)^*(AP)$ is a positive operator, so we must have $\rho_P(A^*A) \geq 0$.

We now claim that each state ρ on $\mathcal{A}$ can be written as a convex combination of states of the form ρ_P. Let R be the density operator of ρ. By Lemma 2.1.6, we can write $R = \sum_{\lambda \in \sigma(R)} \lambda P_\lambda$ where $\{P_\lambda : \lambda \in \sigma(R)\}$ is a partition of the identity and $P_\lambda \in \mathcal{A}$ for each $\lambda \in \sigma(R)$. If some of the projections P_λ is not minimal, then we can find some projection $Q \neq 0, P_\lambda$ with $Q \leq P_\lambda$. Then also $P_\lambda - Q \leq P_\lambda$ is a projection and we can refine our partition of the identity replacing P_λ by Q and $P_\lambda - Q$. Continuing in this way, we can find a partition of the identity $\{P_1, \ldots, P_n\}$ such that each P_i is a minimal projection in $\mathcal{A}$ and R can be written as

$$R = \sum_{i=1}^n q_i P_i,$$

where the q_i are nonnegative real numbers (since R is a positive operator) of which some may occur more than once. Then

$$\begin{aligned}\rho(A) = \mathrm{tr}(RA) &= \sum_{i=1}^n q_i \mathrm{tr}(P_i A) = \sum_{i=1}^n q_i \mathrm{tr}(P_i A P_i) \\ &= \sum_{i=1}^n q_i \mathrm{tr}(P_i)\rho_{P_i}(A) = \sum_{i=1}^n p_i \rho_{P_i}(A) \qquad (A \in \mathcal{A}),\end{aligned}$$

where we have defined $p_i := q_i \mathrm{tr}(P_i)$, which are nonnegative constants that must obviously sum up to one in order to have $\rho(1) = 1$.

To complete the proof, we must show that states of the form ρ_P are extremal. We have already shown that each state on $\mathcal{A}$ can be written as a convex combination of states of the form ρ_P. In view of this, it suffices to prove the following statement. Let P and $P_1, \ldots, P_n$ be minimal projections and $p_1, \ldots, p_n$ be strictly positive constants such that

$$\rho_P = \sum_{i=1}^n p_i \rho_{P_i}.$$

Then $P_i = P$ for each $i = 1, \ldots, n$. Since

$$0 = P(1-P)P = \rho_P(1-P)P = \sum_{i=1}^{n} p_i \rho_{P_i}(1-P)P,$$

we see that the nonnegative numbers $\rho_{P_i}(1-P)$ must all be zero and hence $P_i(1-P)P_i = 0$ for each $i = 1, \ldots, n$. Since $P_i(1-P)P_i = ((1-P)P_i)^*((1-P)P_i)$, it follows that $(1-P)P_i = 0$ and hence by Exercise 1.2.18 $P_i \leq P$. Since P is minimal, we conclude that $P_i = P$. ∎

Exercise 4.1.4 Let $\mathcal{H}$ be an inner product space and let $\mathcal{A} \subset \mathcal{L}(\mathcal{H})$ be a sub-$*$-algebra of $\mathcal{L}(\mathcal{H})$. Let $P_{\mathcal{F}} = P \in \mathcal{A}$ be a mininal projection, where $\mathcal{F} \subset \mathcal{H}$ is the subspace that $P_{\mathcal{F}}$ projects on. Show that the density opererator R of the pure state $\rho_{P_{\mathcal{F}}}$ is given by

$$R = \frac{1}{\dim(\mathcal{F})} P_{\mathcal{F}}.$$

Exercise 4.1.5 Let $\mathcal{A}$ be a Q-algebra. Show that every real linear form ρ on $\mathcal{A}$ can be written as $\rho = \rho_+ - \rho_-$, where ρ_+, ρ_- are positive linear forms. Show that every linear form l on $\mathcal{A}$ be written as $\mathrm{Re}(l) + i\mathrm{Im}(l)$, where $\mathrm{Re}(l), \mathrm{Im}(l)$ are real linear forms.

Exercise 4.1.6 Show that the pure states on a Q-algebra $\mathcal{A}$ span the space of all linear forms on $\mathcal{A}$.

Lemma 4.1.7 (Pure states on $\mathcal{L}(\mathcal{H})$) *Let $\mathcal{H}$ be an inner product space. Then ρ is a pure state on $\mathcal{L}(\mathcal{H})$ if and only if there exists a vector $\psi \in \mathcal{H}$ with $\|\psi\| = 1$ such that*

$$\rho(A) = \rho_\psi(A) := \langle \psi | A | \psi \rangle \qquad (A \in \mathcal{L}(\mathcal{H})).$$

For any state ρ on $\mathcal{L}(\mathcal{H})$ there exists an orthonormal basis $\{e(1), \ldots, e(n)\}$ and nonnegative numbers $p_1, \ldots, p_n$, summing up to one, such that

$$\rho(A) = \sum_i p_i \langle e(i) | A | e(i) \rangle \qquad (A \in \mathcal{L}(\mathcal{H})).$$

Proof The minimal projections on $\mathcal{L}(\mathcal{H})$ are those that project on a one-dimensional subspace, i.e., those of the form $P = |\psi\rangle\langle\psi|$ where $\psi \in \mathcal{H}$ has norm $\|\psi\| = 1$, and maximally fine partitions of the identity consist of projections $P_i = |e(i)\rangle\langle e(i)|$ that project on an orthonormal basis $\{e(1), \ldots, e(n)\}$. Now if $P = |\psi\rangle\langle\psi|$, then

$$\rho_P(A)P = PAP = |\psi\rangle\langle\psi|A|\psi\rangle\langle\psi| = \langle\psi|A|\psi\rangle|\psi\rangle\langle\psi|,$$

which shows that

$$\rho_P(A) = \langle\psi|A|\psi\rangle.$$

The lemma is therefore just a special case of Proposition 4.1.2. ∎

In the special case that our Q-algebra is of the form $\mathcal{L}(\mathcal{H})$, Lemma 4.1.7 shows that every *state vector* $\psi \in \mathcal{H}$ with $\|\psi\| = 1$ defines a pure state ρ_ψ, and every pure state is of this form. This correspondence is almost one-to-one, except that the state vectors

$$\psi \quad \text{and} \quad e^{i\alpha}\psi \qquad (\alpha \in [0, 2\pi)),$$

differing only by the *phase factor* $e^{i\alpha}$ describe the same pure state. Note that by Exersice 4.1.8 below, two states ρ_1, ρ_2 are equal if and only if they give the same probability to every observation (projection) P. Thus, there is no 'redundant information' in states ρ.

State vectors were invented earlier than Q-algebras or C∗-algebras. The celebrated *Copenhagen interpretation* of quantum mechanics says that the state of a quantum mechanical system is described by a unit vector ψ in a Hilbert space $\mathcal{H}$. Real observables correspond to self-adjoint operators A. An observable A can assume values in its spectrum $\sigma(A)$. Let $\mathcal{P}$ be the spectral measure associated with A; in the finite-dimensional case, this means that $\mathcal{P}(D)$ is the orthogonal projection on the space spanned by all eigenvectors with eigenvalues in a set $D \subset \mathbb{R}$. Then

$$\|\mathcal{P}(D)\psi\|^2 = \langle\mathcal{P}(D)\psi|\mathcal{P}(D)\psi\rangle = \langle\psi|\mathcal{P}(D)|\psi\rangle = \rho_\psi(\mathcal{P}(D))$$

is the probability that an ideal measurement of A yields a value in D. Given that we do such an observation, we must describe our sytem with the new state

$$\tilde{\rho}_\psi(A) = \frac{\rho_\psi(\mathcal{P}(D)A\mathcal{P}(D))}{\rho_\psi(\mathcal{P}(D))} = \frac{\langle\mathcal{P}(D)\psi|A|\mathcal{P}(D)\psi\rangle}{\|\mathcal{P}(D)\psi\|^2} = \rho_{\tilde{\psi}}(A),$$

where $\tilde{\psi}$ is the unit vector defined by

$$\tilde{\psi} := \tfrac{1}{\|\mathcal{P}(D)\psi\|}\mathcal{P}(D)\psi.$$

This recipe for conditioning a pure state is known as the *projection postulate* and has been the subject of much discussion. It is worth mentioning that although historically, the word projection postulate refers to the map $\rho \mapsto \tilde{\rho}$ from our interpretation of quantum probability spaces, in modern discussions of this topic, the same term is sometimes used to refer to the map $\rho \mapsto \rho'$ from our interpretation.

Exercise 4.1.8 Show that the projections in a Q-algebra $\mathcal{A}$ span the whole algebra $\mathcal{A}$. (Hint: Exercise 2.1.2.)

Exercise 4.1.9 Let $\mathcal{A}$ be a Q-algebra and let ρ_1, ρ_2 be states on $\mathcal{A}$. Show that $\rho_1(P) = \rho_2(P)$ for all projections $P \in \mathcal{A}$ if and only if $\rho_1 = \rho_2$.

If $\mathcal{A}$ is abelian, then it is easy to see that a state ρ is pure if under ρ, each projection P has either probability zero or one. The next exercise shows that in the nonabelian case, the situation is quite different.

Exercise 4.1.10 (Unprecise states) If $\dim(\mathcal{H}) \geq 2$, then for every state ρ there exists a projection $P \in \mathcal{L}(\mathcal{H})$ such that $0 < \rho(P) < 1$.

4.2 The Bloch sphere

If Ω is a set with $|\Omega| = 4$ elements, then the space of all probability laws on Ω is a three-dimensional convex set with four extremal elements. In the present section, we investigate the shape of the convex set of all states on an algebra of the form $\mathcal{L}(\mathcal{H})$, where $\dim(\mathcal{H}) = 2$. We will see that this set has the form of a three-dimensional ball, with the surface corresponding to the extremal elements, i.e., pure states. We start with some lemmas.

Lemma 4.2.1 (Basis for hermitian operators) *Let $\mathcal{H}$ be an inner product space with $\dim(\mathcal{H}) = 2$ and a given orthonormal basis $\{e(1), e(2)\}$. With respect to this basis, let $S_{\rm x}, S_{\rm y}$, and $S_{\rm z}$ be given by the Pauli matrices introduced in Section 2.3. Then the matrices $\{1, S_{\rm x}, S_{\rm y}, S_{\rm z}\}$ form a basis for the real vector space $\mathcal{A}_{\rm herm} := \{A \in \mathcal{L}(\mathcal{H}) : A^* = A\}$ of hermitian operators on $\mathcal{H}$. This basis is orthonormal with respect to the inner product*

$$\langle A|B\rangle := \tfrac{1}{2}{\rm tr}(AB) \qquad (A, B \in \mathcal{A}_{\rm herm}).$$

Proof Clearly, each hermitian operator A can uniquely be written as

$$A = \begin{pmatrix} r + \theta_{\rm z} & \theta_{\rm x} - i\theta_{\rm y} \\ \theta_{\rm x} + i\theta_{\rm y} & r - \theta_{\rm z} \end{pmatrix}$$

for some $(r, \theta_{\rm x}, \theta_{\rm y}, \theta_{\rm z}) \in \mathbb{R}^4$. This shows that $\{1, S_{\rm x}, S_{\rm y}, S_{\rm z}\}$ is a basis of the real linear space $\{A \in \mathcal{L}(\mathcal{H}) : A^* = A\}$. It is straightforward to check that

$$\begin{aligned} &{\rm tr}(S_{\rm x}) = 0, \quad {\rm tr}(S_{\rm y}) = 0, \quad {\rm tr}(S_{\rm z}) = 0, \\ &{\rm tr}(S_{\rm x}S_{\rm y}) = 0, \quad {\rm tr}(S_{\rm x}S_{\rm z}) = 0, \quad {\rm tr}(S_{\rm y}S_{\rm z}) = 0. \end{aligned}$$

On the other hand, $S_{\rm x}^2 = S_{\rm y}^2 = S_{\rm z}^2 = 1$ and $\mathrm{tr}(1) = 2$, showing that $\{1, S_{\rm x}, S_{\rm y}, S_{\rm z}\}$ is orthonormal with respect to the inner product $\langle A|B\rangle := \frac{1}{2}\mathrm{tr}(AB)$. ∎

Lemma 4.2.2 (Proper projections) *Let $\mathcal{H}$ be an inner product space with $\dim(\mathcal{H}) = 2$ and a given orthonormal basis $\{e(1), e(2)\}$. With respect to this basis, let $S_{\rm x}, S_{\rm y}$, and $S_{\rm z}$ be given by the Pauli matrices introduced in Section 2.3. Then an operator $P \in \mathcal{L}(\mathcal{H})$ is a projector on a one-dimensional subspace if and only if*

$$P = \tfrac{1}{2}1 + \tfrac{1}{2}\theta_{\rm x}S_{\rm x} + \tfrac{1}{2}\theta_{\rm y}S_{\rm y} + \tfrac{1}{2}\theta_{\rm z}S_{\rm z} =: P_\theta \tag{4.2}$$

for some $\theta = (\theta_{\rm x}, \theta_{\rm y}, \theta_{\rm z}) \in \mathbb{R}^3$ with $\|\theta\| = 1$.

Proof Since $\dim(\mathcal{H}) = 2$, an operator $P \in \mathcal{L}(\mathcal{H})$ is a projector on a one-dimensional subspace if and only if its spectrum is $\sigma(P) = \{0, 1\}$. We first calculate the spectrum of a hermitian operator of the form

$$\theta_{\rm x}S_{\rm x} + \theta_{\rm y}S_{\rm y} + \theta_{\rm z}S_{\rm z} = \begin{pmatrix} \theta_{\rm z} & \theta_{\rm x} - i\theta_{\rm y} \\ \theta_{\rm x} + i\theta_{\rm y} & -\theta_{\rm z} \end{pmatrix}$$

To find its eigenvalues, we must solve $\det(\lambda - A) = 0$, which gives

$$\begin{aligned} &(\lambda - \theta_{\rm z})(\lambda + \theta_{\rm z}) - (\theta_{\rm x} - i\theta_{\rm y})(\theta_{\rm x} + i\theta_{\rm y}) = 0 \\ \Leftrightarrow \quad &\lambda^2 - \theta_{\rm z}^2 - \theta_{\rm x}^2 - \theta_{\rm y}^2 = 0 \quad \Leftrightarrow \quad \lambda = \pm\|\theta\|. \end{aligned}$$

It follows that the spectrum of a general hermitian operator is given by

$$\sigma\big(r1 + \theta_{\rm x}S_{\rm x} + \theta_{\rm y}S_{\rm y} + \theta_{\rm z}S_{\rm z}\big) = \{r + \|\theta\|, r - \|\theta\|\}.$$

In particular, the spectrum of such an operator is $\{0, 1\}$ if and only if it is of the form P_θ in (4.2) with $\|\theta\| = 1$. ∎

Proposition 4.2.3 (Convex set of all states) *Let $\mathcal{H}$ be an inner product space with $\dim(\mathcal{H}) = 2$ and a given orthonormal basis $\{e(1), e(2)\}$. With respect to this basis, let $S_{\rm x}, S_{\rm y}$, and $S_{\rm z}$ be given by the Pauli matrices introduced in Section 2.3. Let $\mathcal{R}_{\rm real}$ be the real vector space consisting of all real linear forms on $\mathcal{L}(\mathcal{H})$. Then the map*

$$\mathcal{R}_{\rm real} \ni \rho \mapsto \big(\rho(1), \rho(S_{\rm x}), \rho(S_{\rm y}), \rho(S_{\rm z})\big) \in \mathbb{R}^4 \tag{4.3}$$

is a bijection. This bijection maps the convex set $\mathcal{R}_{\rm state}$ of all states on $\mathcal{L}(\mathcal{H})$ into the convex set

$$\{(1, \theta_{\rm x}, \theta_{\rm y}, \theta_{\rm z}) \in \mathbb{R}^4 : \|\theta\| \leq 1\}.$$

In particular, for each $\theta \in \mathbb{R}^3$ with $\|\theta\| = 1$, there exists a unique pure state $\rho_{(\theta)}$ on $\mathcal{L}(\mathcal{H})$ such that

$$\big(\rho_{(\theta)}(S_{\rm x}), \rho_{(\theta)}(S_{\rm y}), \rho_{(\theta)}(S_{\rm z})\big) = \theta. \tag{4.4}$$

Proof By Exercise 2.1.2, a linear form ρ on $\mathcal{L}(\mathcal{H})$ is uniquely characterized by the linear function $\mathcal{A}_{\rm herm} \ni A \mapsto \rho(A) \in \mathbb{C}^4$. Moreover, ρ is real if and only if this functions maps $\mathcal{A}_{\rm herm}$ into $\mathbb{R}^4$. By Lemma 4.2.1, the matrices $\{1, S_{\rm x}, S_{\rm y}, S_{\rm z}\}$ form a basis for the real vector space $\mathcal{A}_{\rm herm}$, so the map in (4.3) is a bijection.
By Lemmas 4.1.7 and 4.2.2, the set of pure states on $\mathcal{L}(\mathcal{H})$ is given by $\{\rho_{(\theta)} : \theta \in \mathbb{R}^3,\ \|\theta\| = 1\}$, where

$$\rho_{(\theta)}(A) := \mathrm{tr}(P_\theta A) \qquad (A \in \mathcal{L}(\mathcal{H})),$$

and P_θ is defined in (4.2). By Lemma 4.2.1, the matrices $\{1, S_{\rm x}, S_{\rm y}, S_{\rm z}\}$ are orthonormal with respect to the inner product $\langle A|B\rangle := \frac{1}{2}\mathrm{tr}(AB)$. Using this and the definition of P_θ in (4.2), we see that (4.4) holds. Since the set of all states on $\mathcal{L}(\mathcal{H})$ is the convex hull of the set of pure states, we see that this set corresponds to the unit ball in $\mathbb{R}^3$. ∎

Remarks The sphere $\{\rho_{(\theta)} : \theta \in \mathbb{R}^3,\ \|\theta\| = 1\}$ is known as the *Bloch sphere*. Each $\psi \in \mathcal{H}$ (not necessarily of norm one) defines a positive linear form ρ_ψ through the formula $\rho_\psi(A) := \langle\psi|A|\psi\rangle$ $(A \in \mathcal{L}(\mathcal{H}))$. The map

$$\mathcal{H} \ni \psi \mapsto \big(\rho_\psi(S_{\rm x}), \rho_\psi(S_{\rm y}), \rho_\psi(S_{\rm z})\big) =: \theta(\psi)$$

is known as the *Hopf fibration*. One can prove that $|\theta(\psi)| = \|\psi\|^2$ and that the angle between $\theta(\psi)$ and $\theta(\phi)$ is twice the angle between ψ and ϕ. In particular, pure states that are orthogonal lie on opposite sites of the Bloch sphere and $\theta(\psi) = \theta(\phi)$ if and only if $\psi = \lambda\phi$ for some $\lambda \in \mathbb{C}$ with $|\lambda| = 1$.

4.3 The Schrödinger equation

In order to present the Schrödinger equation in a historically correct way, we will make a small detour to infinite dimensions. The Schrödinger equation describes the time evolution of a *nonrelativistic scalar massive point particle* in a *potential*.

Here *nonrelativistic* means that we neglect Einstein's theory of relativity and describe space and time with Carthesian coordinates. Thus, we specify some coordinate system and we descibe a position in space with coordinates $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, where $x_1, \ldots, x_d$ are distances measured in meters ($d = 3$ is the physically relevant dimension). We describe a moment in time with a coordinate $t \in \mathbb{R}$, measuring time in seconds before or after some specified reference time.

Scalar means that we assume that the particle has no internal degrees of freedom (such as spin). *Massive* means that the particle has a mass $m > 0$, measured in

kilogram. A *point particle* means a particle that is so small that we can neglect its spatial extensions. A *potential*, finally, is a function $V : \mathbb{R}^d \to \mathbb{R}$ which tells you that if the particle is at the point $x \in \mathbb{R}^d$, then its potential energy[4] is $V(x)$, measured in Joule, where one Jou is one kg met^2 sec^{-2}.

With these assumptions, the *Schrödinger equation*[5] describing this particle is

$$i\hbar \frac{\partial}{\partial t}\psi_t(x) = V(x)\psi_t(x) - \frac{\hbar^2}{2m}\sum_{k=1}^{d} \frac{\partial^2}{\partial {x_k}^2}\psi_t(x) \qquad (t \in \mathbb{R},\ x \in \mathbb{R}^d). \tag{4.5}$$

Here $\hbar$ is Planck's constant,[6]

$$\hbar = 1.0546 \cdot 10^{-34} \quad \text{Jou sec.}$$

The function $\psi : \mathbb{R} \times \mathbb{R}^d \to \mathbb{C}$ is called the *wave function*. Note that $\hbar$ is very small. However, if we measure space in ångström, which is 10^{-10} meter (a typical distance in atoms), mass in $1.0546 \cdot 10^{-30}$ kilo (which is a bit more than the mass of an electron), and time in 10^{-16} seconds, then $\hbar = 1$.

To relate (4.5) to the formalism of quantum mechanics, consider the linear space

$$\Big\{\psi : \mathbb{R}^d \to \mathbb{C},\ \psi \text{ measurable},\ \int |\psi(x)|^2 \mathrm{d}x < \infty\Big\},$$

and let $\mathcal{H}$ be the space of equivalence classes of a.e. equal functions from this space, equipped with the inner product

$$\langle \psi | \phi \rangle := \int \psi(x)^* \phi(x) \mathrm{d}x.$$

Then $\mathcal{H}$ is a Hilbert space. The linear operators

$$\begin{aligned} V\psi\,(x) &:= V(x)\psi(x), \\ P_k\psi\,(x) &:= -i\hbar \tfrac{\partial}{\partial x_k}\psi(x), \\ Q_k\psi\,(x) &:= x_k\psi(x), \end{aligned}$$

[4]For example, due to the force of gravitation, the potential energy of a particle is proportional to its mass (in kg) times its height (in met) times the gravitational accelaration (which is 9.8 met sec^{-2}). Note that kg×met×met sec^{-2}=Jou, the unit of energy.

[5]Suggested by Schrödinger in 1926 to describe the behavior of an electron in the Coulomb potential from the nucleus of an atom.

[6]Introduced by Max Planck in 1900 in a formula for black body radiation. In fact, Planck introduced the constant $h = 2\pi\hbar$.

can be interpreted as (unbounded, densely defined) self-adjoint linear operators on $\mathcal{H}$. V is interpreted as the observable corresponding to the potential energy of the particle. $(P_1, \ldots, P_d)$ is the observable corresponding to the momentum of the particle (clasically defined as mass times velocity), and $(Q_1, \ldots, Q_d)$ is the observable corresponding to the position of the particle. Now the self-adjoint operator

$$H := V + \frac{1}{2m} \sum_{k=1}^{d} P_k^2$$

corresponds to the energy of the particle. The operator H is also called the *Hamiltonian*. Note that with this notation, equation (4.5) takes the form

$$i\hbar \frac{\partial}{\partial t} \psi_t = H\psi_t \qquad (t \in \mathbb{R}). \tag{4.6}$$

One can show that for each $\psi_0 \in \mathcal{H}$ equation (4.5) (suitably interpreted) has a unique solution $(\psi_t)_{t\in\mathbb{R}}$ in $\mathcal{H}$ satisfying

$$\|\psi_t\| = \|\psi_0\| \qquad (t \in \mathbb{R}).$$

One usually normalizes such that $\|\psi_0\| = 0$ and interprets ψ_t as the pure state describing the particle at time t.

There exist also Schrödinger equations to describe the time evolution of two or more scalar point particles. For example, if we have n particles with masses $m_1, \ldots, m_n$, and we denote the coordinates of the j-th particle in 3-dimensional space by $x_j = (x_{j,1}, x_{j,2}, x_{j,3})$, then the corresponding Schrödinger equation reads

$$i\hbar \frac{\partial}{\partial t} \psi_t(x) = V(x)\psi_t(x) - \sum_{j=1}^{n} \frac{\hbar^2}{2m_j} \sum_{l=1}^{3} \frac{\partial^2}{\partial {x_{j,l}}^2} \psi_t(x) \qquad (t \in \mathbb{R},\ x \in (\mathbb{R}^3)^n).$$

Here $V(x)$ is typically of the form

$$V(x) = \sum_{j=1}^{n} V_j(x_j) + \sum_{j\neq k} V_{jk}(|x_j - x_k|),$$

where $V_1, \ldots, V_n$ are external fields and V_{jk} are two-point potentials describing the forces between the particles j and k.

4.4 Deterministic time evolution

In this section we return to finite dimensions. We moreover choose units such that $\hbar = 1$. Let $\mathcal{H}$ be a finite-dimensional inner product space and let $\mathcal{A} = \mathcal{L}(\mathcal{H})$ be

the $*$-algebra of all linear operators on $\mathcal{H}$. Let $H \in \mathcal{A}$ be any hermitian operator. Motivated by the Schrödinger equation in its form (4.6) we look at the differential equation

$$i\tfrac{\partial}{\partial t}\psi_t = H\psi_t. \tag{4.7}$$

This is just a linear differential equation in finite dimensions with Lipschitz coefficients, so it follows from basic theory of ordinary differential equations that for each initial condition $\psi_0 \in \mathcal{H}$ there is a unique continuously differentiable solution $t \mapsto \psi_t$. In fact, this solution is given by

$$\psi_t = e^{-itH}\psi_0 \qquad (t \in \mathbb{R}),$$

where the linear operator e^{-itH} is defined using the functional calculus for normal operators.

Exercise 4.4.1 Show that the operators e^{-itH} (with $t \in \mathbb{R}$) are unitary. Show that $\|e^{-itH}\psi\| = \|\psi\|$ for all $\psi \in \mathcal{H}$ and $t \in \mathbb{R}$.

We interpret the vector ψ_t, normalized such that $\|\psi_t\| = 1$, as the pure state of the physical system under consideration. More precisely, we interpret (4.7) as saying that if the system at time $t = 0$ is described by the pure state

$$\rho_0(A) = \rho_{\psi_0}(A) = \langle\psi_0|A\psi_0\rangle \qquad (A \in \mathcal{A}),$$

then at time t the state is

$$\rho_t(A) = \langle\psi_t|A\psi_t\rangle = \langle e^{-itH}\psi_0|Ae^{-itH}\psi_0\rangle = \langle\psi_0|e^{itH}Ae^{-itH}\psi_0\rangle = \rho_0(e^{itH}Ae^{-itH}).$$

Note that the right-hand side of this formula is representation independent, and also defined for probabilities ρ that are not pure states. This motivates us to give the following description of deterministic time evolution in quantum mechanics:

Let $\mathcal{A} = \mathcal{L}(\mathcal{H})$ describe the observables of a physical system. Then deterministic time evolution of probabilities on $\mathcal{A}$ is described by an hermitian operator $H \in \mathcal{A}$, called *Hamiltonian*. If ρ_0 is the probability law describing our knowledge about the system at time 0, then our knowledge about the system at time t is described by the probability law

$$\rho_t(A) = \rho_0(e^{itH}Ae^{-itH}) \qquad (t \in \mathbb{R},\ A \in \mathcal{A}).$$

Exercise 4.4.2 Show that adding a constant to the Hamiltonian does not change the time evolution, a fact well-known in physics.

Exercise 4.4.3 Show that deterministic time evolution of the type just described preserves pure states, i.e., if ρ_0 is a pure state then ρ_t is a pure state.

Remark There is a certain similarity between part 4° of our interpretation of quantum probability spaces and deterministic time evolution. Recall from Section 2.3 that if someone performs the ideal measurement $\{P_1, \ldots, P_n\}$ on the system, then the state of the system changes as $\rho \mapsto \rho'$ with $\rho'(A) := \sum_{i=1}^n \rho(P_i A P_i)$. This looks a bit similar to the formula for time evoution $\rho_t(A) := \rho_0(e^{itH} A e^{-itH})$. In Section 8.4 below, we will see that both are special cases of *operations*, which correspond to things one can "do" with a physical system. In particular, in Proposition 8.4.4 we will prove a deeper converse of Exercise 4.4.3: any time evolution, satisfying certain natural conditions, that maps pure states into pure states is of the type just described.

The fact that our time evolution maps pure states, which are probabilities that give 'maximal knowledge' about our system, into pure states, is why we have called the time evolution described by a Hamiltonian deterministic. A remarkable property of this sort of time evolution is that it is *reversible*, in the sense that from the state at time t we can deduce the state at all earlier times. This is true regardless of the fact that if the Schrödinger equation (4.5) with $V \equiv 0$ is started in a pure state ψ_0 that is localized in a small region of space, then as time tends to infinity the solution of (4.5) becomes more and more spread out in space.

The description of time evolution we have just given, where probabilities evolve in time, is called the *Schrödinger picture*. The so-called *Heisenberg picture* takes a somewhat different point of view. Here, there is just one probability law ρ, describing our knowledge about a physical system at all times, but the observables evolve in time, in the following way: If $A_0 \in \mathcal{A}$ is a hermitian operator describing a physical quantity at time 0, then the same physical quantity at another time t is described by the hermitian(!) operator:

$$A_t = e^{itH} A_0 e^{-itH} \qquad (t \in \mathbb{R}).$$

It is not hard to see that both pictures are equivalent, i.e., if we know for every observable the probabilities that an ideal measurement at time 0 yields a certain outcome, and we want to calculate from that the probability that an ideal measurement of an observable at some other time yields a certain outcome, then we get the same answer in both pictures.

Chapter 5

Algebras

5.1 Introduction

Recall that Theorem 2.1.5 says that every Q-algebra has a faithful representation on a complex inner product space $\mathcal{H}$. Assuming the validity of this theorem, in the present chapter, we will determine the general structure of Q-algebra's and their representations. For the information of the reader, we outline a crude sketch of the proof of Theorem 2.1.5 and its infinite dimensional analogue, Theorem 3.4.1, in Section 5.8.

To give the reader a rough idea of what we are up to, recall that the simplest example of a Q-algebra is an algebra of the form $\mathcal{A} \cong \mathcal{L}(\mathcal{H})$, where $\mathcal{H}$ is some complex inner product space. For example, if $\dim(\mathcal{H}) = 3$, then, with respect to a given orthonormal basis, the simplest possible representation of $\mathcal{A}$ consists of all matrices of the form

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad (a_{11}, \ldots, a_{33} \in \mathbb{C}).$$

This is not the only possible representation of $\mathcal{A}$. For example, on a space with dimension 6, we may represent the same algebra as the set of all matrices of the form

$$A' = \begin{pmatrix} a_{11} & a_{12} & a_{13} & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{11} & a_{12} & a_{13} \\ 0 & 0 & 0 & a_{21} & a_{22} & a_{23} \\ 0 & 0 & 0 & a_{31} & a_{32} & a_{33} \end{pmatrix} \quad (a_{11}, \ldots, a_{33} \in \mathbb{C}).$$

of course, we can play the same trick on spaces with dimensions $9, 12, \ldots$. Now let us consider something different. Consider a space $\mathcal{H}$ with $\dim(\mathcal{H}) = 5$, choose an orthonormal basis, and consider matrices of the form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 \\ 0 & 0 & 0 & b_{11} & b_{12} \\ 0 & 0 & 0 & b_{21} & b_{22} \end{pmatrix} \quad (a_{11}, \ldots, b_{22} \in \mathbb{C}).$$

It is not hard to check that the space of all matrices of this form is a sub-$*$-algebra of $\mathcal{L}(\mathcal{H})$. Of course, this algebra is isomorphic to the algebra of all matrices of the form

$$\begin{pmatrix} A & & \\ & B & \\ & & B \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_{11} & b_{12} & 0 & 0 \\ 0 & 0 & 0 & b_{21} & b_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & b_{11} & b_{12} \\ 0 & 0 & 0 & 0 & 0 & b_{21} & b_{22} \end{pmatrix},$$

where we have repeated the second block. Perhaps surprisingly, we will prove that this is about as general as one can get. For any complex inner product space $\mathcal{H}$ and sub-$*$-algebra of $\mathcal{A} \subset \mathcal{L}(\mathcal{H})$, we can find an orthonormal basis of $\mathcal{H}$ such that with respect to this basis, a general element of $\mathcal{A}$ has the *block-diagonal form*

$$\begin{pmatrix} A_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & A_1 & & & & & & \\ & & & A_2 & & & & & \\ & & & & \ddots & & & & \\ & & & & & A_2 & & & \\ & & & & & & \ddots & & \\ & & & & & & & A_n & \\ & & & & & & & & \ddots \\ & & & & & & & & & A_n \end{pmatrix} \quad (A_1 \in \mathcal{L}(\mathcal{H}_1), \ldots, A_n \in \mathcal{L}(\mathcal{H}_n)),$$

where the block A_k is repeated m_k times ($k = 1, \ldots, n$). Note that such an algebra is abelian if and only if $\dim(\mathcal{H}_k) = 1$ for all k, i.e., if each block is a 1×1 matrix.

5.2 Decomposition of algebras

If $\mathcal{A}_1, \ldots, \mathcal{A}_n$ are algebras, then we equip their direct sum $\mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$ with the structure of an algebra by putting

$$(A_1 + \cdots + A_n)(B_1 + \cdots + B_n) := (A_1B_1 + \cdots + A_nB_n).$$

Here we view $\mathcal{A}_1, \ldots, \mathcal{A}_n$ as linear subspaces of $\mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$ with the property that each $A \in \mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$ can in a unique way be written as $A = A_1 + \cdots + A_n$ with $A_1 \in \mathcal{A}_1, \ldots, A_n \in \mathcal{A}_n$. If $\mathcal{A}_1, \ldots, \mathcal{A}_n$ are $*$-algebras, then we make $\mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$ into a $*$-algebra by putting

$$(A_1 + \cdots + A_n)^* := (A_1^* + \cdots + A_n^*).$$

Note that if $\mathcal{H}_1, \mathcal{H}_2$ are complex inner product spaces and $\mathcal{A}_1, \mathcal{A}_2$ are sub-$*$-algebras of $\mathcal{L}(\mathcal{H}_1), \mathcal{L}(\mathcal{H}_2)$, respectively, then the $*$-algebra $\mathcal{A}_1 \oplus \mathcal{A}_2$ is isomorphic to the algebra of all operators on $\mathcal{H}_1 \oplus \mathcal{H}_2$ of the whose matrices (with respect to an obvious basis) have the block-diagonal form

$$\begin{pmatrix} A_1 & \\ & A_2 \end{pmatrix} \qquad (A_1 \in \mathcal{A}_1,\ A_2 \in \mathcal{A}_2).$$

By definition, a *left ideal* (resp. *right ideal*) of an algebra $\mathcal{A}$ is a linear subspace $\mathcal{I} \subset \mathcal{A}$ such that $AB \in \mathcal{I}$ (resp. $BA \in \mathcal{I}$) for all $A \in \mathcal{A}$, $B \in \mathcal{I}$. An *ideal* is a subspace that is both a left and right ideal. If $\mathcal{A}$ is a $*$-algebra, then a *$*$-ideal* is an ideal $\mathcal{I}$ with the property that $A^* \in \mathcal{I}$ for all $A \in \mathcal{I}$.

Note that if an algebra $\mathcal{A}$ is the direct sum of two other algebras, $\mathcal{A} = \mathcal{A}_1 \oplus \mathcal{A}_2$, then $\mathcal{A}_1$ is an ideal of $\mathcal{A}$. It is not a subalgebra, however, since the identity in $\mathcal{A}_1$ is not the identity in $\mathcal{A}$. If $\mathcal{A}_1$ and $\mathcal{A}_2$ are $*$-algebras and $\mathcal{A}$ is their direct sum (equipped with the standard adjoint operation), then $\mathcal{A}_1$ is a $*$-ideal of $\mathcal{A}$. By definition, an algebra is a *factor algebra* if it has no *proper ideals*, i.e., its only ideals are $\{0\}$ and $\mathcal{A}$.

Proposition 5.2.1 (Decomposition into factor algebras) *Every ideal of a Q-algebra is also a $*$-ideal. Every Q-algebra $\mathcal{A}$ can be written as a direct sum of factor algebras*

$$\mathcal{A} \cong \mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n.$$

Proof Imagine that $\mathcal{A}$ has a proper ideal $\mathcal{I}$. By Theorem 2.1.5, we can choose a faithful pseudotrace τ on $\mathcal{A}$ (in particular, we may take for τ the trace in any

faithful representation of $\mathcal{A}$). Let $\mathcal{I}$ be the orthogonal complement of $\mathcal{I}$ with respect to the inner product $\langle\cdot|\cdot\rangle_\tau$ defined in (4.1), i.e.,

$$\mathcal{I}^\perp := \{C \in \mathcal{A} : \tau(C^*B) = 0 \ \forall B \in \mathcal{I}\}.$$

We claim that $\mathcal{I}^\perp$ is another ideal of $\mathcal{A}$. Indeed, for each $A \in \mathcal{A}$, $B \in \mathcal{I}$ and $C \in \mathcal{I}^\perp$, we have $\tau((AC)^*B) = \tau(C^*(A^*B)) = 0$ and $\tau((CA)^*B) = \tau(C^*(BA^*)) = 0$, from which we see that $AC \in \mathcal{I}^\perp$ and $CA \in \mathcal{I}^\perp$. Since $\mathcal{I}^\perp$ is the orthogonal complement of $\mathcal{I}$ in the inner product $\langle\cdot|\cdot\rangle_\tau$, every element $A \in \mathcal{A}$ can in a unique way be written as $A = A_1 + A_2$ with $A_1 \in \mathcal{I}$ and $A_2 \in \mathcal{I}^\perp$. We observe that

$$(A_1 + A_2)(B_1 + B_2) = (A_1B_1 + A_2B_2) \qquad (A_1, B_1 \in \mathcal{I},\ A_2, B_2 \in \mathcal{I}^\perp) \tag{5.1}$$

where we have used that $A_1B_2, A_2B_1 \in \mathcal{I} \cap \mathcal{I}^\perp = \{0\}$. Write $1 = 1_1 + 1_2$, where $1_1 \in \mathcal{I}$ and $1_2 \in \mathcal{I}^\perp$. It is easy to see that 1_1 is a unit element in $\mathcal{I}$ and 1_2 is a unit element in $\mathcal{I}^\perp$, and that $\mathcal{I}$ and $\mathcal{I}^\perp$ (equipped with these unit elements) are algebras. This shows that $\mathcal{A}$ is the direct sum of $\mathcal{A}_1 := \mathcal{I}$ and $\mathcal{A}_2 := \mathcal{I}^\perp$ in the sense of algebras.
To complete the proof, we must show that $\mathcal{I}$ and $\mathcal{I}^\perp$ are $*$-ideals; then it will follow that $\mathcal{A}_1 = \mathcal{I}$ and $\mathcal{A}_2 = \mathcal{I}^\perp$ are Q-algebras and that $\mathcal{A}$ is the direct sum of $\mathcal{A}_1$ and $\mathcal{A}_2$ in the sense of $*$-algebras. If either $\mathcal{A}_1$ or $\mathcal{A}_2$ has a proper ideal, then we can continue the process until we have only factor algebras left. By symmetry, it suffices to show that $\mathcal{I}$ is a $*$-ideal.
We claim that for any $A \in \mathcal{A}$,

$$A \in \mathcal{I} \text{ if and only if } \langle B|AC\rangle_\tau = 0 \quad \text{for all } B, C \in \mathcal{I}^\perp. \tag{5.2}$$

To prove this, write $A = A_1 + A_2$ with $A_1 \in \mathcal{I}$ and $A_2 \in \mathcal{I}^\perp$. Then, for any $B, C \in \mathcal{I}^\perp$, one has $\langle B|AC\rangle_\tau = \langle B|A_2C\rangle_\tau$ by (5.1), which is zero if $A_2 = 0$, and nonzero if $C = 1_2$ and $B = A_2$. Now, if $A \in \mathcal{I}$, then by (5.2), $0 = \langle B|AC\rangle_\tau = \tau(B^*AC) = \tau((A^*B)^*C) = \langle A^*B|C\rangle_\tau = \langle C|A^*B\rangle_\tau^*$ for all $B, C \in \mathcal{I}^\perp$, which shows that $A^* \in \mathcal{I}$. ∎

5.3 Decomposition of representations

Recall that a representation of an algebra (resp. $*$-algebra) $\mathcal{A}$ is a pair $(\mathcal{H}, l)$ where $\mathcal{H}$ is a linear space (resp. inner product space) and $l : \mathcal{A} \to \mathcal{L}(\mathcal{H})$ is an algebra homomorphism (resp. $*$-algebra homomorphism). A somewhat different way of looking at representations is as follows. Let $\mathcal{A}$ be an algebra and let $\mathcal{H}$ be a linear

space. Imagine that we are given a map $(A, \phi) \to A\phi$ from $\mathcal{A} \times \mathcal{H}$ to $\mathcal{H}$ with the following properties:

$$\begin{array}{lll} \text{(a)} & A(a\phi + b\psi) = aA\phi + bA\psi & (a, b \in \mathbb{K},\ A \in \mathcal{A},\ \phi, \psi \in \mathcal{H}), \\ \text{(b)} & (aA + bB)\phi = aA\phi + bB\phi & (a, b \in \mathbb{K},\ A, B \in \mathcal{A},\ \phi \in \mathcal{H}), \\ \text{(c)} & (AB)\phi = A(B\phi) & (A, B \in \mathcal{A},\ \phi \in \mathcal{H}), \\ \text{(d)} & 1\phi = \phi & (\phi \in \mathcal{H}). \end{array}$$

Then the map $l : \mathcal{A} \to \mathcal{L}(\mathcal{H})$ defined by $l(A)\phi := A\phi$ is an algebra homomorphism. If moreover, $\mathcal{H}$ is equipped with an inner product such that

$$\text{(e)} \quad \langle \phi | A\psi \rangle = \langle A^*\phi | \psi \rangle \qquad (A \in \mathcal{A},\ \phi, \psi \in \mathcal{H}),$$

then l is a $*$-algebra homomorphism. Conversely, if $l : \mathcal{A} \to \mathcal{L}(\mathcal{H})$ is an algebra homomorphism (resp. $*$-algebra homomorphism), then setting $A\phi := l(A)\phi$ defines a map from $\mathcal{A} \times \mathcal{H}$ to $\mathcal{H}$ with the properties (a)–(d) (resp. (a)–(e)). We call such a map an *action* of the algebra $\mathcal{A}$ on $\mathcal{H}$. Thus, we can view representations of an algebra (resp. $*$-algebra) $\mathcal{A}$ as linear spaces (resp. inner product spaces) on which there is defined an action of $\mathcal{A}$. Which is a long way of saying that from now on, we will often drop the map l from our notation, write $A\phi$ instead of $l(A)\phi$, and write phrases like: 'let $\mathcal{H}$ be a representation of $\mathcal{A}$'.

Exercise 5.3.1 Let $\mathcal{A}$ be an algebra. Show that $\mathcal{A}$, equipped with the action $(A, B) \mapsto AB$, becomes a representation of itself. If $\mathcal{A}$ is a $*$-algebra and τ is a faithful pseudotrace on $\mathcal{A}$, then show that $\mathcal{A}$ equipped with the inner product $\langle \cdot | \cdot \rangle_\tau$ is a faithful representation of itself as a $*$-algebra.

If $\mathcal{H}_1, \ldots, \mathcal{H}_n$ are representations of an algebra (resp. $*$-algebra) $\mathcal{A}$, then we equip the direct sum $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$ with the structure of a representation of $\mathcal{A}$ by putting

$$A(\phi(1) + \cdots + \phi(n)) := A\phi(1) + \cdots + A\phi(n),$$

where $\phi(1) \in \mathcal{H}_1, \ldots, \phi(n) \in \mathcal{H}_n$. It is not hard to see that this action of $\mathcal{A}$ on $\mathcal{H}_1 \oplus \ldots \oplus \mathcal{H}_n$ has the properties (a)–(d) (resp. (a)–(e)). If $\mathcal{H}_1, \mathcal{H}_2$ are representations and $l_i : \mathcal{A} \to \mathcal{L}(\mathcal{H}_i)$ $(i = 1, 2)$ are the associated algebra homomorphisms, then the algebra homomorphism associated with $\mathcal{H}_1 \oplus \mathcal{H}_2$ takes the form

$$\mathcal{A} \ni A \mapsto \begin{pmatrix} l_1(A) & \\ & l_2(A) \end{pmatrix} \in \mathcal{L}(\mathcal{H}_1 \oplus \mathcal{H}_2).$$

Note that in particular, we can have $\mathcal{H}_1 = \mathcal{H}_2$, in which case the same block $l_1(A) = l_2(A)$ is repeated two times.

By definition, an *invariant subspace* of a representation $\mathcal{H}$ of some algebra $\mathcal{A}$ is a linear subspace $\mathcal{F} \subset \mathcal{H}$ such that

$$\phi \in \mathcal{F} \text{ implies } A\phi \in \mathcal{F} \qquad (A \in \mathcal{A}).$$

Note that $\mathcal{F}$, equipped with the obvious action, is itself a representation of $\mathcal{A}$. We say that a representation $\mathcal{H}$ of an algebra $\mathcal{A}$ is *irreducible* if it has no *proper invariant subspaces*, i.e., invariant subspaces that are not $\{0\}$ or $\mathcal{H}$.

Lemma 5.3.2 (Decomposition of representations) *Every representation of a Q-algebra can be written as a direct sum of irreducible representations.*

Proof If $\mathcal{H}$ is a representation of a $*$-algebra and $\mathcal{F}$ is an invariant subspace, then we claim that $\mathcal{F}^\perp$ is also an invariant subspace. Indeed, $\psi \in \mathcal{F}^\perp$ implies $\langle A\psi|\phi\rangle = \langle\psi|A^*\phi\rangle = 0$ for all $\phi \in \mathcal{F}$, which implies $A\psi \in \mathcal{F}^\perp$. It follows that $\mathcal{H} \cong \mathcal{F} \oplus \mathcal{F}^\perp$. We can continue this process until we arrive at a decomposition of $\mathcal{H}$ into invariant subspaces that have no further proper invariant subspaces. ∎

5.4 Von Neumann's bicommutant theorem

In the next section, we will study the structure of factor algebras and irreducible representations. To prepare for this, we need a result that is known as Von Neumann's bicommutant theorem.
Let $\mathcal{H}$ be an (as usual finite dimensional) inner product space over $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$. For any set $\mathcal{A} \subset \mathcal{H}$, we let

$$\mathcal{A}^{\rm c} := \{B \in \mathcal{L}(\mathcal{H}) : [A, B] = 0 \ \forall A \in \mathcal{A}\}$$

denote the *commutant* of $\mathcal{A}$.

Exercise 5.4.1 Show that for any set $\mathcal{A} \subset \mathcal{L}(\mathcal{H})$, the commutant $\mathcal{A}^{\rm c}$ is a sub-algebra of $\mathcal{L}(\mathcal{H})$. Show that if $\mathcal{A}$ is closed under taking of adjoints, then the same is true for $\mathcal{A}^{\rm c}$. In particular, if $\mathcal{A}$ is a sub-$*$-algebra of $\mathcal{L}(\mathcal{H})$, then so is $\mathcal{A}^{\rm c}$.

We call $(\mathcal{A}^{\rm c})^{\rm c}$ the *bicommutant* of $\mathcal{A}$. The following result is known as Von Neumann's bicommutant theorem.

Theorem 5.4.2 (Bicommutant theorem) *Let $\mathcal{H}$ be an inner product space over $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$ and let $\mathcal{A}$ be a sub-$*$-algebra of $\mathcal{L}(\mathcal{H})$. Then $(\mathcal{A}^{\rm c})^{\rm c} = \mathcal{A}$.*

We start with two preparatory lemmas. By definition, we say that a linear subspace $\mathcal{F} \subset \mathcal{H}$ is *invariant* under an operator $A \in \mathcal{L}(\mathcal{H})$ if $\psi \in \mathcal{F}$ implies $A\psi \in \mathcal{F}$.

Lemma 5.4.3 *Let $A \in \mathcal{L}(\mathcal{H})$ be an operator, let $\mathcal{F} \subset \mathcal{H}$ be a linear subspace, and let $P_{\mathcal{F}}$ denote the orthogonal projection on $\mathcal{F}$. Then one has $[A, P_{\mathcal{F}}] = 0$ if and only if $\mathcal{F}$ and $\mathcal{F}^{\perp}$ are invariant under A.*

Proof We observe that for any $\psi \in \mathcal{H}$:

$$\psi \in \mathcal{F} \quad \Leftrightarrow \quad P_{\mathcal{F}}\psi = \psi. \tag{5.3}$$

Moreover

$$\mathcal{F} = \{P_{\mathcal{F}}\psi : \psi \in \mathcal{H}\}. \tag{5.4}$$

Now if $[A, P_{\mathcal{F}}] = 0$ and $\psi \in \mathcal{F}$ then by (5.3) $A\psi = AP_{\mathcal{F}}\psi = P_{\mathcal{F}}A\psi \in \mathcal{F}$ which shows that $\mathcal{F}$ is invariant under A. Moreover, since

$$P_{\mathcal{F}^{\perp}} = 1 - P_{\mathcal{F}},$$

we have $[A, P_{\mathcal{F}^{\perp}}] = [A, 1 - P_{\mathcal{F}}] = [A, 1] - [A, P_{\mathcal{F}}] = 0$, so the same argument as before shows that also $\mathcal{F}^{\perp}$ is invariant. Conversely, assume that $\mathcal{F}$ and $\mathcal{F}^{\perp}$ are invariant under A. We can uniquely decompose a general element $\psi \in \mathcal{H}$ as $\psi = \phi + \eta$ with $\phi \in \mathcal{F}$ and $\eta \in \mathcal{F}^{\perp}$. Then $AP_{\mathcal{F}}\psi = A\phi$. On the other hand, $A\psi = A\phi + A\eta$ where by the assumption that $\mathcal{F}$ and $\mathcal{F}^{\perp}$ are invariant under A, we have $A\phi \in \mathcal{F}$ and $A\eta \in \mathcal{F}^{\perp}$. It follows that $P_{\mathcal{F}}(A\psi) = A\phi = AP_{\mathcal{F}}\psi$, showing that A and $P_{\mathcal{F}}$ commute. ∎

Exercise 5.4.4 Let $A \in \mathcal{L}(\mathcal{H})$ be an operator, let $\mathcal{F} \subset \mathcal{H}$ be a linear subspace, and let $P_{\mathcal{F}}$ denote the orthogonal projection on $\mathcal{F}$. Assume that $\mathcal{F}$ is invariant under A. Show by counterexample that this does *not* imply that A and $P_{\mathcal{F}}$ commute.

The next lemma is not as trivial as the previous one.

Lemma 5.4.5 *Let $\mathcal{H}$ be an inner product space over $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$ and let $\mathcal{A}$ be a sub-$*$-algebra of $\mathcal{L}(\mathcal{H})$. Then, for all $\psi \in \mathcal{H}$ and $B \in (\mathcal{A}^{\mathrm{c}})^{\mathrm{c}}$, there exists an $A \in \mathcal{A}$ such that $A\psi = B\psi$.*

Proof Fix $\psi \in \mathcal{H}$, and consider the linear subspace $\mathcal{F} := \{A\psi : A \in \mathcal{A}\}$. We claim that the algebra $\mathcal{A}$ leaves the spaces $\mathcal{F}$ and $\mathcal{F}^{\perp}$ invariant, i.e., $\phi \in \mathcal{F}$ implies $A\phi \in \mathcal{F}$ and $\phi \in \mathcal{F}^{\perp}$ implies $A\phi \in \mathcal{F}^{\perp}$ for all $A \in \mathcal{A}$. Indeed, if $\phi \in \mathcal{F}$ then ϕ is of the form $\phi = A'\psi$ for some $A' \in \mathcal{A}$, hence $A\phi = AA'\psi \in \mathcal{F}$, and if $\phi \in \mathcal{F}^{\perp}$ then $\langle \phi | A\psi \rangle = 0$ for all $A' \in \mathcal{A}$, hence $\langle A\phi | A'\psi \rangle = \langle \phi | A^*A'\psi \rangle = 0$ for all $A' \in \mathcal{A}$, hence $A\phi \in \mathcal{F}^{\perp}$. It follows that each element of $\mathcal{A}$ commutes with the orthogonal projection $P_{\mathcal{F}}$ on $\mathcal{F}$, i.e., $P_{\mathcal{F}} \in \mathcal{A}^{\mathrm{c}}$. Hence, if $B \in (\mathcal{A}^{\mathrm{c}})^{\mathrm{c}}$, then B commutes with

$P_{\mathcal{F}}$, which implies that B leaves the spaces $\mathcal{F}$ and $\mathcal{F}^{\perp}$ invariant. In particular, $B\psi \in \mathcal{F}$, which shows that $B\psi = A\psi$ for some $A \in \mathcal{A}$. ∎

Proof of Theorem 5.4.2 Lemma 5.4.5 says that for each $B \in (\mathcal{A}^{\mathrm{c}})^{\mathrm{c}}$ and $\psi \in \mathcal{H}$ we can find an $A \in \mathcal{A}$ such that A and B agree on ψ. In order to prove the theorem, we must show that we can find an $A \in \mathcal{A}$ such that A and B agree on *all* vectors in $\mathcal{H}$. By linearity, it suffices to do this for a basis of $\mathcal{H}$. Thus, we need to show that for any $B \in (\mathcal{A}^{\mathrm{c}})^{\mathrm{c}}$ and $\psi(1), \ldots, \psi(n) \in \mathcal{H}$, there exists an $A \in \mathcal{A}$ such that $A\psi(i) = B\psi(i)$ for all $i = 1, \ldots, n$.
Let $\mathcal{H}_1, \ldots, \mathcal{H}_n$ be n identical copies of $\mathcal{H}$, and consider the direct sum $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$. Let $\mathcal{A}^{(n)}$ denote the sub-$*$-algebra of $\mathcal{L}(\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n)$ consisting of all operators of the form

$$A^{(n)}(\phi(1), \ldots, \phi(n)) := (A\phi(1), \ldots, A\phi(n))$$

for some $A \in \mathcal{A}$. We wish to describe the commutant $(A^{(n)})^{\mathrm{c}}$. With respect to an obvious orthonormal basis for $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$, each $A^{(n)} \in \mathcal{A}^{(n)}$ has the block-diagonal form (for example for $n = 3$):

$$A^{(n)} = \begin{pmatrix} A & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & A \end{pmatrix}.$$

Now any $C \in \mathcal{L}(\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n)$ can be written as

$$C = \begin{pmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{pmatrix},$$

where the C_{ij} are linear maps from $\mathcal{H}$ into $\mathcal{H}$. We see that

$$A^{(n)}C = \begin{pmatrix} AC_{11} & AC_{12} & AC_{13} \\ AC_{21} & AC_{22} & AC_{23} \\ AC_{31} & AC_{32} & AC_{33} \end{pmatrix} \quad \text{and} \quad CA^{(n)} = \begin{pmatrix} C_{11}A & C_{12}A & C_{13}A \\ C_{21}A & C_{22}A & C_{23}A \\ C_{31}A & C_{32}A & C_{33}A \end{pmatrix},$$

and therefore C commutes with each $A^{(n)}$ in $\mathcal{A}^{(n)}$ if and only if $C_{ij} \in \mathcal{A}^{\mathrm{c}}$ for each i, j.
Now let $B \in (\mathcal{A}^{\mathrm{c}})^{\mathrm{c}}$ and $\psi(1), \ldots, \psi(n) \in \mathcal{H}$. By what we have just proved, it is easy to see that $B^{(n)} \in ((A^{(n)})^{\mathrm{c}})^{\mathrm{c}}$. Therefore, applying Lemma 5.4.5 to $B^{(n)}$ and the vector $(\psi(1), \ldots, \psi(n)) \in \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$, we conclude that there exists an $A^{(n)} \in \mathcal{A}^{(n)}$ such that

$$A^{(n)}(\psi(1), \ldots, \psi(n)) = B^{(n)}(\psi(1), \ldots, \psi(n))$$

i.e., $A\psi(i) = B\psi(i)$ for all $i = 1, \ldots, n$, as desired. ∎

5.5 Factor algebras

We have seen that any Q-algebra $\mathcal{A}$ can be decomposed into factor algebras and that any representation of $\mathcal{A}$ can be decomposed into irreducible representations. In the present section, we will see that there is a close connection between these two objects.
By definition, the *center* of a Q-algebra is the abelian sub-$*$-algebra $\mathcal{C}(\mathcal{A}) \subset \mathcal{A}$ given by

$$\mathcal{C}(\mathcal{A}) := \{C \in \mathcal{A} : [A, C] = 0 \ \forall A \in \mathcal{A}\},$$

i.e., $\mathcal{C}(\mathcal{A})$ consists of those elements of $\mathcal{A}$ that commute with all elements of $\mathcal{A}$. We say that the center is *trivial* if $\mathcal{C}(\mathcal{A}) = \{a1 : a \in \mathbb{C}\}$.

Theorem 5.5.1 (Factor algebras) *Let $\mathcal{A}$ be a Q-algebra. Then the following statements are equivalent.*

(1) *$\mathcal{A}$ is a factor algebra.*

(2) *$\mathcal{A}$ has a faithful irreducible representation.*

(3) *$\mathcal{A} \cong \mathcal{L}(\mathcal{H})$ for some inner product space $\mathcal{H}$.*

(4) *$\mathcal{A}$ has a trivial center.*

Proof (1)⇒(2): By finite dimensionality each algebra has an irreducible representation. We claim that representations of factor algebras are always faithful. Indeed, if $(\mathcal{H}, l)$ is a representation of an algebra $\mathcal{A}$, then it is easy to see that the kernel $\mathrm{Ker}(l) = \{A \in \mathcal{A} : l(A) = 0\}$ is an ideal of $\mathcal{A}$. In particular, if $\mathcal{A}$ is a factor algebra, we must have $\mathrm{Ker}(l) = \mathcal{A}$ or $\mathrm{Ker}(l) = \{0\}$. Since $l(1) = 1 \neq 0$, the first option can be excluded, hence $(\mathcal{H}, l)$ is faithful.
(2)⇒(3): It suffices to show that if $\mathcal{H}$ is an inner product space and $\mathcal{A} \subset \mathcal{L}(\mathcal{H})$ is a sub-$*$-algebra with no proper invariant subspaces, then $\mathcal{A} = \mathcal{L}(\mathcal{H})$. By Von Neumann's bicommutant theorem, it suffices to prove that $\mathcal{A}^{\mathrm{c}} = \{a1 : a \in \mathbb{C}\}$. By Lemma 4.1.3, it suffices to show that $0, 1$ are the only projections in $\mathcal{A}$. Imagine that $0, 1 \neq P \in \mathcal{A}$ is a projection. Let $\mathcal{F}$ be the space that P projects upon. By Lemma 5.4.3, $\mathcal{F}$ is invariant under each $A \in \mathcal{A}$, so we contradict the assumption that $\mathcal{A}$ has no proper invariant subspaces.
(3)⇒(4): It suffices to show that $\mathcal{L}(\mathcal{H})$ has a trivial center. By Lemma 4.1.3, it suffices to show that $0, 1$ are the only projections in the center. By the same argument as used for the previous implication, if the center contains a nontrivial projection, then $\mathcal{L}(\mathcal{H})$ has a proper invariant subspace. It is easy to see, however, that $\mathcal{L}(\mathcal{H})$ has no proper invariant subspaces.

(4)⇒(1): If $\mathcal{A}$ is not a factor algebra, then by Proposition 5.2.1, we can write $\mathcal{A} \cong \mathcal{A}_1 \oplus \mathcal{A}_2$, where $\mathcal{A}_1$ and $\mathcal{A}_2$ are Q-algebras. Now the identity $1_1 \in \mathcal{A}_1$ is a nontrivial element of the center $\mathcal{C}(\mathcal{A})$, hence the latter is not trivial. ∎

The proof of Theorem 5.5.1 has a useful corollary.

Corollary 5.5.2 (Representations of factors) *Let $\mathcal{A}$ be a factor algebra. Then each representation $(\mathcal{H}, l)$ of $\mathcal{A}$ is faithful. If moreover $\mathcal{A}$ is a Q-algebra and $(\mathcal{H}, l)$ is irreducible, then $l : \mathcal{A} \to \mathcal{L}(\mathcal{H})$ is surjective.*

Proof This follows from the steps (1)⇒(2) and (2)⇒(3) of the proof of Theorem 5.5.1. ∎

Exercise 5.5.3 Show that on a factor algebra, there exists up to a multiplicative constant a unique pseudotrace. Hint: choose an orthonormal basis $\{e(1), \ldots, e(n)\}$ and a vector ϕ of norm one, and write $|e(i)\rangle\langle e(j)| = |e(i)\rangle\langle\phi|\phi\rangle\langle e(j)|$.

5.6 Structure of Q-algebras

Let $\mathcal{A}$ be an algebra and let $\mathcal{H}_1, \mathcal{H}_2$ be representations of $\mathcal{A}$. By definition, a *representation homomorphism* is a linear map $U : \mathcal{H}_1 \to \mathcal{H}_2$ such that

$$UA\phi = AU\phi \qquad (\phi \in \mathcal{H}_1,\ A \in \mathcal{A}).$$

Note that this says that U preserves the action of the algebra $\mathcal{A}$. If $\mathcal{A}$ is a $*$-algebra then we also require that U is unitary, i.e., U preserves the inner product. If U is a bijection then one can check that U^{-1} is also a representation homomorphism. In this case we call U a representation isomorphism and we say that $\mathcal{H}_1$ and $\mathcal{H}_2$ are *equivalent* representations of $\mathcal{A}$. Note that if $(\mathcal{H}_1, l_1)$ and $(\mathcal{H}_2, l_2)$ are equivalent representations, then

$$l_1(A) = U^{-1} l_2(A) U \qquad (A \in \mathcal{A}).$$

Lemma 5.6.1 (Irreducible representations of factor algebras) *All irreducible representations of a factor algebra $\mathcal{A}$ are equivalent.*

Proof We observe that each left ideal $\mathcal{I} \neq \{0\}$ of an algebra $\mathcal{A}$ becomes a representation of $\mathcal{A}$ if we equip it with the obvious action $(A, B) \mapsto AB$ $(A \in \mathcal{A},\ B \in \mathcal{I})$. Since a subspace $\mathcal{I}' \subset \mathcal{I}$ is invariant under the action of $\mathcal{A}$ if and only if $\mathcal{I}'$ is a left

ideal, we see that $\mathcal{I}$ is irreducible if and only if $\mathcal{I}$ is a *minimal left ideal*, i.e., the only left ideals $\mathcal{I}'$ of $\mathcal{A}$ such that $\mathcal{I}' \subset \mathcal{I}$ are $\mathcal{I}' = 0$ and $\mathcal{I}' = \mathcal{I}$. Such a minimal left ideal exists by finite dimensionality and the fact that $\mathcal{A}$ is a left ideal of itself. Now let $\mathcal{A}$ be a factor algebra and let $\mathcal{H}$ be an irreducible representation of $\mathcal{A}$. By the previous remarks, $\mathcal{A}$ has a minimal left ideal, and each minimal left ideal $\mathcal{I}$ is an irreducible representation of $\mathcal{A}$. We will show that $\mathcal{H}$ and $\mathcal{I}$ are equivalent. Since $\mathcal{H}$ is arbitrary, this proves that all irreducible representations of $\mathcal{A}$ are equivalent. Fix $0 \neq C \in \mathcal{I}$. By Corollary 5.5.2, $\mathcal{H}$ is faithful, so we can choose $\phi \in \mathcal{H}$ such that $C\phi \neq 0$. Define $U : \mathcal{I} \to \mathcal{H}$ by

$$UB := B\phi \qquad (B \in \mathcal{I}).$$

Then U is a representation homomorphism. It follows that $\mathrm{Ran}(U)$ is an invariant subspace of $\mathcal{H}$ and $\mathrm{Ker}(U)$ is an invariant subspace of $\mathcal{I}$. Since $C\phi \neq 0$, we see that $\mathrm{Ran}(U) \neq \{0\}$ and $\mathrm{Ker}(U) \neq \mathcal{I}$. Since $\mathcal{H}$ and $\mathcal{I}$ are irreducible, it follows that $\mathrm{Ran}(U) = \mathcal{H}$ and $\mathrm{Ker}(U) = \{0\}$, hence U is a linear bijection.
This completes the proof in case $\mathcal{A}$ is an algebra. In case $\mathcal{A}$ is a Q-algebra, we must additionaly show that U is unitary. Indeed, if $(\mathcal{H}_1, l_1)$ and $(\mathcal{H}_2, l_2)$ are irreducible representations of a Q-algebra $\mathcal{A}$, then by what we have just shown, there exists a linear bijection $U : \mathcal{H}_1 \to \mathcal{H}_2$ such that

$$l_2(A) = Ul_1(A)U^{-1} \qquad (A \in \mathcal{A}).$$

By Corollary 5.5.2, l_1 and l_2 are surjective, so the composition $l = l_2 \circ l_1^{-1}$ is a $*$-algebra isomorphism from $\mathcal{L}(\mathcal{H}_1)$ to $\mathcal{L}(\mathcal{H}_2)$, and

$$l(A) = UAU^{-1} \qquad (A \in \mathcal{L}(\mathcal{H}_1)).$$

Let $\{e(1), \ldots, e(n)\}$ be an orthonormal basis of $\mathcal{H}_1$. Then

$$l(|e(i)\rangle\langle e(i)|) = U|e(i)\rangle\langle e(i)|U^{-1} = |Ue(i)\rangle\langle (U^{-1})^* e(i)|.$$

Since l is a $*$-algebra isomorphism, $l(|e(i)\rangle\langle e(i)|)$ is a projection, which is only possible if

$$Ue(i) = (U^{-1})^* e(i).$$

Since this holds for each i, $U^* = U^{-1}$, i.e., U is unitary. ∎

The following theorem describes the general structure of Q-algebras and their representations.

Theorem 5.6.2 (Structure theorem for Q-algebras) *Let $\mathcal{A}$ be a Q-algebra. Then $\mathcal{A}$ has finitely many nonequivalent irreducible representations $(\mathcal{H}_1, l_1), \ldots, (\mathcal{H}_n, l_n)$, and the map*

$$A \mapsto (l_1(A), \ldots, l_n(A))$$

defines a $*$-algebra isomorphism

$$\mathcal{A} \cong \mathcal{L}(\mathcal{H}_1) \oplus \cdots \oplus \mathcal{L}(\mathcal{H}_n).$$

Every representation of $\mathcal{A}$ is equivalent to a representation of the form

$$\mathcal{H} = (\underbrace{\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_1}_{m_1 \text{ times}}) \oplus \cdots \oplus (\underbrace{\mathcal{H}_n \oplus \cdots \oplus \mathcal{H}_n}_{m_n \text{ times}}),$$

with $m_i \geq 0$ $(i = 1, \ldots, n)$. $\mathcal{H}$ is faithful if and only if $m_i \geq 1$ for all $i = 1, \ldots, n$.

The numbers $m_1, \ldots, m_n$ are called the *multiplicities* of the irreducible representations $\mathcal{H}_1, \ldots, \mathcal{H}_n$.

Proof of Theorem 5.6.2 By Proposition 5.2.1, $\mathcal{A}$ is isomorphic to a direct sum of factor algebras $\mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$. Let $(\mathcal{H}, l)$ be a representation of $\mathcal{A}$. Let $1_1, \ldots, 1_n$ denote the identities in $\mathcal{A}_1, \ldots, \mathcal{A}_n$, respectively. Then $\{l(1_1), \ldots, l(1_n)\}$ is a partition of the identity on $\mathcal{H}$. Let $\mathcal{F}_i$ be the space that $l(1_i)$ projects on (which may be zero-dimensional for some i). Then $\mathcal{H} = \mathcal{F}_1 \oplus \cdots \oplus \mathcal{F}_n$, where $\mathcal{F}_i$ is a representation of $\mathcal{A}_i$. By Lemma 5.3.2, we can split $\mathcal{F}_i$ into irreducible representations of $\mathcal{A}_i$, say $\mathcal{F}_i = \mathcal{F}_{i1} \oplus \cdots \oplus \mathcal{F}_{im(i)}$, where possibly $m(i) = 0$. Let $l_{ij} : \mathcal{A}_i \to \mathcal{L}(\mathcal{F}_{ij})$ denote the corresponding $*$-algebra homomorphism. By Corollary 5.5.2, the representations $(\mathcal{F}_{i1}, l_{i1}), \ldots, (\mathcal{F}_{im(i)}, l_{im(i)})$ are faithful and $l_{i1}, \ldots, l_{im(i)}$ are surjective. By Lemma 5.6.1, the $(\mathcal{F}_{i1}, l_{i1}), \ldots, (\mathcal{F}_{im(i)}, l_{im(i)})$ are equivalent. It is not hard to see that $(\mathcal{F}_{ij}, l_{ij})$ and $(\mathcal{F}_{i'j'}, l_{i'j'})$ are not equivalent if $i \neq i'$. From these observations the statements of the theorem follow readily. ∎

5.7 Abelian algebras

In this section, we look at abelian algebras. In particular, we will prove Theorem 2.4.1.

Theorem 5.7.1 (Abelian algebras) *Let $\mathcal{H}$ be an inner product space over $\mathbb{C}$ and let $\mathcal{A}$ be an abelian sub-$*$-algebra of $\mathcal{L}(\mathcal{H})$. Then there exists a partition of the identity $\{P_1, \ldots, P_n\}$ such that*

$$\mathcal{A} = \Big\{ \sum_{i=1}^{n} a_i P_i : a_i \in \mathbb{C} \ \forall i = 1, \ldots, n \Big\}. \tag{5.5}$$

Proof Immediate from Theorem 5.6.2. ∎

Theorem 5.7.1 has a useful corollary.

Theorem 5.7.2 (Simultaneous diagonalization of normal operators) *Let $\mathcal{H}$ be an inner product space over $\mathbb{C}$ and let $A(1), \ldots, A(k)$ be a collection of mutually commuting normal operators. Then there exists an orthonormal basis $\{e(1), \ldots, e(n)\}$ such that for each $j = 1, \ldots, k$ there exist complex numbers $\lambda_1(j), \ldots, \lambda_n(j)$ with*

$$A(k) = \sum_{i=1}^{n} \lambda_i(k) |e(i)\rangle\langle e(i)|.$$

Proof Let $\mathcal{A}$ be the $*$-algebra generated by $A(1), \ldots, A(k)$, i.e., $\mathcal{A}$ consists of all linear combinations of finite products of the operators $A(1), \ldots, A(k)$ and their adjoints. We claim that $\mathcal{A}$ is abelian. This is not quite as obvious as it may seem, since we have assumed that $A(j)$ commutes with $A(j')$ for each j, j', but not that $A(j)$ commutes with $A(j')^*$. For general operators A, B, it is not always true that A^* commutes with B if A commutes with B. For normal operators this is true, however. To see this, choose an orthonormal basis such that A is diagonal. Then $AB = BA$ implies $A_{ii}B_{ij} = B_{ij}A_{jj}$ for all i, j, hence, for each i, j we have either $B_{ij} = 0$ or $A_{ii} = A_{jj}$. It follows that $A_{ii}^* B_{ij} = B_{ij} A_{jj}^*$ for all i, j, hence $A^*B = BA^*$. Once this little complication is out of the way, the proof is easy. Since $\mathcal{A}$ is abelian, there exists a partition of the identity $\{P_1, \ldots, P_n\}$ such that each element of $\mathcal{A}$, in particular each operator $A(j)$, is a linear combination of the $P_1, \ldots, P_n$. Let $\mathcal{F}_1, \ldots \mathcal{F}_n$ be the orthogonal subspaces upon which the $P_1, \ldots, P_n$ project. Choosing an orthonormal basis of $\mathcal{H}$ that is a union of orthonormal bases of the $\mathcal{F}_1, \ldots \mathcal{F}_n$, we arrive at the desired result. ∎

We can now also easily give the:

Proof of Theorem 2.4.1 By Theorem 5.7.1, there exists a partition of the identity $\{P_1, \ldots, P_n\}$ such that $\mathcal{A}$ consists of all linear combinations of the $P_1, \ldots, P_n$. Set $\Omega = \{1, \ldots, n\}$ and define a map $l : \mathbb{C}^\Omega \to \mathcal{A}$ by

$$l(f) := \sum_{i=1}^{n} f(i) P_i.$$

It is easy to see that l is an isomorhism for $*$-algebras. ∎

Exercise 5.7.3 Let $\mathcal{A}$ be the real $*$-algebra consisting of all matrices of the form

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \qquad (a, b \in \mathbb{R}).$$

Show that $\mathcal{A}$ is abelian, but not isomorphic to $\mathbb{R}^\Omega$ for some finite set Ω. Does $\mathcal{A}$ remind you of some algebra you know?

5.8 Proof of the representation theorems*

In this section, we give a brief sketch of the proofs of Theorems 2.1.5 and 3.4.1. The proof of Theorem 3.4.1 is standard and can be found in any book on C∗-algebras (e.g. [Con90, Dav96]). Theorem 2.1.5 is rather obscure; I am indebted to Roberto Conti for pointing out its proof in [GHJ89, Appendix IIa].

By definition, an algebra $\mathcal{A}$ is semisimple if it is the direct sum of factor algebras. Not every algebra is semisimple; a counterexample is the algebra of all matrices of the form

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \qquad (a, b, c \in \mathbb{K}).$$

Proposition 5.2.1 says that every Q-algebra is semisimple. Unfortunately, our proof of Proposition 5.2.1 leans heavily on the fact that every Q-algebra has a faithful representation. The crucial step in the proof of Theorem 2.1.5 is to show that Q-algebras are semisimple using only the properties (i)–(viii) from Section 2.1.
By definition, the *Jacobson radical* $\mathcal{J}$ of an algebra $\mathcal{A}$ is the intersection of all maximal (proper) ideals in $\mathcal{A}$. It is known that $\mathcal{A}$ is semi-simple if and only if $\mathcal{J} = \{0\}$. Thus, we need to show that the Jacobson radical $\mathcal{J}$ of a Q-algebra is trivial.
It is easy to see that if $\mathcal{I}$ is a left ideal in $\mathcal{A}$, then $\mathcal{I}^* := \{A^* : A \in \mathcal{I}\}$ is a right ideal. Thus, if $\mathcal{I}$ is an ideal, then $\mathcal{I}^*$ is also an ideal. If $\mathcal{I}$ is maximal, then $\mathcal{I}^*$ is also maximal. Hence

$$\mathcal{J}^* = \bigcap\{\mathcal{I}^* : \mathcal{I} \text{ maximal ideal}\} = \bigcap\{\mathcal{I} : \mathcal{I} \text{ maximal ideal}\} = \mathcal{J}.$$

Now imagine that $0 \neq A \in \mathcal{J}$. By what we have just proved $A^* \in \mathcal{J}$ and therefore $A^*A \in \mathcal{J}$. By the positivity condition (viii) from Section 2.1, $A^*A \neq 0$, $(A^*A)^*(A^*A) = (A^*A)^2 \neq 0$, and by induction, $(A^*A)^{2^n} \neq 0$ for all $n \geq 1$. However, it is known (see e.g. [Lan71]) that the Jacobson radical of a finite-dimensional algebra is nilpotent, i.e., $\mathcal{J}^n = \{0\}$ for some n. We arrive at a contradiction.
Using again the positivity condition (viii) from Section 2.1, one can show that the adjoint operation on a Q-algebra $\mathcal{A}$ must respect the factors in the decomposition $\mathcal{A} \cong \mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$, i.e., $A \in \mathcal{A}_i$ implies $A^* \in \mathcal{A}_i$. It follows from general theory of algebras that each $\mathcal{A}_i$ is of the form $\mathcal{L}(\mathcal{V}_i)$, where $\mathcal{V}_i$ is a complex linear space. To complete the proof, it then suffices to show that the adjoint operation on $\mathcal{L}(\mathcal{V}_i)$

arises from an inner product on $\mathcal{V}_i$. To show this, choose any inner product $\langle \cdot, \cdot \rangle$ on $\mathcal{V}_i$ and let $A \mapsto A^\dagger$ denote the adjoint operation with respect to this inner product. Then $A \mapsto (A^*)^\dagger$ is an algebra isomorphism from $\mathcal{L}(\mathcal{V}_i)$ into itself. It follows from Lemma 5.6.1 that every algebra isomorphism from $\mathcal{L}(\mathcal{V}_i)$ into itself is an *inner isomorphism,* i.e., $(A^*)^\dagger = UAU^{-1}$ for some linear bijection $U : \mathcal{V}_i \to \mathcal{V}_i$. Setting $\langle x, y \rangle' := \langle Ux, Uy \rangle$ then yields an inner product on $\mathcal{V}_i$ such that $A \mapsto A^*$ is the adjoint operation with respect to this inner product.

The proof of Theorem 3.4.1 follows a completely different strategy. Let $\mathcal{A}$ be a C$*$-algebra and let ρ be a probability law (state) on $\mathcal{A}$. We claim that then there exists a representation $\mathcal{H}$ of $\mathcal{A}$ and a vector $\phi \in \mathcal{H}$ such that

$$\rho(A) = \langle \phi | A\phi \rangle \qquad (A \in \mathcal{A}).$$

To prove this, put

$$\mathcal{N} := \{A \in \mathcal{A} : \rho(A^*A) = 0\}.$$

One can check that $\mathcal{N}$ is a closed linear subspace of $\mathcal{A}$, and a left ideal. Moreover,

$$\langle A + \mathcal{N}, B + \mathcal{N} \rangle := \rho(A^*B) \tag{5.6}$$

defines an inner product on the quotient space

$$\mathcal{A}/\mathcal{N} := \{A + \mathcal{N} : A \in \mathcal{A}\}$$

Let $\mathcal{H}$ be the completion of $\mathcal{A}/\mathcal{N}$ in this inner product. Then one checks that

$$A(B + \mathcal{N}) := AB + \mathcal{N} \qquad (A, B \in \mathcal{A}) \tag{5.7}$$

defines an action of $\mathcal{A}$ on $\mathcal{H}$. Setting $\phi = 1 + \mathcal{N}$ now yields the claims. This construction is known as the *GNS-construction.*

The strategy of the proof of Theorem 3.4.1 is now to show that there exist enough states ρ on $\mathcal{A}$ so that the direct sum of their corresponding representations, obtained with the GNS-construction, is faithful. The proof is not easy; one more or less has to derive the whole spectral theory of normal elements of $\mathcal{A}$ without knowing that $\mathcal{A}$ has a faithful representation, before one can prove Theorem 3.4.1.

Chapter 6

Subsystems and independence

6.1 Subsystems

As we have seen in Section 2.3, we use a Q-algebra to describe all properties of a physical system that are of interest to us. Often, a physical system is made up of several smaller systems. And, of course, since we rarely consider the universe as a whole, any system we look at will be a subsystem of something larger. In quantum probability, we describe such subsystems with sub-$*$-algebras. Such sub-$*$-algebras may describe all aspects of our system that can be measured in a certain part of space, or that refer to one particular particle, or physical quantity, etc.
Thus, if $\mathcal{A}$ is a Q-algebra and $\mathcal{B} \subset \mathcal{A}$ is a sub-$*$-algebra, then we may interpret $\mathcal{B}$ as a subsystem of $\mathcal{A}$. A partition of the identity $\{P_1, \ldots, P_n\}$ such that $P_i \in \mathcal{B}_1$ for all i is interpreted as an ideal measurement on the subsystem $\mathcal{B}$. If ρ is a state (probability law) on $\mathcal{A}$, then the restriction of ρ to $\mathcal{B}$ describes our knowledge about $\mathcal{B}$.
If $\mathcal{A}$ is a Q-algebra and $\mathcal{D} \subset \mathcal{A}$ is some set, then we let $\alpha(\mathcal{D})$ denote the smallest sub-$*$-algebra of $\mathcal{A}$ containing $\mathcal{D}$. It is not hard to see that

$$\alpha(\mathcal{D}) := \text{span}(\{1\} \cup \{D_1 \cdots D_n : n \geq 1,\ D_i \in \mathcal{D} \text{ or } D_i^* \in \mathcal{D}\ \forall i = 1, \ldots, n\}),$$

i.e., $\alpha(\mathcal{D})$ is the linear span of all finite products of elements of $\mathcal{D}$ and their adjoints. We call $\alpha(\mathcal{D})$ the sub-$*$-algebra *generated* by $\mathcal{D}$. For example, if $\mathcal{B}_1, \mathcal{B}_2$ are sub-$*$-algebras of some larger Q-algebra $\mathcal{A}$, then $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$ is the smallest sub-$*$-algebra containing both $\mathcal{B}_1$ and $\mathcal{B}_2$.
In this section, we will in particular be interested in the case when subsystems that are independent, i.e., when measurements on one subsystem give no information about the other.

Recall from Section 2.3 that if we perform an ideal measurement $\{P_1, \ldots, P_n\}$ on a system described by a quantum probability space $(\mathcal{A}, \rho)$, then in general we perturb our system, which we describe by replacing the state ρ by the state $\rho'(A) := \sum_i \rho(P_i A P_i)$. We ask ourselves under which conditions performing a measurement on one subsystem does not perturb another subsystem.

Lemma 6.1.1 (Commuting subalgebras) *Let $\mathcal{A}$ be a Q-algebra and let $\mathcal{B}_1, \mathcal{B}_2$ be sub-$*$-algebras of $\mathcal{A}$. Then the following are equivalent:*

(i)	$\sum_{i=1}^n \rho(P_{2,i} P_1 P_{2,i}) = \rho(P_1)$	$\forall P_1 \in \mathcal{B}_1$ *projection,* $\{P_{2,1}, \ldots, P_{2,n}\} \subset \mathcal{B}_2$ *partition of the identity,* ρ *state on* $\mathcal{A}$,
(ii)	$P_1 P_2 = P_2 P_1$	$\forall P_1 \in \mathcal{B}_1,\ P_2 \in \mathcal{B}_2,\ P_1, P_2$ *projections,*
(iii)	$B_1 B_2 = B_2 B_1$	$\forall B_1 \in \mathcal{B}_1,\ B_2 \in \mathcal{B}_2$.

Proof (i)⇒(ii): In particular, setting $n = 2$, we have for any projections $P_1 \in \mathcal{B}_1$, $P_2 \in \mathcal{B}_2$ and for any probability ρ on $\mathcal{A}$

$$\begin{aligned} & \rho(P_2 P_1 P_2) + \rho((1 - P_2) P_1 (1 - P_2)) = \rho(P_1) \\ \Leftrightarrow\ & \rho(P_2 P_1 P_2) + \rho(P_1) + \rho(P_2 P_1 P_2) - \rho(P_1 P_2) - \rho(P_2 P_1) = \rho(P_1) \\ \Leftrightarrow\ & 2\rho(P_2 P_1 P_2) = \rho(P_1 P_2) + \rho(P_2 P_1). \end{aligned}$$

By Excersice 4.1.6, this holds for every state ρ if and only if it holds for every linear form ρ. Hence, this holds if and only if

$$2 P_2 P_1 P_2 = P_1 P_2 + P_2 P_1. \tag{6.1}$$

It follows that $P_1 P_2 = 2 P_2 P_1 P_2 - P_2 P_1$ hence $P_2(P_1 P_2) = P_2(2 P_2 P_1 P_2 - P_2 P_1) = 2 P_2 P_1 P_2 - P_2 P_1$, hence $P_2 P_1 P_2 = P_2 P_1$, which together with (6.1) implies that $P_2 P_1 = P_1 P_2$.

(ii)⇒(iii): This follows from Excersice 4.1.8.

(iii)⇒(i): Obvious, since

$$\sum_{i=1}^n \rho(P_{2,i} B_1 P_{2,i}) = \sum_{i=1}^n \rho(P_{2,i} P_{2,i} B_1) = \sum_{i=1}^n \rho(P_{2,i} B_1) = \rho(1 B_1) = \rho(B_1)$$

for any $B_1 \in \mathcal{B}_1$ and any partition of the identity $\{P_{2,1}, \ldots, P_{2,n}\} \subset \mathcal{B}_2$. ∎

If $\mathcal{B}_1$ and $\mathcal{B}_2$ are sub-$*$-algebras that commute with each other, then performing a measurement on $\mathcal{B}_1$ does not disturb $\mathcal{B}_2$, and vice versa. Thus, it should be possible to do *simultaneous measurements* on $\mathcal{B}_1$ and $\mathcal{B}_2$. Indeed, if $\{P_1, \ldots, P_n\}$

and $\{Q_1, \ldots, Q_m\}$ are ideal measurements such that $P_i \in \mathcal{B}_1$ and $Q_j \in \mathcal{B}_2$ for each i, j, then since $\mathcal{B}_1$ and $\mathcal{B}_2$ commute with each other, it is easy to see that

$$\{P_i Q_j : 1 \leq i \leq n,\ 1 \leq j \leq m\}$$

is an ideal measurement (partition of the identity). We interpret this as a measurement that carries out $\{P_1, \ldots, P_n\}$ and $\{Q_1, \ldots, Q_m\}$ simultaneously, i.e., at some point in time we perform $\{P_1, \ldots, P_n\}$ and at some point in time we perform $\{Q_1, \ldots, Q_m\}$; the order doesn't matter. If P, Q are projections that commute with each other, then we interpret PQ as the simultaneous observation of both P and Q. Note that for any state ρ, one has

$$\rho(PQ) = \frac{\rho(QPQ)}{\rho(Q)}\rho(Q) = \rho(P|Q)\rho(Q),$$

which is a well-known formula from classical probability. If P and Q do not commute, then PQ is not a projection, so we say that simultaneous measurements with noncommuting observations are not possible. In this case, $\rho(P|Q)\rho(Q)$ is still well-defined and can be interpreted as the probability of first doing the observation Q and then P, which may be different from $\rho(Q|P)\rho(P)$ (first P, then Q).

6.2 Joint measurement

We recall from Section 2.3 that in quantum mechanics, a (real-valued) *observable* is described by a hermitian operator A. The *spectrum* $\sigma(A)$ of such a hermitian operator is a finite[1] subset of the real line. The spectrum, and in fact the whole *spectral decomposition*

$$A = \sum_{\lambda \in \sigma(A)} P_\lambda$$

of such a hermitian operator are representation-independent. We interpret the partition of the identity $\{P_\lambda : \lambda \in \sigma(A)\}$ as an *ideal measurement* of the observable A, where P_λ corresponds to the *observation* that the observable A assumes the value λ. In view of this, and to stay close to classical probabilistic notation, we will sometimes use the notation

$$\{A = \lambda\} := P_\lambda$$

to denote the observation (projection operator) that A assumes the value λ. We leave the following simple fact as an excercise to the reader.

[1]The fact that $\sigma(A)$ is finite follows from our assumption that all spaces are finite dimensional. If one allows infinite dimensional spaces, then the spectrum of a self-adjoint operator may be an infinite, even uncountable subset of $\mathbb{R}$.

Exercise 6.2.1 Let $\mathcal{A}$ be a Q-algebra and let $B, C \in \mathcal{A}$ be hermitian operators. Then B and C commute (i.e., $BC = CB$) if and only if for each $b \in \sigma(B)$ and $c \in \sigma(C)$, the projection operators $\{B = b\}$ and $\{C = c\}$ commute (i.e., $\{B = b\}\{C = c\} = \{C = c\}\{B = b\}$).

Lemma 6.1.1 and Excercise 6.2.1 show that if B, C are hermitian operators in some Q-algebra $\mathcal{A}$, then the following statements are equivalent:

1° B and C commute.

2° For any state ρ on $\mathcal{A}$, performing an ideal measurement on B does not change any of the probabilities $\rho(\{C = c\})$.

In view of this, if (and only if) B and C are commuting observables, then it does not matter in which order we measure them. For each $b \in \sigma(B)$ and $c \in \sigma(C)$, We may interpret the projection (!) operator

$$\{B = b\}\{C = c\} = \{C = c\}\{B = b\} =: \{(B, C) = (b, c)\}$$

as the observation that B assumes the value b and simultaneously, C assumes the value c. In general, it may happen that $\{B = b\}\{C = c\} = 0$. Let us write

$$\sigma(B, C) := \big\{(b, c) : b \in \sigma(B),\ c \in \sigma(C),\ \{B = b\}\{C = c\} \neq 0\big\}.$$

Then we may interpret the partition of the identity

$$\big\{\{(B, C) = (b, c)\} : (b, c) \in \sigma(B, C)\big\}$$

as a *joint measurement* of the observables B and C. More generally, we may interpret any finite sequence $(B_1, \ldots, B_n)$ of mutually commuting hermitian operators as a *vector-valued* observable that can assume values in $\sigma(B_1, \ldots, B_n) \subset \mathbb{R}^n$. Such vector-valued observables are quite common in quantum mechanics. For example, the position and momentum of a single particle are observables taking values in $\mathbb{R}^3$.

More generaly, for collections of (mutually) commuting normal operators, one can define a generalization of the *functional calculus*. If $A_1, \ldots, A_n$ are (mutually) commuting normal operators and $f : \mathbb{C}^n \to \mathbb{C}$ is any function, then one defines a normal operator $f(A_1, \ldots, A_n)$ by

$$f(A_1, \ldots, A_n) := \sum_{(a_1, \ldots, a_n) \in \sigma(A_1, \ldots, A_n)} f(a_1, \ldots, a_n)\{(A_1, \ldots, A_n) = (a_1, \ldots, a_n)\}$$

where we define $\{(A_1, \ldots, A_n) = (a_1, \ldots, a_n)\}$ and $\sigma(A_1, \ldots, A_n)$ in the same way as for hermitian operators. The next excercise shows that this functional calculus yields the 'right' answers if applied to linear functions or the product function.

Exercise 6.2.2 Fix $\lambda_1, \lambda_2 \in \mathbb{C}$. Define $f : \mathbb{C}^2 \to \mathbb{C}$ and $g : \mathbb{C}^2 \to \mathbb{C}$ by $f(a_1, a_2) := \lambda_1 a_1 + \lambda_2 a_2$ and $g(a_1, a_2) = a_1 a_2$. Let A_1, A_2 be commuting normal operators. Show that $f(A_1, A_2) = \lambda_1 A_1 + \lambda_2 A_2$ and $g(A_1, A_2) = A_1 A_2$.

The following excercise demonstrates that if A, B are commuting hermitian operators, which we interpret as real-valued observables, then we may interpret AB as an observable that assumes the value ab if A assumes the value a and B assumes the value b.

Exercise 6.2.3 Let A, B be commuting observables. Show that AB is a hermitian operator and that

$$\{AB = c\} = \sum_{\substack{(a,b)\in\sigma(A,B)\\ ab=c}} \{(A, B) = (a, b)\},$$

and hence, for any state ρ,

$$\rho(\{AB = c\}) = \sum_{\substack{(a,b)\in\sigma(A,B)\\ ab=c}} \rho(\{(A, B) = (a, b)\}).$$

Concluding, we may summarize the present section by saying that *joint measurement of two or more observables is well-defined if and only if these observables commute.*

6.3 Independence

By Lemma 6.1.1, performing a measurement on a sub-$*$-algebras $\mathcal{B}_1$ does not have any effect on a sub-$*$-algebra $\mathcal{B}_2$ if and only if $\mathcal{B}_1$ and $\mathcal{B}_2$ commute with each other. We now ask under which circumstances these subsystems are independent, i.e., doing an observation on one subsystem gives no information about the other subsystem. Recall that if in some ideal measurement we do the observation P, we must describe our new knowledge about the system with the conditioned probability law $\tilde{\rho} = \rho(\cdot|P)$ defined by

$$\rho(A|P) := \frac{\rho(PAP)}{\rho(P)} \qquad (A \in \mathcal{A}).$$

Lemma 6.3.1 (Independent subalgebras) *Let $\mathcal{A}$ be a Q-algebra and let $\mathcal{B}_1, \mathcal{B}_2$ be sub-$*$-algebras of $\mathcal{A}$ that commute with each other. Then the following are equivalent:*

(i) $\rho(P_1|P_2) = \rho(P_1)$ *for all projections $P_1 \in \mathcal{B}_1$, $P_2 \in \mathcal{B}_2$ with $\rho(P_2) \neq 0$.*

(ii) $\rho(B_1B_2) = \rho(B_1)\rho(B_2)$ $\forall B_1 \in \mathcal{B}_1,\ B_2 \in \mathcal{B}_2$.

Proof Since $\mathcal{B}_1$ and $\mathcal{B}_2$ commute, $\rho(P_1|P_2) = \rho(P_2P_1P_2) = \rho(P_1P_2P_2) = \rho(P_1P_2)$, so (i) is equivalent to

$$\rho(P_1P_2) = \rho(P_1)\rho(P_2) \tag{6.2}$$

for all projections $P_1 \in \mathcal{B}_1$, $P_2 \in \mathcal{B}_2$ with $\rho(P_2) \neq 0$. In fact, (6.2) is automatically satisfied if $\rho(P_2) = 0$; to see this, note that since $\mathcal{B}_1$ and $\mathcal{B}_2$ commute, P_1P_2 is a projection. Now $P_1P_2 \leq P_2$, hence $\rho(P_1P_2) \leq \rho(P_2) = 0$. Thus, (i) holds if and only if (6.2) holds for all projections $P_1 \in \mathcal{B}_1$, $P_2 \in \mathcal{B}_2$. Since the Q-algebras $\mathcal{B}_1, \mathcal{B}_2$ are spanned by their projections (Excercise 4.1.8), this is equivalent to (ii). ∎

If $\mathcal{B}_1, \mathcal{B}_2$ are sub-$*$-algebras of some larger Q-algebra $\mathcal{A}$, and $\mathcal{B}_1$ and $\mathcal{B}_2$ commute with each other, then we observe that

$$\alpha(\mathcal{B}_1 \cup \mathcal{B}_2) = \mathcal{B}_1\mathcal{B}_2,$$

where for any subsets $\mathcal{D}_1, \mathcal{D}_2$ of a Q-algebra $\mathcal{A}$ we introduce the notation

$$\mathcal{D}_1\mathcal{D}_2 := \mathrm{span}\{D_1D_2 : D_1 \in \mathcal{D}_1,\ D_2 \in \mathcal{D}_2\}.$$

Therefore, by Lemma 6.3.1 (ii), if ρ_1 and ρ_2 are states on $\mathcal{B}_1$ and $\mathcal{B}_2$, respectively, then by linearity, there exists at most one state ρ on $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$ such that $\mathcal{B}_1$ and $\mathcal{B}_2$ are independent under ρ, and the restrictions of ρ to $\mathcal{B}_1$ and $\mathcal{B}_2$ are ρ_1 and ρ_2, respectively. We now ask under which conditions such a state ρ exists.

Lemma 6.3.2 (Logically independent algebras) *Let $\mathcal{B}_1, \mathcal{B}_2$ be sub-$*$-algebras of some larger Q-algebra, which commute with each other. Then the following statements are equivalent:*

(i) *$P_1P_2 \neq 0$ for all projections $P_1 \in \mathcal{B}_1$ and $P_2 \in \mathcal{B}_2$ with $P_1 \neq 0$ and $P_2 \neq 0$.*

(ii) *For all states ρ_1 on $\mathcal{B}_1$ and ρ_2 on $\mathcal{B}_2$ there exists a unique state ρ on $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$ such that $\rho(B_1B_2) = \rho_1(B_1)\rho_2(B_2)$ for all $B_1 \in \mathcal{B}_1,\ B_2 \in \mathcal{B}_2$.*

Proof (i)⇒(ii): We first prove the statement when ρ_1 and ρ_2 are pure states, i.e., $\rho_1 = \rho_{P_1}$ and $\rho_2 = \rho_{P_2}$, where P_1 and P_2 are minimal projections in $\mathcal{B}_1$ and $\mathcal{B}_2$, respectively. Using the fact that $\mathcal{B}_1$ and $\mathcal{B}_2$ commute, it is easy to see that P_1P_2 is a projection in $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$. Now

$$(P_1P_2)(B_1B_2)(P_1P_2) = P_1B_1P_1P_2B_2P_2 = \rho_1(B_1)\rho_2(B_2)P_1P_2 \quad (B_1 \in \mathcal{B}_1,\ B_2 \in \mathcal{B}_2).$$

Since $P_1P_2 \neq 0$, and since $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$ is spanned by elements of the form B_1B_2, there exists a function $\rho : \alpha(\mathcal{B}_1 \cup \mathcal{B}_2) \to \mathbb{C}$ such that

$$(P_1P_2)A(P_1P_2) = P_1B_1P_1P_2B_2P_2 = \rho(A)P_1P_2 \qquad (A \in \alpha(\mathcal{B}_1 \cup \mathcal{B}_2)).$$

From this it is easy to see that P_1P_2 is a minimal projection in $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$, and $\rho = \rho_{P_1P_2}$ is the pure state asociated with P_1P_2.

In the general case, when ρ_1 and ρ_2 are not pure states, we write

$$\rho_1 = \sum_{i=1}^{n_1} p_i\rho_{1,i} \quad \text{and} \quad \rho_2 = \sum_{j=1}^{n_2} q_j\rho_{2,j}$$

where the $\rho_{1,i}$ and $\rho_{2,j}$ are pure states. By what we have just proved, there exist pure states ρ_{ij} on $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$ such that $\rho_{ij}(B_1B_2) = \rho_{1,i}(B_1)\rho_{2,j}(B_2)$ for all $B_1 \in \mathcal{B}_1,\ B_2 \in \mathcal{B}_2$. Putting

$$\rho := \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} p_iq_j\rho_{ij}$$

now defines a state with the required property.

To see that (i) is also necessary for (ii), imagine that $P_1P_2 = 0$ for some nonzero projections $P_1 \in \mathcal{B}_1$ and $P_2 \in \mathcal{B}_2$. Then we can find states ρ_1, ρ_2 on $\mathcal{B}_1, \mathcal{B}_2$ such that $\rho_1(P_1) = 1$ and $\rho_2(P_2) = 1$. However, any state ρ on $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$ satisfies $0 = \rho(0) = \rho(P_1P_2) \neq \rho(P_1)\rho_2(P_2)$. ∎

Let us say that two sub-$*$-algebras $\mathcal{B}_1, \mathcal{B}_2$ of some larger Q-algebra $\mathcal{A}$ are *logically independent* if $\mathcal{B}_1$ and $\mathcal{B}_2$ commute with each other and satisfy the equivalent properties (i)–(ii) from Lemma 6.3.2. In classical probability, property (i) is sometimes called 'qualitative independence' [Ren70]. Note that this says that if no probability ρ on $\mathcal{A}$ is specified, then by doing an observation on system $\mathcal{B}_1$ we can never rule out an observation on system $\mathcal{B}_2$. If $\mathcal{B}_1, \mathcal{B}_2$ are logically independent sub-$*$-algebras of some larger Q-algebra $\mathcal{A}$, then we can give a nice description of the algebra $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$ in terms of $\mathcal{B}_1$ and $\mathcal{B}_2$.

Recall from Section 1.3 that the *tensor product* of two linear spaces $\mathcal{V}, \mathcal{W}$ is a linear space $\mathcal{V} \otimes \mathcal{W}$, equipped with a bilinear map $(\phi, \psi) \mapsto \phi \otimes \psi$ from $\mathcal{V} \times \mathcal{W}$ into $\mathcal{V} \otimes \mathcal{W}$

satisfying the equivalent conditions of Proposition 1.3.8. Such a tensor product is unique up to equivalence. Now let $\mathcal{A}_1, \mathcal{A}_2$ be Q-algebras and let $\mathcal{A}_1 \otimes \mathcal{A}_2$ be their tensor product (in the sense of linear spaces). We equip $\mathcal{A}_1 \otimes \mathcal{A}_2$ with the structure of a Q-algebra by putting

$$(A_1 \otimes A_2)(B_1 \otimes B_2) := (A_1 B_1) \otimes (A_2 B_2) \qquad (A_1, B_1 \in \mathcal{A}_1,\ A_2, B_2 \in \mathcal{A}_2)$$

and

$$(A_1 \otimes A_2)^* := (A_1^*) \otimes (A_2^*).$$

By the properties of the tensor product, these definitions extend linearly to all of $\mathcal{A}_1 \otimes \mathcal{A}_2$, making it into a Q-algebra. If $\mathcal{H}_1$ and $\mathcal{H}_2$ are representations of $\mathcal{A}_1$ and $\mathcal{A}_2$, respectively, then setting

$$(A_1 \otimes A_2)(\phi_1 \otimes \phi_2) := (A_1 \phi_1) \otimes (A_2 \phi_2) \qquad (A_1 \in \mathcal{A}_1,\ A_2 \in \mathcal{A}_2,\ \phi_1 \in \mathcal{H}_1,\ \phi_2 \in \mathcal{H}_2) \tag{6.3}$$

makes $\mathcal{H}_1 \otimes \mathcal{H}_2$ into a representation of $\mathcal{A}_1 \otimes \mathcal{A}_2$. This leads to the natural isomorphism

$$\mathcal{L}(\mathcal{H}_1) \otimes \mathcal{L}(\mathcal{H}_2) \cong \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2).$$

Note that if $\{e(1), \ldots, e(n)\}$ and $\{f(1), \ldots, f(m)\}$ are orthonormal bases of $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, then a basis for $\mathcal{L}(\mathcal{H}_1) \otimes \mathcal{L}(\mathcal{H}_2)$ is formed by all elements of the form $(|e(i)\rangle\langle e(j)|) \otimes (|f(k)\rangle\langle f(l)|)$, while a basis for $\mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ is formed by all elements of the form $|e(i) \otimes f(k)\rangle\langle e(j) \otimes f(l)|$. The dimension of both spaces is $\dim(\mathcal{H}_1)^2 \dim(\mathcal{H}_2)^2$.

Lemma 6.3.3 (Logical independence and tensor product) *If $\mathcal{B}_1, \mathcal{B}_2$ are logically independent sub-$*$-algebras of some larger Q-algebra $\mathcal{A}$, then the map*

$$B_1 B_2 \mapsto B_1 \otimes B_2$$

is a $*$-algebra isomorphism from $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$ to the tensor product algebra $\mathcal{B}_1 \otimes \mathcal{B}_2$.

Proof By Lemma 6.3.2 and Excercise 4.1.6, if l_1, l_2 are linear forms on $\mathcal{B}_1, \mathcal{B}_2$, respectively, then there exists a unique linear form l on $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$ such that $l(B_1 B_2) = l_1(B_1) l_2(B_2)$ for all $B_1 \in \mathcal{B}_1$, $B_2 \in \mathcal{B}_2$. Therefore, by Proposition 1.3.12 (iv) and Lemma 1.3.9, $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2) \cong \mathcal{B}_1 \otimes \mathcal{B}_2$. ∎

If ρ_1, ρ_2 are states (probability laws) on Q-algebras $\mathcal{A}_1, \mathcal{A}_2$, respectively, then we define a unique *product state* (product law) on $\mathcal{A}_1 \otimes \mathcal{A}_2$ by

$$(\rho_1 \otimes \rho_2)(A_1 \otimes A_2) := \rho_1(A_1) \rho_2(A_2) \qquad (A_1 \in \mathcal{A}_1,\ A_2 \in \mathcal{A}_2).$$

(This is good notation, since we can interpret $\rho_1 \otimes \rho_2$ as an element of the tensor product $\mathcal{A}_1' \otimes \mathcal{A}_2'$, where $\mathcal{A}_1'$ and $\mathcal{A}_2'$ are the dual spaces of $\mathcal{A}_1$ and $\mathcal{A}_2$, respectively.) Products of three and more Q-algebras and states are defined analoguously.

Chapter 7

Quantum paradoxes

7.1 Hidden variables

As we have already seen, the 'states' of quantum probability are something quite different from the states of classical probability. Rather, what is called a state in quantum probability corresponds to a probability law in classical probability. Pure states are probability laws that cannot be written as a mixture of other probability laws, hence a pure state ρ on a Q-algebra $\mathcal{A}$ corresponds, in a way, to maximal knowledge. If $\mathcal{A}$ is abelian, then pure states have the property that they assign probability one or zero to every observation (projection operator $P \in \mathcal{A}$). Hence, in the classical case, it is, at least theoretically, possible to know everything we want to know about a system. In Excercise 4.1.10, we have seen that in the quantum case this is not so.

Of course, in practice, even for classical systems, our knowledge is often not perfect. Especially when systems get large (e.g. contain 10^{22} molecules), it becomes impossible to know the exact value of every observable that could be of interest of us. Also, continuous observables can be measured only with limited precision. Nevertheless, it is intuitively very helpful to *imagine* that all observables *have* a value -we just don't know which one. This intuition is very much behind classical probability theory. In quantum probability, it can easily lead us astray.

Many physicists have felt uncomfortable with this aspect of quantum mechanics. Most prominently, Einstein had a deep feeling that on the grounds mentioned above, quantum theory must be incomplete. While his attempts to show that quantum mechanics is inconsistent failed, the 'Einstein-Podolsky-Rosen paradox' put forward in [EPR35] has led to a better understanding of quantum probability, and the invention of the Bell inequalities.

The absence of 'perfect knowledge' in quantum probability has prompted many

attempts to replace quantum mechanics by some more fundamental theory, in which, at least theoretically, it is possible to have extra information that allows us to predict the outcome of any experiment with certainty. Such an extended theory would be called a *hidden variable theory*, since it would involve adding some extra variables that give more information than the pure states of quantum mechanics. These extra variables can presumably never be measured so they are called *hidden* variables. It *is* possible to construct such hidden variable theories (the hidden variable theory of Bohm enjoys some popularity), but we will see that any hidden variable theory must have strange properties, making it rather unattractive.

7.2 The Kochen-Specker paradox

The Kochen-Specker paradox [KS67] shows that we run into trouble if we assume that every observable has a well-defined value. In other words, the next theorem shows that we cannot think about the observations (projection operators) from quantum probability in the same way as we think about events in classical probability.

Theorem 7.2.1 (Kochen-Specker paradox) *Let $\mathcal{H}$ be an inner product space of dimension at least 3. Then there exists a finite set $\mathcal{P}$ whose elements are projections $P \in \mathcal{L}(\mathcal{H})$, such that it is not possible to assign to every element $P \in \mathcal{P}$ a value 'true' or 'false', in such a way that in every ideal measurement $\{P_1, \ldots, P_n\}$ consisting of elements of $\mathcal{P}$, exactly one projection has the value 'true' and all others have the value 'false'.*

Remark I The essential assumption is that the value ('true' or 'false') of a projection P does not depend on the ideal measurement that it occurs in. Thus, if $\{P_1, \ldots, P_n\}$ and $\{Q_1, \ldots, Q_m\}$ are ideal measurements and $P_i = Q_j$, then P_i and Q_j should either both be 'true' or both 'false'. If one drops this assumption there is no paradox.

Remark II The fact that we run into trouble even for a finite set $\mathcal{P}$ shows that the paradox is not the result of some (perhaps unnatural) continuity or set-theoretic assumption.

Remark III The assumption that $\dim(\mathcal{H}) \geq 3$ is necessary. In the next section, when we discuss the Bell inequality, we will even need spaces of dimension at least 4. It seems that for spaces of dimension 2, there are no serious quantum paradoxes.

Proof of Theorem 7.2.1 As will be obvious from our proof, it suffices to prove the statement for the case $\dim(\mathcal{H}) = 3$. Choose an orthonormal basis $\{e(1), e(2), e(3)\}$

of $\mathcal{H}$ and consider projections of the form

$$P := |\psi\rangle\langle\psi| \quad \text{with} \quad \psi = x_1 e(1) + x_2 e(2) + x_3 e(3),$$

where $x = (x_1, x_2, x_3)$ lies on the surface of the three dimensional real unit sphere:

$$(x_1, x_2, x_3) \in S_2 := \{x \in \mathbb{R}^3 : \|x\| = 1\}.$$

Note that x and $-x$ correspond to the same projection. If three points $x, y, z \in S_2$ are orthogonal, then the corresponding projections form an ideal measurement. Therefore, we need to assign the values 'true' or 'false' to the points $x \in S_2$ in such a way that x and $-x$ always get the same value, and if three points x, y, z are orthogonal, then one of them gets the value 'true' and the other two get the value 'false'. We will show that there exists a finite set $\mathcal{P} \subset S_2$ such that it is not possible to assign the values 'true' or 'false' to the points in $\mathcal{P}$ in this way.

Note that if two points x, y are orthogonal, then by adding a third point z that is orthogonal to x and y, we see that x and y cannot both be 'true'. Therefore, it suffices to show that there exists a finite set $\mathcal{P}' \subset S_2$ such that we cannot assign values to the points in $\mathcal{P}'$ according to the following rules:

(i) Two orthogonal points are never both 'true',
(ii) Of three orthogonal points, exactly one has the value 'true'.

If we cannot assign values to $\mathcal{P}'$ according to these rules then by adding finitely many points we get a set $\mathcal{P}$ that cannot be assigned values to according to our earlier rules.

Since we are only interested in orthogonality relations between finite subsets of S_2, let us represent such subsets by a graph, where the vertices are points in S_2 and there is a bond between two vertices if the corresponding points in S_2 are orthogonal. We claim that if $x(1), x(2) \in S_2$ are close enough together, in particular, when the angle $\alpha_{1,2}$ between $x(1)$ and $x(2)$ satisfies

$$0 \leq \sin(\alpha_{1,2}) \leq \tfrac{1}{3},$$

then we can find points $x(3), \ldots, x(10)$ such that the orthogonality relations in Figure 7.1 hold.

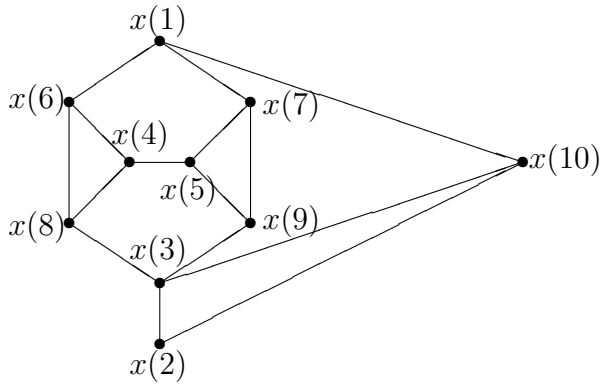

Figure 7.1: Kochen-Specker diagram

To prove this formally, take

$$\begin{aligned}
x(4) &= (1,0,0)\\
x(5) &= (0,0,1)\\
x(6) &= (0,1,\lambda)(1+\lambda^2)^{-1/2}\\
x(7) &= (1,\lambda,0)(1+\lambda^2)^{-1/2}\\
x(8) &= (0,\lambda,-1)(1+\lambda^2)^{-1/2}\\
x(9) &= (\lambda,-1,0)(1+\lambda^2)^{-1/2}\\
x(1) &= (\lambda^2,-\lambda,1)(1+\lambda^2+\lambda^4)^{-1/2}\\
x(3) &= (1,\lambda,\lambda^2)(1+\lambda^2+\lambda^4)^{-1/2},
\end{aligned}$$

where $\lambda \geq 0$ is a parameter to be determined later. It is easy to check that orthogonality relations as in Figure 7.1 hold between these points. Since $x(10)$ is orthogonal to $x(1), x(2)$, and $x(3)$, we need to take $x(2)$ in the plane spanned by $x(1)$ and $x(3)$. Denote the angle between $x(1)$ and $x(3)$ by $\alpha_{1,3}$. Then the inner product of $x(1)$ and $x(3)$ is

$$\langle x(1)|x(3)\rangle = \cos(\alpha_{1,3}).$$

We calculate

$$\langle x(1)|x(3)\rangle = \frac{\lambda^2}{1+\lambda^2+\lambda^4},$$

which is zero for $\lambda = 0$ and $\frac{1}{3}$ for $\lambda = 1$. It is not hard to see that for $\lambda = 1$ the angle between $x(1)$ and $x(3)$ is sharp so by varying λ, we can construct the diagram in Figure 7.1 for any sharp angle $\alpha_{1,3}$ with $0 \leq \cos(\alpha_{1,3}) \leq \frac{1}{3}$. Since $x(2)$

and $x(3)$ are orthogonal, it follows that we can choose $x(2)$ for any sharp angle $\alpha_{1,2}$ between $x(1)$ and $x(2)$ with $0 \leq \sin(\alpha_{1,2}) \leq \frac{1}{3}$, as claimed.

We now claim that if orthogonality relations as in Figure 7.1 hold between points $x(1), \ldots, x(10)$, and $x(1)$ has the value 'true', then $x(2)$ must also have the value 'true'.

To prove this, assume that $x(1)$ is 'true' and $x(2)$ is 'false'. Then $x(6)$, $x(7)$, and $x(10)$ must be 'false' since they are orthogonal to $x(1)$. But then $x(3)$ must be 'true' since $x(2)$ and $x(10)$ are already 'false'. Then $x(8)$ and $x(9)$ must be 'false' since they are orthogonal to $x(3)$. Now $x(4)$ must be 'true' since $x(8)$ and $x(6)$ are already 'false' and $x(5)$ must be 'true' since $x(9)$ and $x(7)$ are already false. But $x(4)$ and $x(5)$ are orthogonal, so they are not allowed to be both 'true'. We arrive at a contradiction.

We see that if two points are close enough together, then using only finitely many other points we can argue that if one is 'true' then the other one must also be 'true'. Now choose three points x, y, z that are orthogonal to each other. Then we can choose $x(1), x(2), \ldots, x(n)$ close enough together, such that x is 'true' $\Rightarrow x(1)$ is 'true' $\Rightarrow \cdots \Rightarrow x(n)$ is 'true' $\Rightarrow y$ is 'true'. (In fact, it turns out that $n = 4$ points suffice.) In the same way, using finitely many points, we can argue that y is 'true' $\Rightarrow z$ is 'true' and z is 'true' $\Rightarrow x$ is 'true'. Since x, y, and z are orthogonal, exactly one of them must be true, so we arrive at a contradiction. (In fact, it turns out that a set $\mathcal{P}'$ with 117 points suffices. For our original set $\mathcal{P}$ we need even more points, but still finitely many.) ∎

7.3 The Bell inequality

The Kochen-Specker paradox shows that the ideal measurements of quantum mechanics cannot be interpreted as classical ideal measurements. The attribute 'ideal' is essential here: if we assume that our measurements perturb our system, i.e., if the system can react differently on different measurements, there is no paradox. In this section we discuss a 'paradox' that is more compelling, since in this case, if we want to keep our classical intuition upright, we would have to assume that a system can react on a measurement that is performed in another system -potentially very far away.

Entanglement

Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be Q-algebras and let $\mathcal{A}_1 \otimes \mathcal{A}_2$ be their tensor product. We have seen that such product algebras are used to model two logically independent subsystems of a larger physical system. The systems $\mathcal{A}_1$ and $\mathcal{A}_2$ are independent under a state (probability law) ρ if and only if ρ is of product form, $\rho = \rho_1 \otimes \rho_2$ where ρ_1, ρ_2 are states on $\mathcal{A}_1, \mathcal{A}_2$, respectively. By definition, a state ρ is *entangled* if ρ can *not* be written as a convex combination of product states, i.e., if ρ is not of the form

$$\rho = \sum_{k=1}^{n} p_k \rho_{1,k} \otimes \rho_{2,k},$$

where $\rho_{1,k}, \rho_{2,k}$ are states on $\mathcal{A}_1, \mathcal{A}_2$, respectively, and the p_k are nonnegative numbers summing up to one. In classical probability, entangled states do not exist:

Exercise 7.3.1 Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be Q-algebras and assume that $\mathcal{A}_1$ is abelian. Show that there exist no entangled states on $\mathcal{A}_1 \otimes \mathcal{A}_2$.

On the other hand, if $\mathcal{A}_1$ and $\mathcal{A}_2$ are both nonabelian, then entangled states *do* exist. To see this, it suffices to consider the case that $\mathcal{A}_1 = \mathcal{L}(\mathcal{H}_1)$ and $\mathcal{A}_2 = \mathcal{L}(\mathcal{H}_2)$ where $\mathcal{H}_1, \mathcal{H}_2$ are inner product spaces of dimension at least two. Recall that $\mathcal{L}(\mathcal{H}_1) \otimes \mathcal{L}(\mathcal{H}_2) \cong \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2)$. Let $\{e, e'\}$ be orthonormal vectors in $\mathcal{H}_1$ and let $\{f, f'\}$ be orthonormal vectors in $\mathcal{H}_2$. Define a unit vector $\psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ by

$$\psi := \tfrac{1}{\sqrt{2}}\, e \otimes f + \tfrac{1}{\sqrt{2}}\, e' \otimes f', \tag{7.1}$$

and let $\rho = \rho_\psi(A) = \langle \psi | A | \psi \rangle$ be the pure state associated with ψ. We claim that ρ cannot be written as a convex combination of product states. Since ρ is pure, it suffices to show that ρ is not a product state itself. If it were, it would have to be the product of its *marginals* ρ_1, ρ_2. Here ρ_1 is the state on $\mathcal{A}_1$ defined by

$$\begin{aligned}
\rho_1(A_1) &= \langle \psi | A_1 \otimes 1 | \psi \rangle \\
&= \tfrac{1}{2}\langle e \otimes f | A_1 \otimes 1 | e \otimes f \rangle + \tfrac{1}{2}\langle e' \otimes f' | A_1 \otimes 1 | e' \otimes f' \rangle \\
&\quad + \tfrac{1}{2}\langle e \otimes f | A_1 \otimes 1 | e' \otimes f' \rangle + \tfrac{1}{2}\langle e' \otimes f' | A_1 \otimes 1 | e \otimes f \rangle \\
&= \tfrac{1}{2}\langle e | A_1 | e \rangle \langle f | f \rangle + \tfrac{1}{2}\langle e' | A_1 | e' \rangle \langle f' | f' \rangle + 0 + 0 \\
&= \tfrac{1}{2}\langle e | A_1 | e \rangle + \tfrac{1}{2}\langle e' | A_1 | e' \rangle \qquad (A_1 \in \mathcal{L}(\mathcal{H}_1)),
\end{aligned}$$

i.e., $\rho_1 = \frac{1}{2}\rho_e + \frac{1}{2}\rho_{e'}$. In the same way we see that $\rho_2 = \frac{1}{2}\rho_f + \frac{1}{2}\rho_{f'}$. In particular, ρ_1 and ρ_2 are not pure states! It is not hard to see that

$$\rho_1 \otimes \rho_2 = \tfrac{1}{4}\big(\rho_{e\otimes f} + \rho_{e'\otimes f} + \rho_{e\otimes f'} + \rho_{e'\otimes f'}\big)$$

is not a pure state, hence $\rho_1 \otimes \rho_2 \neq \rho$, so ρ is entangled.

The Bell inequality

The Bell inequality is a test on entanglement. If $(\mathcal{A}, \rho)$ is a quantum probability space and $P, Q \in \mathcal{A}$ are projections that commute with each other, then we define their *correlation coefficient* $c_\rho(P, Q)$ by

$$c_\rho(P, Q) := \rho(PQ) + \rho((1-P)(1-Q)) - \rho(P(1-Q)) - \rho((1-P)Q).$$

Note that since P and Q commute, we can interpret PQ as the simultaneous observation of P and Q. The next result is due to Bell [Bel64].

Theorem 7.3.2 (Bell inequality) *Let $\mathcal{B}_1, \mathcal{B}_2$ be logically independent sub-$*$-algebras of some larger Q-algebra and let ρ be a state on $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$. If ρ is not entangled, then for any projections $P_1, P_1' \in \mathcal{B}_1$ and $P_2, P_2' \in \mathcal{B}_2$, one has*

$$|c_\rho(P_1, P_2) + c_\rho(P_1', P_2) + c_\rho(P_1, P_2') - c_\rho(P_1', P_2')| \leq 2. \tag{7.2}$$

Proof We first prove the inequality for product states. Set

$$S_1 := 2P_1 - 1$$

and define S_1', S_2, S_2' similarly. Note that $S_1 = P_1 - (1 - P_1)$, so S_1 is a hermitian operator with spectrum $\sigma(S_1) = \{-1, +1\}$, i.e., S_1 is an observable that can take on the values ± 1, such that P_1 (resp. $1 - P_1$) corresponds to the observation that $S_1 = +1$ (resp. $S_1 = -1$). Then

$$c_\rho(P_1, P_2) = \rho(S_1 S_2),$$

etc., so if ρ is a product measure, then

$$\begin{aligned} &c_\rho(P_1, P_2) + c_\rho(P_1', P_2) + c_\rho(P_1, P_2') - c_\rho(P_1', P_2') \\ &\quad = \rho(S_1 S_2) + \rho(S_1' S_2) + \rho(S_1 S_2') - \rho(S_1' S_2') \\ &\quad = \rho(S_1)\rho(S_2) + \rho(S_1')\rho(S_2) + \rho(S_1)\rho(S_2') - \rho(S_1')\rho(S_2') \\ &\quad = \rho(S_1)(\rho(S_2) + \rho(S_2')) + \rho(S_2')(\rho(S_2) - \rho(S_2')), \end{aligned}$$

so the quantity in (7.2) can be estimated by

$$|\rho(S_2) + \rho(S_2')| + |\rho(S_2) - \rho(S_2')|.$$

If $\rho(S_2) + \rho(S_2')$ and $\rho(S_2) - \rho(S_2')$ have the same sign, then we get $2|\rho(S_2)|$, while otherwise we get $2|\rho(S_2')|$. At any rate, our estimate shows that the quantity in (7.2) is less or equal than 2.

More generally, if ρ is a convex combination of product states, $\rho = \sum_k p_k \rho_k$, say, then

$$\begin{aligned}&|c_\rho(P_1, P_2) + c_\rho(P_1', P_2) + c_\rho(P_1, P_2') - c_\rho(P_1', P_2')| \\ &\quad \le \sum_k p_k \big| c_{\rho_k}(P_1, P_2) + c_{\rho_k}(P_1', P_2) + c_{\rho_k}(P_1, P_2') - c_{\rho_k}(P_1', P_2') \big| \le 2\end{aligned}$$

by what we have just proved. ∎

We next show that entangled states can violate the Bell inequality. We will basically use the same entangled state as in (7.1), which we interpret in terms of two polarized photons. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two-dimensional inner product spaces with orthonormal bases $\{e(1), e(2)\}$ and $\{f(1), f(2)\}$, respectively. For $\gamma \in [0, \pi)$, define $\eta_\gamma \in \mathcal{H}_1$ and $\zeta_\gamma \in \mathcal{H}_2$ by

$$\eta_\gamma := \cos(\gamma) e(1) + \sin(\gamma) e(2) \quad \text{and} \quad \zeta_\gamma := \cos(\gamma) f(1) + \sin(\beta) f(2).$$

Set $P_\gamma := |\eta_\gamma\rangle\langle\eta_\gamma|$ and $Q_\beta := |\zeta_\beta\rangle\langle\eta_\beta|$. For each $\gamma, \tilde{\gamma}$ we may interpret $\{P_\gamma, P_{\gamma+\pi/2}\}$ and $\{Q_\gamma, Q_{\tilde{\gamma}+\pi/2}\}$ as an ideal measurements of the polarization of our first photon and second photon, respectively, in the directions γ and $\tilde{\gamma}$ (see Section 2.3). We prepare our system in the entangled state

$$\psi := \tfrac{1}{\sqrt{2}}\, e(1) \otimes f(1) + \tfrac{1}{\sqrt{2}}\, e(2) \otimes f(2).$$

We claim that for any γ,

$$\psi = \tfrac{1}{\sqrt{2}}\, \eta_\gamma \otimes \zeta_\gamma + \tfrac{1}{\sqrt{2}}\, \eta_{\gamma+\pi/2} \otimes \zeta_{\gamma+\pi/2}. \tag{7.3}$$

Note that this says that if we measure the polarization of both photons along the same direction, we will always find that both photons are polarized in the same way! To see this, we observe that

$$\begin{aligned}\eta_\gamma \otimes \zeta_\gamma &= (\cos(\gamma) e(1) + \sin(\gamma) e(2)) \otimes (\cos(\gamma) f(1) + \sin(\gamma) f(2)) \\ &= \cos(\gamma)^2\, e(1) \otimes f(1) + \sin(\gamma)^2\, e(2) \otimes f(2) \\ &\quad + \cos(\gamma)\sin(\gamma) e(1) \otimes f(2) + \sin(\gamma)\cos(\gamma) e(2) \otimes f(1)\end{aligned}$$

and

$$\begin{aligned}\eta_{\gamma+\pi/2} \otimes \zeta_{\gamma+\pi/2} &= (-\sin(\gamma) e(1) + \cos(\gamma) e(2)) \otimes (-\sin(\gamma) f(1) + \cos(\gamma) f(2)) \\ &= \sin(\gamma)^2\, e(1) \otimes f(1) + \cos(\gamma)^2\, e(2) \otimes f(2) \\ &\quad - \sin(\gamma)\cos(\gamma) e(1) \otimes f(2) - \cos(\gamma)\sin(\gamma) e(2) \otimes f(1).\end{aligned}$$

Adding both expressions and dividing by $\sqrt{2}$ we arrive at (7.3).
The probability of finding one photon polarized in the direction γ and the other photon in the direction $\tilde{\gamma}$ is given by

$$\begin{aligned}\rho_\psi(P_\gamma \otimes Q_{\tilde{\gamma}}) &= \rho_\psi(P_0 \otimes Q_{\tilde{\gamma}-\gamma}) \\ &= \tfrac{1}{2}\langle e(1)\otimes f(1)|P_0 \otimes Q_{\tilde{\gamma}-\gamma}|e(1)\otimes f(1)\rangle \\ &\quad + \tfrac{1}{2}\langle e(2)\otimes f(2)|P_0 \otimes Q_{\tilde{\gamma}-\gamma}|e(2)\otimes f(2)\rangle \\ &= \tfrac{1}{2}\langle e(1)|e(1)\rangle\langle f(1)|\zeta_{\tilde{\gamma}-\gamma}\rangle\langle\zeta_{\tilde{\gamma}-\gamma}|f(1)\rangle \\ &= \tfrac{1}{2}\cos(\tilde{\gamma}-\gamma)^2.\end{aligned}$$

(Compare Excercise 2.3.2.) Hence

$$c_{\rho_\psi}(P_\gamma \otimes 1, 1\otimes Q_{\tilde{\gamma}}) = \cos(\tilde{\gamma}-\gamma)^2 - \sin(\tilde{\gamma}-\gamma)^2 = 2\cos(\tilde{\gamma}-\gamma)^2 - 1 = \cos(2(\tilde{\gamma}-\gamma)).$$

We now check that for an appropriate choice of the angles, these correlation coefficients violate the Bell inequality (7.2). We take

$$\begin{aligned} P_1 &= P_0 \otimes 1, & P_1' &= P_{\alpha+\beta}\otimes 1, \\ P_2 &= 1\otimes Q_\alpha, & P_2' &= 1\otimes Q_{-\beta}.\end{aligned}$$

The expression in (7.2) then becomes

$$\big|\cos(2\alpha) + 2\cos(2\beta) - \cos(4\beta+2\alpha)\big|.$$

We want to maximize the expression inside the brackets. Setting the derivatives with respect to α and β equal to zero yields the equations

$$\begin{aligned} -2\sin(2\alpha) + 2\sin(4\beta+2\alpha) &= 0, \\ -4\sin(2\beta) + 4\sin(4\beta+2\alpha) &= 0.\end{aligned}$$

It follows that $\sin(2\beta) = \sin(4\beta+2\alpha) = \sin(2\alpha)$. We choose

$$\beta = \alpha.$$

The expression to be maximized then becomes

$$3\cos(2\alpha) - \cos(6\alpha).$$

Differentiating and setting equal to zero yields

$$-6\sin(2\alpha) + 6\sin(6\alpha) = 0 \;\Rightarrow\; \sin(2\alpha) = \sin(6\alpha).$$

Setting $z = e^{i2\alpha}$, we need to solve

$$\begin{aligned} \tfrac{1}{2i}\big(e^{i2\alpha} - e^{-i2\alpha}\big) &= \tfrac{1}{2i}\big(e^{i6\alpha} - e^{-i6\alpha}\big) \\ \Leftrightarrow \quad z - z^{-1} &= z^3 - z^{-3} \\ \Leftrightarrow \quad z^6 - z^4 + z^2 - 1 &= 0. \end{aligned}$$

Setting $y = z^2 = e^{i4\alpha}$, we obtain the cubic equation

$$y^3 - y^2 + y - 1 = 0.$$

We know that $y = e^{i20} = 1$ is a trivial solution, so factorising this out we get

$$(y-1)(y^2+1) = 0,$$

which has nontrivial solutions $y = \pm i = e^{\pm i\pi/2}$. Therefore, the maximum we are interested in occurs at $\alpha = \frac{1}{8}\pi$. The expression in (7.2) then becomes

$$3\cos(\tfrac{1}{4}\pi) - \cos(\tfrac{3}{4}\pi) = 3\tfrac{1}{\sqrt{2}} - -\tfrac{1}{\sqrt{2}} = 2\sqrt{2} \approx 2.82847,$$

which is indeed larger than 2, the bound from the Bell inequality. Correlations between single photons passing through prismas can be measured, and this violation of the Bell inequality has been verified experimentally [Red87, CS78].

7.4 The GHZ paradox

In classical probability theory, every probability law can be written as the convex combination of 'precise' states, where the values of all observables are known. In particular, on a product space, such precise states are product states, so there can be no entanglement. The Bell inequality is a test on entanglement, but it is only a statistical test, which requires us to measure certain probabilities that lie strictly between zero and one. The appeal of the Kochen-Specker paradox is that it talks about events that have probability zero or one only. The Greenberger-Horne-Zeilinger paradox [G–Z90] gives us a test on entanglement that involves only events that have probability zero or one. In fact, things are happening with *certainty* here that in classical probability can at most have probability 3/4. A slight complication that we have to deal with to achieve this is that we need three subsystems, rather than two.

Theorem 7.4.1 (GHZ paradox) *Let $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ be logically independent sub-$*$-algebras of some larger Q-algebra and let ρ be a state on $\mathcal{B}_1\mathcal{B}_2\mathcal{B}_3$. For each*

$i = 1, 2, 3$, let $S_i, S_i' \in \mathcal{B}_i$ be hermitian operators with spectrum $\sigma(S_i) = \sigma(S_i') = \{-1, 1\}$. If ρ is not entangled, then

$$\begin{aligned}\rho(\{S_1S_2'S_3' = 1\}) + \rho(\{S_1'S_2S_3' = 1\}) \\ +\rho(\{S_1'S_2'S_3 = 1\}) + \rho(\{S_1S_2S_3 = -1\}) \leq 3.\end{aligned} \tag{7.4}$$

On the other hand, if each $\mathcal{B}_i$ is of the form $\mathcal{B}_i \cong \mathcal{L}(\mathcal{H})$ where $\mathcal{H}$ is a 2-dimensional inner product space, then there exists a state ρ on $\mathcal{B}_1\mathcal{B}_2\mathcal{B}_3$ such that

$$\begin{aligned}\rho(\{S_1S_2'S_3' = 1\}) + \rho(\{S_1'S_2S_3' = 1\}) \\ +\rho(\{S_1'S_2'S_3 = 1\}) + \rho(\{S_1S_2S_3 = -1\}) = 4.\end{aligned} \tag{7.5}$$

Remark 1 Since the algebras $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ are logically independent, the operators S_1, S_2', S_3' all commute with each other. It follows that their product $S_1S_2'S_3'$ is also a hermitian operator. If S_1, S_2', S_3' assume values s_1, s_2', s_3', then the observable $S_1S_2'S_3'$ assumes the value $s_1s_2's_3'$. In fact, by Excercise 6.2.3, we have that

$$\begin{aligned}\{S_1S_2'S_3' = 1\} = \{S_1 = 1\}\{S_2' = 1\}\{S_3' = 1\} + \{S_1 = 1\}\{S_2' = -1\}\{S_3' = -1\} \\ +\{S_1 = -1\}\{S_2' = 1\}\{S_3' = -1\} + \{S_1 = -1\}\{S_2' = -1\}\{S_3' = 1\}.\end{aligned}$$

The same applies to the operators $S_1'S_2S_3'$, $S_1'S_2'S_3$, and $S_1S_2S_3$, so in this way, we can write the left-hand side of (7.4) as the sum of 16 observations of the form $P_1P_2P_3$, where $P_i \in \mathcal{B}_i$ is a projection. Note that by our assumption that the algebras $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ are logically independent, all of these projections are nonzero.

Remark 2 Since (even in quantum probability) probabilities can never be larger than one, formula (7.5) simply says that each of the four observations there has probability one. If we think classically, and presume that the observables $S_1, S_2, S_3, S_1', S_2', S_3'$ have some hidden 'true' values $s_1, s_2, s_2, s_1', s_2', s_3'$ in $\{-1, 1\}$, then this is clearly nonsense, since it implies that with probability one

$$(s_1s_2's_3')(s_1's_2s_3')(s_1's_2's_3) = 1 \neq -1 = (s_1s_2s_3).$$

Remark 3 The GHZ paradox has been experimentally tested [B–Z99].

Proof of Theorem 7.4.1 We start by constructing a state ρ for which (7.5) holds. Let $\mathcal{H}$ be a two-dimensional inner product space with orthonormal basis $\{e(1), e(2)\}$. Let $\rho = \rho_\psi$ be the pure state desccribed by the state vector

$$\psi := \tfrac{1}{\sqrt{2}}\big(e(1) \otimes e(1) \otimes e(1) + e(2) \otimes e(2) \otimes e(2)\big). \tag{7.6}$$

Let $S, S' \in \mathcal{L}(\mathcal{H})$ be the hermitian operators given with respect to the basis $\{e(1), e(2)\}$ by the matrices

$$S := \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad S' := \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

It is easy to check that $S^2 = 1 = S'^2$, so $\sigma(S)$ and $\sigma(S)$ are subsets of $\{-1, 1\}$. Since S and S' are clearly not equal to the operators ± 1, we conclude that $\sigma(S) = \{-1, 1\} = \sigma(S)$. It is easy to check that

$$SS' = -S'S,$$

i.e., the operators S and S' *anticommutate*. We now represent the product algebra $\mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{H})$ in the standard way on the product space $\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}$ and set

$$S_1 := S \otimes 1 \otimes 1, \quad S_2 := 1 \otimes S \otimes 1, \quad \text{and} \quad S_3 := 1 \otimes 1 \otimes S,$$

and similarly

$$S'_1 := S' \otimes 1 \otimes 1, \quad S'_2 := 1 \otimes S' \otimes 1, \quad \text{and} \quad S'_3 := 1 \otimes 1 \otimes S'.$$

Then $\rho_\psi(S_1 S'_2 S'_3) = \langle \psi | S_1 S'_2 S'_3 | \psi \rangle$ is given by

$$\begin{aligned}
&\tfrac{1}{2}\langle e(1) \otimes e(1) \otimes e(1) + e(2) \otimes e(2) \otimes e(2)| \\
&\quad \cdot S \otimes S' \otimes S' | e(1) \otimes e(1) \otimes e(1) + e(2) \otimes e(2) \otimes e(2)\rangle \\
&= \tfrac{1}{2}\langle e(1) \otimes e(1) \otimes e(1) + e(2) \otimes e(2) \otimes e(2)| \\
&\quad \cdot |(Se(1)) \otimes (S'e(1)) \otimes (S'e(1)) + (Se(2)) \otimes (S'e(2)) \otimes (S'e(2))\rangle \\
&= \tfrac{1}{2}\langle e(1)|Se(1)\rangle\langle e(1)|S'e(1)\rangle\langle e(1)|S'e(1)\rangle \\
&\quad + \tfrac{1}{2}\langle e(1)|Se(2)\rangle\langle e(1)|S'e(2)\rangle\langle e(1)|S'e(2)\rangle \\
&\quad + \tfrac{1}{2}\langle e(2)|Se(1)\rangle\langle e(2)|S'e(1)\rangle\langle e(2)|S'e(1)\rangle \\
&\quad + \tfrac{1}{2}\langle e(2)|Se(2)\rangle\langle e(2)|S'e(2)\rangle\langle e(2)|S'e(2)\rangle.
\end{aligned}$$

Here

$$\langle e(1)|Se(1)\rangle = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \middle| \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \middle| \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\rangle = 0,$$

$$\langle e(1)|Se(2)\rangle = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \middle| \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \middle| \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\rangle = -1,$$

$$\langle e(2)|Se(2)\rangle = \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \middle| \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \middle| \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\rangle = 0,$$

and

$$\langle e(1)|S'e(1)\rangle = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \middle| \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \middle| \begin{pmatrix} 0 \\ -i \end{pmatrix} \right\rangle = 0,$$

$$\langle e(1)|S'e(2)\rangle = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \middle| \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \middle| \begin{pmatrix} i \\ 0 \end{pmatrix} \right\rangle = i,$$

$$\langle e(2)|S'e(2)\rangle = \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \middle| \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \middle| \begin{pmatrix} i \\ 0 \end{pmatrix} \right\rangle = 0.$$

Since S and S' are hermitian, $\langle e(2)|S|e(1)\rangle = \langle e(1)|S|e(2)\rangle^* = (-1)^* = -1$ and $\langle e(2)|S'|e(1)\rangle = \langle e(1)|S'|e(2)\rangle^* = i^* = -i$, so we conclude that our earlier expression for $\rho_\psi(S_1S_2'S_3')$ is given by

$$\rho_\psi(S_1S_2'S_3') = 0 + \tfrac{1}{2}(-1)\cdot i\cdot i + \tfrac{1}{2}(-1)\cdot(-i)\cdot(-i) + 0 = 1.$$

By the symmetry of the state ψ, it is now clear that

$$\rho_\psi(S_1S_2'S_3') = \rho_\psi(S_1'S_2S_3') = \rho_\psi(S_1'S_2'S_3) = 1.$$

On the other hand, the same calculations as above give for $S_1S_2S_3$

$$\rho_\psi(S_1S_2S_3') = 0 + \tfrac{1}{2}(-1)^3 + \tfrac{1}{2}(-1)^3 + 0 = -1.$$

Since $S_1S_2'S_3'$ is an $\{-1,1\}$-valued observable, the fact that the expected value

$$\rho_\psi(S_1S_2'S_3') = \rho(\{S_1S_2'S_3' = 1\}) - \rho(\{S_1S_2'S_3' = -1\})$$

equals one implies that $S_1S_2'S_3'$ assumes the value 1 with probability one. The same arguments apply to S_1', S_2, S_3', S_1', S_2', S_3, and S_1, S_2, S_3, so we conclude that (7.5) holds.

Because the explcit calculations above are rather involved, we now give some algebraic arguments (inspired by [Gil06]) why (7.5) should perhaps have been expected. We start by observing that the hermitian operators

$$(S_1S_2'S_3'), \quad (S_1'S_2S_3'), \quad (S_1'S_2'S_3), \quad \text{and} \quad (S_1S_2S_3)$$

all commute with each other. Indeed, by the anticommutation relation between S and S'

$$\begin{aligned}(S_1S_2'S_3')(S_1'S_2S_3') &= (S_1S_1')(S_2'S_2)(S_3')^2 \\ &= (-1)^2(S_1'S_1)(S_2S_2')(S_3')^2 = (S_1'S_2S_3')(S_1S_2'S_3').\end{aligned}$$

Since we always have to permute the order of two primed and unprimed operators, we see in this way that (S_1', S_2, S_3'), (S_1', S_2', S_3), and (S_1, S_2, S_3) mutually commute. It follows that we can find an orthonormal basis that diagonalizes these operators simultaneously. In particular, we can find some vector ψ and constants $s_{011}, s_{101}, s_{110}, s_{000}$ such that

$$(S_1 S_2' S_3')\psi = s_{011}\psi, \quad (S_1' S_2 S_3')\psi = s_{101}\psi,$$
$$(S_1' S_2' S_3)\psi = s_{110}\psi, \quad \text{and} \quad (S_1 S_2 S_3)\psi = s_{000}\psi.$$

But, again using the anticommutation relation between S and S' as well as the fact that $S^2 = S'^2 = 1$, we see that

$$\begin{aligned}(S_1 S_2' S_3')(S_1' S_2 S_3')(S_1' S_2' S_3) &= (S_1 S_1' S_1')(S_2' S_2 S_2')(S_3' S_3' S_3) \\ &= -(S_1 S_1' S_1')(S_2 S_2' S_2')(S_3 S_3' S_3') = -(S_1 S_2 S_3),\end{aligned}$$

which implies that

$$s_{011} s_{101} s_{110} = -s_{000}.$$

Our earlier, explicit calculations show that in fact it is possible to choose $s_{011} = s_{101} = s_{110} = 1$ and $s_{000} = -1$, and moreover tell us that this is realized for the nice entangled state in (7.6).

We still have to show that states that are not entangled must satisfy (7.4). By linearity, it suffices to show that product states satisfy (7.4). Define probability laws μ_i, μ_i' $(i = 1, 2, 3)$ on $\{-1, 1\}$ by

$$\mu_i(s) := \rho(\{S_i = s\}) \quad \text{and} \quad \mu_i'(s) := \rho(\{S_i' = s\}) \quad (s \in \{-1, 1\}).$$

As explained in Remark 1, the left-hand side of (7.4) is really the sum of 16 probabilities of the form

$$\rho(\{(S_1, S_2', S_3') = (1, -1, -1)\}),$$

etc., where there is always an even number of primed variables and an even or odd number of $+1$'s and -1's depending on which observation we are considering. Since ρ is a product state

$$\rho(\{(S_1, S_2', S_3') = (s_1, s_2, s_3)\}) = \mu_1(s_1)\mu_2'(s_2)\mu_3'(s_3).$$

Let $T_1, T_2, T_3, T_1', T_2', T_3'$ be random variables on a classical probability space, all independent of each other, with laws $\mu_1, \mu_2, \mu_3, \mu_1', \mu_2', \mu_3'$, respectively. Then

$$\rho(\{(S_1, S_2', S_3') = (s_1, s_2, s_3)\}) = \mathbb{P}[(T_1, T_2', T_3') = (s_1, s_2, s_3)],$$

etc. So it suffices to show that in classical probability,

$$\mathbb{P}[T_1T_2'T_3'=1]+\mathbb{P}[T_1'T_2T_3'=1]+\mathbb{P}[T_1'T_2'T_3=1]+\mathbb{P}[T_1T_2T_3=-1]\leq 3.$$

Since the first three events imply that

$$(T_1T_2T_3)=(T_1T_2'T_3')(T_1'T_2T_3')(T_1'T_2'T_3)=1\cdot 1\cdot 1=1,$$

at most three of the events $\{T_1T_2'T_3'=1\}$, $\{T_1'T_2T_3'=1\}$, $\{T_1'T_2'T_3=1\}$, and $\{T_1T_2T_3=-1\}$ can be true at the same time, so

$$\mathbb{E}\big[1_{\{T_1T_2'T_3'=1\}}+1_{\{T_1'T_2T_3'=1\}}+1_{\{T_1'T_2'T_3=1\}}+1_{\{T_1T_2T_3=-1\}}\big]\leq 3.$$

∎

Exercise 7.4.2 For the entangled state $\rho=\rho_\psi$ from Theorem 7.4.1, calculate the marginal $\rho_{12}(A_1\otimes A_2):=\rho(A_1\otimes A_2\otimes 1)$ $(A_1\in\mathcal{B}_1,\ A_2\in\mathcal{B}_2)$. Is this marginal entangled?

Exercise 7.4.3 For the entangled state $\rho=\rho_\psi$ from Theorem 7.4.1, calculate $\rho(\{S_1'S_2S_3=1\})$ and $\rho(\{S_1'S_2'S_3'=1\})$.

Exercise 7.4.4 Three astronauts, who are good friends, are separating for many years to explore different parts of space. To stay at least symbolically in contact, they have bought three equal-looking boxes, one for each of them, on which there are two buttons, a blue and a red one, and also two lights in these colors. One day, when they are far from each other, each will push one of the buttons, either the blue or the red one. The boxes are constructed so that they react only to the first button that is pushed on each box. Once a button is pushed, either the blue or red light, or both, or none, light up. Things work in such way that after all friends have chosen their buttons, if all three have pushed their red buttons, then on the three boxes in total an odd number of red lights will be lit, while if exactly one of the friends has chosen the red button, an even number of red lights will be lit. For blue buttons and blue lights, the rules are the same. Can such boxes be constructed in such a way that no signal needs to pass from one box to the other boxes once the buttons are pushed, and if yes, how?

Bell versus Tsirelson

We have seen that in classical probability theory, the quantity in (7.2) is less or equal than 2, while in quantum probability, it can be $2\sqrt{2}$. Note that a priori,

this is just a sum of four correlations, each of which could take values between -1 and 1, so it is conceivable that this quantity could be as high as 4. Nevertheless, the violation of Bell's inequality that we have found is maximal, as was proved by B. Tsirelson [Cir80]. Similarly, we have seen that in classical probability theory, and more generally for nonentangled states, the quantity in (7.4) is less or equal than 3, while in quantum probability, it can be 4.

Another way of looking at these inequalities is as follows. Imagine that we have s physical systems (separated in space), such on each system, m different ideal measurements are possible, each of which yields one of n different possible outcomes. The Bell inequality (7.2) considers the case $s = m = n = 2$, while the GHZ paradox is concerned with $s = 3$, $m = n = 2$. Numbering the systems, measurements, and outcomes in some arbitrary way, we are interested in the $(mn)^s$ conditional probabilities, say

$$p(a_1, \ldots, a_s | b_1 \ldots, b_s),$$

that experiments $b_1, \ldots, b_s \in \{1, \ldots, m\}$ yield outcomes $a_1, \ldots, a_s \in \{1, \ldots, n\}$. We are interested in the case that chosing which measurement to perform on one system does not influence probabilities on another system. For example, in the case $s = m = n = 2$, this yields the 'no signalling' requirement

$$p(1,1|1,1) + p(1,2|1,1) = p(1,1|1,2) + p(1,2|1,2),$$

which says that the conditional probability of outcome 1 given that on system 1 we perform measurement 1, does not depend on the choice of the measurement at the second system. There are other requirements coming from the fact that probabilities must be nonnegative and sum up to one. Together, these requirements define a convex set $\mathcal{P}_{\text{nosignal}}$ of functions p that assign probabilities $p(a_1, \ldots, a_s | b_1 \ldots, b_s)$ to the outcomes of different measurements.

It turns out that not all these probability functions p can arise from classical probability. More precisely, classically, we imagine that there are certain 'hidden variables' that deterministically predict the outcome of each measurement. Thus, we imagine that

$$p(a_1, \ldots, a_s | b_1 \ldots, b_s) = \sum_h P(h) p_h(a_1, \ldots, a_s | b_1 \ldots, b_s) \tag{7.7}$$

where h represents the 'hidden' variables, $P(h)$ is the probability that these hidden variables take the value h, and p_h is a function satisfying the 'no signalling' and other requirements mentioned above, such that in addition, $p_h(a_1, \ldots, a_s | b_1 \ldots, b_s)$ is either 0 or 1 for each choice of $a_1, \ldots, a_s, b_1 \ldots, b_s$. Since there are only finitely many such functions, the collection of functions p of the form (7.7) is a convex set

$\mathcal{P}_{\text{classical}}$ with finitely many extreme points, which are the functions p_h. It turns out that $\mathcal{P}_{\text{classical}}$ is strictly smaller than $\mathcal{P}_{\text{nosignal}}$. Here, an essential assumption is that the functions p_h also satisfy our 'no signalling' requirements. If we allow hidden variables to communicate at a distance (possibly with a speed larger than the speed of light), then there is no problem.
'Interesting' faces of $\mathcal{P}_{\text{classical}}$ correspond to inequalities that are not satisfied by general elements of $\mathcal{P}_{\text{nosignal}}$. In fact, the Tsirelson inequalities show that $\mathcal{P}_{\text{quantum}}$, the quantum analogue of $\mathcal{P}_{\text{classical}}$, is also not equal to $\mathcal{P}_{\text{nosignal}}$. The geometric structure of these convex sets is still very much a topic of research, see [Gil06]. Another interesting question (that I do not know the answer to) is whether there exist good, consistent probability theories that violate the Tsirelson inequalities.

Chapter 8

Operations

8.1 Completely positive maps

Let $\mathcal{A}$ be a Q-algebra and let $\mathcal{A}'_{\text{prob}}$ denote the space of all states on $\mathcal{A}$. As we know, a state $\rho \in \mathcal{A}'_{\text{prob}}$ describes incomplete knowledge about the physical system $\mathcal{A}$. We will be interested in *operations* on $\mathcal{A}'_{\text{prob}}$, i.e., we want to know how ρ can change due to the effects of our physical interference with the system $\mathcal{A}$. In general, such an operation will be described by a map $f : \mathcal{A}'_{\text{prob}} \to \mathcal{A}'_{\text{prob}}$ that has the following interpretation: If our knowledge about the system before we performed the operation was described by the state ρ, then our knowledge after we have performed the operation is described by $f(\rho)$. A natural requirement on f is that it be linear, in the sense that

$$f(p\rho_1 + (1-p)\rho_2) = pf(\rho_1) + (1-p)f(\rho_2) \qquad (\rho_1, \rho_2 \in \mathcal{A}'_{\text{prob}},\ 0 \leq p \leq 1). \quad (8.1)$$

This says that if before we performed our operation f, our knowledge about the system is with probability p described by ρ_1 and with probability $1-p$ by ρ_2, then after we have performed the operation f, the system is with probability p described by $f(\rho_1)$ and with probability $1-p$ by $f(\rho_2)$.

At first, it might seem that nothing more can be said about f and that any map $f : \mathcal{A}'_{\text{prob}} \to \mathcal{A}'_{\text{prob}}$ that is linear in the sense of (8.1) describes a legal operation on the system $\mathcal{A}$. However, it turns out that this is not the case. Assume that $\mathcal{B}$ is some other system, logically independent of $\mathcal{A}$. (Since our algebra $\mathcal{A}$ will typically not describe the whole universe, there will typically be lots of such systems!) Then we must be able to say what happens with a probability ρ on $\mathcal{A} \otimes \mathcal{B}$ when we perform our operation on $\mathcal{A}$ and do nothing with $\mathcal{B}$. For product probabilities, it is natural to require that the effect of our operation on $\mathcal{A}$, doing nothing with $\mathcal{B}$,

is to map $\rho_1 \otimes \rho_2$ to

$$F(\rho_1 \otimes \rho_2) := f(\rho_1) \otimes \rho_2.$$

Now we need that F can be extended to a map $F : (\mathcal{A} \otimes \mathcal{B})'_{\rm prob} \to (\mathcal{A} \otimes \mathcal{B})'_{\rm prob}$ that is linear in the sense of (8.1). It turns out that such a linear extension does not always exist, and this leads to a nontrivial requirement on f! As one might guess, problems occur for entangled states.

Slightly generalizing our set-up, let us consider two Q-algebras $\mathcal{A}$ and $\mathcal{B}$ and maps $f : \mathcal{A}'_{\rm prob} \to \mathcal{B}'_{\rm prob}$ that are linear in the sense of (8.1).

Lemma 8.1.1 (Linear maps acting on states) *Let $\mathcal{A}$ and $\mathcal{B}$ be Q-algebras and let $T : \mathcal{B} \to \mathcal{A}$ be a linear map satisfying*

$$\begin{array}{ll} \text{(i)} & T(B^*) = T(B)^*, \\ \text{(ii)} & B \geq 0 \;\Rightarrow\; T(B) \geq 0, \\ \text{(iii)} & T(1) = 1. \end{array} \tag{8.2}$$

Let $T' : \mathcal{A}' \to \mathcal{B}'$ be the dual of T, i.e.,

$$T'(l)(B) = l(T(B))) \qquad (l \in \mathcal{A}',\ B \in \mathcal{B}).$$

Then T' maps $\mathcal{A}'_{\rm prob}$ into $\mathcal{B}'_{\rm prob}$, and conversely, every map from $\mathcal{A}'_{\rm prob}$ into $\mathcal{B}'_{\rm prob}$ that is linear in the sense of (8.1) arises in this way.

Proof Assume that $f : \mathcal{A}'_{\rm prob} \to \mathcal{B}'_{\rm prob}$ is linear in the sense of (8.1). Since $\mathcal{A}'_{\rm prob}$ spans $\mathcal{A}'$, the map f can uniquely be extended to a linear map $\hat{f} : \mathcal{A}' \to \mathcal{B}'$. It is not hard to see that $\hat{f}$ satisfies

$$\begin{array}{ll} \text{(i)} & l \text{ real} \Rightarrow \hat{f}(l) \text{ real}, \\ \text{(ii)} & l \text{ positive} \Rightarrow \hat{f}(l) \text{ positive}, \\ \text{(iii)} & \hat{f}(l)(1) = l(1), \end{array} \tag{8.3}$$

and conversely, if a linear map $\hat{f} : \mathcal{A}' \to \mathcal{B}'$ satisfies (8.3), then $\hat{f}$ maps $\mathcal{A}'_{\rm prob}$ into $\mathcal{B}'_{\rm prob}$. Let T denote the dual of $\hat{f}$, i.e., the unique linear map $T : \mathcal{B} \to \mathcal{A}$ such that

$$\hat{f}(l)(B) = l(T(B)) \qquad (l \in \mathcal{A}',\ B \in \mathcal{B}).$$

Then it is not hard to see that the conditions (8.2) (i)–(iii) are equivalent to the conditions (8.3) (i)–(iii). ∎

Exercise 8.1.2 Check that (8.2) is equivalent to (8.3) if $\hat{f} = T'$, the dual of T.

A linear map $T : \mathcal{B} \to \mathcal{A}$ satisfying (8.2) (i) and (ii) (but not necessarily (iii)) is called *positive*.

Coming back to our earlier discussion, let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be Q-algebras, let $T : \mathcal{B} \to \mathcal{A}$ be a positive linear map, and let T' be its dual. In view of Lemma 8.1.1, we want to know if there exists a positive linear map $S : \mathcal{B} \otimes \mathcal{C} \to \mathcal{A} \otimes \mathcal{C}$ whose dual S' satisfies

$$S'(\rho_1 \otimes \rho_2) := T'(\rho_1) \otimes \rho_2 \qquad (\rho_1 \in \mathcal{A}'_{\text{prob}},\ \rho_2 \in \mathcal{C}'_{\text{prob}}).$$

Since $\mathcal{A}'_{\text{prob}} \otimes \mathcal{C}'_{\text{prob}}$ is dense in $(\mathcal{A}' \otimes \mathcal{C}') \cong (\mathcal{A} \otimes \mathcal{C})'$, we need that $S' = (T' \otimes 1) = (T \otimes 1)'$, which is equivalent to $S = T \otimes 1$. This leads to the following definition: We say that a linear map $T : \mathcal{B} \to \mathcal{A}$ is *completely positive* if the map

$$T \otimes 1 : \mathcal{B} \otimes \mathcal{L}(\mathcal{K}) \to \mathcal{A} \otimes \mathcal{L}(\mathcal{K})$$

is positive for every inner product space $\mathcal{K}$. Surprisingly, we will see below that positivity does not imply complete positivity. We note that if T is completely positive and $\mathcal{C}$ is some sub-$*$-algebra of $\mathcal{L}(\mathcal{K})$, then $T \otimes 1 : \mathcal{B} \otimes \mathcal{C} \to \mathcal{A} \otimes \mathcal{C}$ is positive. Thus, since every Q-algebra can be embedded in some $\mathcal{L}(\mathcal{K})$, it suffices to consider only factor algebras $\mathcal{L}(\mathcal{K})$.

8.2 A counterexample

We set out to show that positivity does not imply complete positivity and to characterise all completely positive maps. Before we show this, we first give a slightly different formulation of complete positivity, that is often useful. Let $\mathcal{A}$ be a positive *-algebra and let $\mathcal{K}$ be an inner product space. Let $\{e(1), \ldots, e(m)\}$ be an orthonormal basis for $\mathcal{K}$. Then the linear operators

$$|e(i)\rangle\langle e(j)| \qquad (i, j = 1, \ldots, m)$$

form a basis for $\mathcal{L}(\mathcal{K})$, and we can decompose $\mathcal{A} \otimes \mathcal{L}(\mathcal{K})$ as a linear space as

$$\mathcal{A} \otimes \mathcal{L}(\mathcal{K}) \cong \bigoplus_{i,j=1,\ldots,m} \mathcal{A} \otimes |e(i)\rangle\langle e(j)|.$$

Thus, every $A \in \mathcal{A} \otimes \mathcal{L}(\mathcal{K})$ has a unique decomposition

$$A = \sum_{i,j=1}^{m} A_{ij} \otimes |e(i)\rangle\langle e(j)|,$$

with $A_{ij} \in \mathcal{A}$. Now

$$\Big(\sum_{ij} A_{ij} \otimes |e(i)\rangle\langle e(j)|\Big)\Big(\sum_{kl} B_{kl} \otimes |e(k)\rangle\langle e(l)|\Big) = \sum_{il} \Big(\sum_{j} A_{ij}B_{jl}\Big)|e(i)\rangle\langle e(l)|,$$

which shows that

$$(AB)_{ij} = \sum_k A_{ik}B_{kj}.$$

Moreover,

$$\Big(\sum_{ij} A_{ij} \otimes |e(i)\rangle\langle e(j)|\Big)^* = \sum_{ij} A_{ij}^* \otimes |e(j)\rangle\langle e(i)|,$$

which shows that

$$(A^*)_{ij} = (A_{ji})^*.$$

Thus, we see that, with respect to an orthonormal basis for $\mathcal{K}$, there is a natural isomorphism between the positive *-algebra $\mathcal{A} \otimes \mathcal{L}(\mathcal{K})$ and the space of $m \times m$ matrices with entries from $\mathcal{A}$. Now if $T : \mathcal{B} \to \mathcal{A}$ is a linear map, then the linear map $(T \otimes 1) : \mathcal{B} \otimes \mathcal{L}(\mathcal{K}) \to \mathcal{A} \otimes \mathcal{L}(\mathcal{K})$ satisfies

$$(T \otimes 1)\Big(\sum_{ij} B_{ij} \otimes |e(i)\rangle\langle e(j)|\Big) = \Big(\sum_{ij} T(B_{ij}) \otimes |e(i)\rangle\langle e(j)|\Big),$$

which shows that

$$\big((T \otimes 1)(B)\big)_{ij} = T(B_{ij}). \tag{8.4}$$

If $\mathcal{A}$ is a positive *-algebra then we let $M_m(\mathcal{A})$ denote the space of $m \times m$ matrices with entries from $\mathcal{A}$, equipped with the structure of a $*$-algebra by putting $(AB)_{ij} := \sum_k A_{ik}B_{kj}$ and $(A^*)_{ij} := (A_{ji})^*$.

Lemma 8.2.1 (Different formulation of complete positivity) *Let $\mathcal{A}, \mathcal{B}$ be Q-algebras and let $T : \mathcal{B} \to \mathcal{A}$ be a linear map. Then T is completely positive if and only if the map from $M_m(\mathcal{B})$ to $M_m(\mathcal{A})$ given by*

$$\begin{pmatrix} B_{11} & \cdots & B_{1m} \\ \vdots & & \vdots \\ B_{m1} & \cdots & B_{mm} \end{pmatrix} \mapsto \begin{pmatrix} T(B_{11}) & \cdots & T(B_{1m}) \\ \vdots & & \vdots \\ T(B_{m1}) & \cdots & T(B_{mm}) \end{pmatrix}$$

is positive for each m.

Counterexample 8.2.2 (A positive map that is not completely positive) Let $\mathcal{H}$ be an inner product space of dimension 2 and let $\{e(1), e(2)\}$ be an orthonormal basis. Then the linear map $T : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ given, with respect to this basis, by $T(A)_{ij} := A_{ji}$, is positive and satisfies $T(1) = 1$, but T is not completely positive.

Proof If A is positive then

$$\sum_{ij} x_i^* A_{ij} x_j \geq 0 \qquad \forall x \in \mathbb{C}^2.$$

This implies (take $y_i := x_i^*$)

$$\sum_{ij} y_i^* T(A)_{ij} y_j = \sum_{ij} y_i A_{ij} y_j^* \geq 0 \qquad \forall y \in \mathcal{H},$$

which shows that $T(A)$ is positive. However, the map

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \mapsto \begin{pmatrix} T(A_{11}) & T(A_{12}) \\ T(A_{21}) & T(A_{22}) \end{pmatrix}$$

is not positive, since under this map

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The matrix on the left is

$$\left| \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right|,$$

which is clearly a positive operator. On the other hand, the matrix on the right has an eigenvalue -1:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} = - \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}$$

and is therefore not positive. ∎

Remark An operator $T : \mathcal{B} \to \mathcal{A}$ is called n-positive if the map from $M_n(\mathcal{A})$ to $M_n(\mathcal{B})$ in Lemma 8.2.1 is positive. One can show that $(n+1)$-positive $\Rightarrow$ n-positive but n-positive $\not\Rightarrow$ $(n+1)$-positive.

8.3 Characterization of complete positivity

Our aim in this section is to describe the form of a general completely positive map on a given Q-algebra $\mathcal{A}$. For simplicity, we restrict ourselves to the case when $\mathcal{A}$ is a factor algebra.

The next theorem, which describes completely positive maps between factor algebras, is due to Stinespring [Sti55]; see also [Tak79, Thm IV.3.6].

Theorem 8.3.1 (Stinespring) *Let $\mathcal{H}$ and $\mathcal{F}$ be inner product spaces and let $V(1), \ldots, V(n)$ be linear maps in $\mathcal{L}(\mathcal{H}, \mathcal{F})$. Then*

$$T(A) := \sum_{m=1}^{n} V(m) A V(m)^* \qquad (A \in \mathcal{L}(\mathcal{H}))$$

defines a completely positive linear map $T : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{F})$ and conversely, every completely positive linear map from $\mathcal{L}(\mathcal{H})$ to $\mathcal{L}(\mathcal{F})$ is of this form.

Proof This proof is best done with tensor calculus. By Lemma 1.4.3, we have $\mathcal{L}(\mathcal{H}) \cong \mathcal{H} \otimes \mathcal{H}'$ and $\mathcal{L}(\mathcal{F}) \cong \mathcal{F} \otimes \mathcal{F}'$. A linear map T from $\mathcal{L}(\mathcal{H})$ to $\mathcal{L}(\mathcal{F})$ is therefore an element of

$$\mathcal{L}(\mathcal{H} \otimes \mathcal{H}', \mathcal{F} \otimes \mathcal{F}) \cong \mathcal{F} \otimes \mathcal{F}' \otimes (\mathcal{H} \otimes \mathcal{H}')' \cong \mathcal{F} \otimes \mathcal{F}' \otimes \mathcal{H}' \otimes \mathcal{H}.$$

Thus, there is a tensor

$$\tilde{T} \in \mathcal{F} \otimes \mathcal{F}' \otimes \mathcal{H}' \otimes \mathcal{H}$$

such that with respect to bases for $\mathcal{F}, \mathcal{H}$ and the corresponding dual bases for $\mathcal{F}', \mathcal{H}'$, one has

$$(T(A))_{ij} = \sum_{kl} \tilde{T}_{ijkl} A_{kl}.$$

Note that $A \in \mathcal{L}(\mathcal{H}) \cong \mathcal{H} \otimes \mathcal{H}'$. We have contracted the third coordinate of $\tilde{T}$ with the first coordinate of A, which corresponds to contracting $\mathcal{H}'$ with $\mathcal{H}$, and the fourth coordinate of $\tilde{T}$ with the second coordinate of A, which corresponds to contracting $\mathcal{H}$ with $\mathcal{H}'$. In view of what follows, it will be convenient to order the the spaces in the tensor product $\mathcal{F} \otimes \mathcal{F}' \otimes \mathcal{H}' \otimes \mathcal{H}$ in a different way. Let

$$T \in \mathcal{F} \otimes \mathcal{H}' \otimes \mathcal{H} \otimes \mathcal{F}'$$

be the tensor defined by $T_{ijkl} = \tilde{T}_{iljk}$. Then

$$(T(A))_{ij} = \sum_{kl} T_{iklj} A_{kl}. \tag{8.5}$$

Now T is of the form $T(A) = \sum_m V(m)AV(m)^*$ if and only if

$$\sum_{kl} T_{iklj}A_{kl} = (T(A))_{ij} = \big(\sum_m V(m)AV(m)^*\big)_{ij} = \sum_m \sum_{kl} V_{ik}(m)A_{kl}\overline{V}_{jl}(m),$$

which is equivalent to

$$T_{iklj} = \sum_m V_{ik}(m)\overline{V}_{jl}(m). \tag{8.6}$$

We must now formulate complete positivity in the language of tensor calculus. If $\mathcal{K}$ is another inner product space then $\mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{K}) \cong \mathcal{L}(\mathcal{H} \otimes \mathcal{K})$ and

$$\mathcal{L}(\mathcal{H} \otimes \mathcal{K}) \cong (\mathcal{H} \otimes \mathcal{K}) \otimes (\mathcal{H} \otimes \mathcal{K})' \cong \mathcal{H} \otimes \mathcal{K} \otimes \mathcal{K}' \otimes \mathcal{H}'.$$

If we order the spaces $\mathcal{H}, \mathcal{K}, \mathcal{K}', \mathcal{H}'$ in the tensor product in this way, then the map $(T \otimes 1) : \mathcal{L}(\mathcal{H} \otimes \mathcal{K}) \to \mathcal{L}(\mathcal{F} \otimes \mathcal{K})$ takes the form

$$\big((T \otimes 1)A\big)_{ijkl} = \sum_{mn} T_{imnl}A_{mjkn}, \tag{8.7}$$

i.e., $T \otimes 1$ acts only on the coordinates corresponding to $\mathcal{H} \otimes \mathcal{H}'$ (compare formula (8.4)). By definition, T is completely positive if for each $\mathcal{K}$, $T \otimes 1$ maps positive operators into positive operators. We need the following simple fact, which we leave as an exercise.

Exercise 8.3.2 Show that an operator $A \in \mathcal{L}(\mathcal{H}) \cong \mathcal{H} \otimes \mathcal{H}'$ is positive if and only if there exist $x(n) \in \mathcal{H}$ such that $A = \sum_n |x(n)\rangle\langle x(n)|$. Show that in coordinates this says that $A_{ij} = \sum_n x_i(n)x_j^*(n)$. Show that the $x(n)$ can be chosen orthogonal.

For the space $\mathcal{L}(\mathcal{H} \otimes \mathcal{K}) \cong \mathcal{H} \otimes \mathcal{K} \otimes \mathcal{K}' \otimes \mathcal{H}'$ this means that A is positive if and only if A is of the form

$$A_{ijkl} = \sum_n B_{ij}(n)\overline{B}_{lk}(n) \tag{8.8}$$

for some $B(n) \in \mathcal{H} \otimes \mathcal{K}$.

So far we have only translated all definitions into the language of tensor calculus. Now, we must show that the map $T \otimes 1$ defined in (8.7) maps A of the form (8.8) into operators of the form (8.8) again if and only if T is of the form (8.6).

If T is of the form (8.6) and A is of the form (8.8) then

$$\begin{aligned}\big((T \otimes 1)A\big)_{ijkl} &= \sum_n \sum_m \sum_{pq} V_{ip}(m)B_{pj}(n)\overline{V}_{lq}(m)\overline{B}_{qk}(n) \\ &= \sum_{nm} (V(m)B(n))_{ij}\overline{(V(m)B(n))}_{lk},\end{aligned}$$

which is again something of the form (8.8).

To prove that conversely every completely positive T is of the form (8.6) we use a trick. For the space $\mathcal{K}$ we may in particular choose $\mathcal{K} = \mathcal{H}'$. Let 1 denote the identity in $\mathcal{L}(\mathcal{H}) \cong \mathcal{H} \otimes \mathcal{H}' = \mathcal{H} \otimes \mathcal{K}$. Then

$$A_{ijkl} := 1_{ij}\overline{1}_{kl}$$

defines an operator of the form (8.8). Here $1_{ij} = \overline{1}_{ij} = \delta_{ij}$. Since T is completely positive, $T \otimes 1$ maps A into an operator of the form (8.8), and therefore

$$T_{ijkl} = \sum_{mn} T_{imnl}1_{mj}\overline{1}_{kn} = \big((T \otimes 1)A\big)_{ijkl} = \sum_{p} V_{ij}(p)\overline{V}_{kl}(p)$$

for some $V(p) \in \mathcal{F} \otimes \mathcal{K} = \mathcal{F} \otimes \mathcal{H}'$. This shows that T is of the form (8.6). ∎

Exercise 8.3.3 Show that it is possible to choose the $V(n)$ in Theorem 8.3.1 orthogonal with respect to a suitable inner product on $\mathcal{L}(\mathcal{H}, \mathcal{F})$.

8.4 Operations

We now return to our original aim of studying completely positive maps. Let $\mathcal{A}, \mathcal{B}$ be Q-algebras. By definition, an *operation* from $\mathcal{A}$ to $\mathcal{B}$ is a map of the form $T' : \mathcal{A}' \to \mathcal{B}'$, where $T : \mathcal{B} \to \mathcal{A}$ is a completely positive linear map satisfying $T(1) = 1$. Operations correspond to things that one can do with a physical system, changing its state, and possibly even the algebra needed to describe the system.

Proposition 8.4.1 (Operations on factor algebras) *Let $\mathcal{A}$ be a Q-algebra and let $V(1), \ldots, V(n) \in \mathcal{A}$ satisfy $\sum_{m=1}^{n} V(m)V(m)^* = 1$. Let $T : \mathcal{A} \to \mathcal{A}$ be given by*

$$TA := \sum_{m=1}^{n} V(m)AV(m)^*$$

and let $T' : \mathcal{A}' \to \mathcal{A}'$ be given by

$$(T'\rho)(A) := \rho\Big(\sum_{m=1}^{n} V(m)AV(m)^*\Big) \quad (A \in \mathcal{L}(\mathcal{H})). \tag{8.9}$$

Then T' is an operation from $\mathcal{A}$ to $\mathcal{A}$. If $\mathcal{A}$ is a factor algebra, then every operation from $\mathcal{A}$ to $\mathcal{A}$ is of this form.

Proof It is not hard to check that T is completely positive and $T(1) = 1$. The fact that every operation has this form if $\mathcal{A}$ is a factor algebra follows from Stinespring's theorem. ∎

Remark 1 Ideal measurements are operations, where in this case $V(1), \ldots, V(n)$ are projections which form a partition of the identity.

Remark 2 The composition of two ideal measurements, performed one after the other, is an operation. However, unless the measurements commute, this composition is in general not itself an ideal measurement.

Remark 3 If $\mathcal{A}$ is not a factor algebra, then not all operations are of the form (8.9). Indeed, if $\mathcal{A}$ is abelian and $\dim(\mathcal{A}) \geq 2$, then the only operation of the form (8.9) is the identity map $\rho \mapsto \rho$, while by Proposition 8.4.2 below there are many nontrivial operations on $\mathcal{A}$.

Proposition 8.4.2 (Operations on abelian algebras) *Consider the abelian Q-algebras $\mathbb{C}^n$ and $\mathbb{C}^m$. Let $(\pi_{ij})_{i=1,\ldots,n,\ j=1,\ldots,m}$ be nonnegative numbers such that $\sum_j \pi_{ij} = 1$ for each i. Define $T : \mathbb{C}^m \to \mathbb{C}^n$ by*

$$T(b_1, \ldots, b_m) := \Big(\sum_j \pi_{1j} b_j, \ldots, \sum_j \pi_{nj} b_j \Big).$$

Then $T' : (\mathbb{C}^n)' \to (\mathbb{C}^m)'$ is an operation from $\mathbb{C}^n$ to $\mathbb{C}^m$ such that

$$T'\delta_i = \sum_j \pi_{ij} \delta_j \qquad (i = 1, \ldots, n), \tag{8.10}$$

and every operation from $\mathbb{C}^n$ to $\mathbb{C}^m$ is of this form.

Proof It is easy to see that every operation from $\mathbb{C}^n$ to $\mathbb{C}^m$ must be of the form (8.10). Conversely, it is straightforward to check that (8.10) defines an operation. Note that (8.10) says that if the state before we perform our operation is the delta measure in i, then after we perform our operation we are in state j with probability π_{ij}. ∎

The next lemma is almost a direct consequence of the way complete positivity has been defined, so we skip the proof.

Lemma 8.4.3 (Operations on logically independent algebras) *Let $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ be Q-algebras and let $T' : \mathcal{A}' \to \mathcal{B}'$ be an operation. Then there exists a unique operation $S' : (\mathcal{A} \otimes \mathcal{C})' \to (\mathcal{B} \otimes \mathcal{C})'$ such that*

$$S'(\rho_1 \otimes \rho_2) = (T'\rho_1) \otimes \rho_2 \qquad (\rho_1 \in \mathcal{A}'_{\text{prob}},\ \rho_2 \in \mathcal{C}'_{\text{prob}}).$$

This operation is given by $S' = (T \otimes 1)'$. *If* ρ *is any probability on* $\mathcal{A} \otimes \mathcal{C}$, *then*

$$\big((T \otimes 1)'\rho\big)(1 \otimes C) = \rho(1 \otimes C) \qquad (C \in \mathcal{C}).$$

Note that the second formula says that if $\mathcal{A}$ and $\mathcal{C}$ are logically independent, then performing an operation T on $\mathcal{A}$ does not change our knowledge about $\mathcal{C}$.

We next take a look at deterministic operations. Let $\mathcal{A}, \mathcal{B}$ be Q-algebras and let $T' : \mathcal{A}' \to \mathcal{B}'$ be an operation. We will say T' is *deterministic* if for each inner product space $\mathcal{K}$, the operation

$$(T' \otimes 1) : (\mathcal{A} \otimes \mathcal{L}(\mathcal{K}))' \to (\mathcal{B} \otimes \mathcal{L}(\mathcal{K}))'$$

maps pure states into pure states. Since pure states are probabilities describing maximal information about a physical system, deterministic operations are operations without loss of information.

Proposition 8.4.4 (Deterministic operations on factor algebras) *Let* $\mathcal{A} \cong \mathcal{L}(\mathcal{H}))$ *and* $\mathcal{B} \cong \mathcal{L}(\mathcal{F})$ *be factor algebras and assume that* $\mathcal{H}$ *and* $\mathcal{F}$ *have the same dimension. Then, for each unitary operator* $U : \mathcal{H} \to \mathcal{F}$, *the formula*

$$T(B) := UBU^{-1} \qquad (B \in \mathcal{B})$$

defines a completely positive map $T : \mathcal{B} \to \mathcal{A}$ *such that the dual map* $T' : \mathcal{A}' \to \mathcal{B}'$ *is a deterministic operation. Conversely, every deterministic operation arises in this way.*

Proof There is a one-to-one correspondence between linear forms $l \in \mathcal{A}'$ and their densities $L \in \mathcal{A}$ with respect to the trace on $\mathcal{H}$, given by

$$l(A) = \mathrm{tr}(L^*A) \qquad (A \in \mathcal{A}), \tag{8.11}$$

so the linear map $T' : \mathcal{A}' \to \mathcal{B}'$ gives rise to a linear map (also denoted by T') from $\mathcal{A}$ to $\mathcal{B}$, such that

$$(T'l)(A) = \mathrm{tr}((T'L)^*A) \qquad (A \in \mathcal{A}).$$

The fact that $(T' \otimes 1) : (\mathcal{A} \otimes \mathcal{L}(\mathcal{K}))' \to (\mathcal{B} \otimes \mathcal{L}(\mathcal{K}))'$ maps pure states into pure states now means that the corresponding map from $(\mathcal{A} \otimes \mathcal{L}(\mathcal{K}))$ to $(\mathcal{B} \otimes \mathcal{L}(\mathcal{K}))$ maps densities of the form $|\psi\rangle\langle\psi|$ with $\psi \in \mathcal{H} \otimes \mathcal{K}$ into densities of the form $|\phi\rangle\langle\phi|$ with $\phi \in \mathcal{F} \otimes \mathcal{K}$. (Recall Lemma 4.1.7.) Taking $\mathcal{K} = \mathcal{H}'$ and mimicking the proof of Stinespring's theorem, we see that must be of the form

$$T'(L) = U^*LU$$

for some $U \in \mathcal{L}(\mathcal{H}, \mathcal{F})$. Thus, the linear form l in (8.11) is mapped under T' to

$$(T'l)(B) = \mathrm{tr}\big((U^*LU)^*B\big) = l(UBU^*) \qquad (B \in \mathcal{B}),$$

which shows that

$$T(B) = UBU^* \qquad (B \in \mathcal{B}).$$

Since $T(1) = 1$ we have $U = U^{-1}$ so U is unitary. Conversely, it is easy to check that any operation of this form maps pure states into pure states. ∎

Chapter 9

Quantum peculiarities

9.1 Cloning, coding, and teleportation

In this section we will meet two surprising quantum impossibilities, and one possibility. The first theorem that we will prove says that it is not possibly to *copy* a quantum system. This result goes under the fancy name *'no cloning'*.

Theorem 9.1.1 (No cloning) *Let $\mathcal{A}$ be a Q-algebra and let $T' : \mathcal{A}' \to (\mathcal{A} \otimes \mathcal{A})'$ be an operation such that*

$$(T'\rho)(A \otimes 1) = (T'\rho)(1 \otimes A) = \rho(A) \qquad (A \in \mathcal{A},\ \rho \in \mathcal{A}'_{\text{prob}}). \tag{9.1}$$

Then $\mathcal{A}$ is abelian.

Note that (9.1) says that the operation T' takes a single system in the state ρ and produces two logically independent (but not necessarily independent) systems, both in the state ρ. The claim is that if the algebra is not abelian, then there is no operation that does this for any state ρ. On the other hand, in classical probability, copying is always possible:

Exercise 9.1.2 Show that if $\mathcal{A}$ is abelian, there exists an operation T' such that (9.1) holds. Show that T' can be chosen in such a way that

$$(T'\rho)(P \otimes 1) = (T'\rho)(P \otimes P) = (T'\rho)(1 \otimes P)$$

for every projection $P \in \mathcal{A}$ and $\rho \in \mathcal{A}'_{\text{prob}}$. If T' is chosen in this way, then are the two subsystems in general independent under $T'\rho$?

We postpone the proof of Theorem 9.1.1 and first state another, similar theorem. This theorem says that it is not possible to extract all information from a non-abelian algebra, and 'write it down' in an abelian algebra, such that with this information the original nonablian system can later be reconstructed. This result goes under the name *'no coding'*.

Theorem 9.1.3 (No coding) *Let $\mathcal{A}, \mathcal{B}$ be a Q-algebras and let $T' : \mathcal{A}' \to \mathcal{B}'$ and $S' : \mathcal{B}' \to \mathcal{A}'$ be operations such that*

$$S'T'\rho = \rho \qquad (\rho \in \mathcal{A}'_{\text{prob}}). \tag{9.2}$$

Then if $\mathcal{B}$ is abelian, so is $\mathcal{A}$.

Remark As in the previous theorem, it is important that S' and T' work for *any* $\rho \in \mathcal{A}'_{\text{prob}}$.

Again we postpone the proof, and first state another theorem. This theorem says that it is possible to extract information from a quantum system, 'write it down' in an abelian algebra, and then send it to someone else, so that he can reconstruct the original system. This seems to contradict our previous theorem, but the trick is that the two people who want to send information to each other have prepared two entangled particles, and each of them keeps one particle with them. By making use of these entangled particles, they can send the quantum system. This result goes under the science fiction-like name *'teleportation'* [B–W93]. For simplicity, we will only teleport states on an algebra of the form $\mathcal{L}(\mathcal{H})$ where $\mathcal{H}$ is two-dimensional. As is standard in this sort of communication problems, the sender will be called *Alice* and the recipient will be called *Bob*. (Which explains why I wrote that '*he* can reconstruct the original system'.)

Theorem 9.1.4 (Quantum teleportation) *Let $\mathcal{A}$ and $\mathcal{C}$ be Q-algebras of the form $\mathcal{L}(\mathcal{H})$, where $\mathcal{H}$ is two-dimensional, and let $\mathcal{B}$ be a four-dimensional abelian Q-algebra. Then there exists a state $\eta \in (\mathcal{C} \otimes \mathcal{C})'_{\text{prob}}$ and operations $T' : (\mathcal{A} \otimes \mathcal{C})' \to \mathcal{B}'$ and $S' : (\mathcal{B} \otimes \mathcal{C})' \to \mathcal{A}'$, such that*

$$S' \circ (T' \otimes 1)\, \rho \otimes \eta = \rho \qquad (\rho \in \mathcal{A}'_{\text{prob}}). \tag{9.3}$$

Remark 1 In (9.3), $(T' \otimes 1)$ is map from $(\mathcal{A} \otimes \mathcal{C} \otimes \mathcal{C})'_{\text{prob}}$ to $(\mathcal{B} \otimes \mathcal{C})'_{\text{prob}}$, so that the composition of operations $S' \circ (T' \otimes 1)$ is a map from $(\mathcal{A} \otimes \mathcal{C} \otimes \mathcal{C})'_{\text{prob}}$ to $\mathcal{A}'_{\text{prob}}$. The abelian algebra $\mathcal{B}$ contains all information that is sent from Alice to Bob. Therefore, the operation $T' \otimes 1$ acts only on objects that are under Alice's control and S' acts only on objects that are under the control of Bob.

Remark 2 It follows from Lemma 8.4.3 that entangled states cannot be used to send information from Alice to Bob, in spite of their seemingly nonlocal behavior in relation with the Bell inequality. Nevertheless, quantum teleportation shows that entangled states can be used to upgrade 'classical information' to 'quantum information'.

We now set out to prove Theorems 9.1.1–9.1.4. We start with a preparatory result. Recall the partial order for hermitian operators defined on page 14.

Theorem 9.1.5 (Cauchy-Schwarz for operations) *Let $\mathcal{A}, \mathcal{B}$ be Q-algebras and let $T' : \mathcal{A}' \to \mathcal{B}'$ be an operation. Then*

$$T(B^*B) \geq T(B)^*T(B) \qquad (B \in \mathcal{B}).$$

If equality holds for some $B_1 \in \mathcal{B}$, then

$$T(B_1^*B_2) = T(B_1)^*T(B_2) \quad \textit{and} \quad T(B_2^*B_1) = T(B_2)^*T(B_1).$$

Remark 1 The inequality is called 'Cauchy-Schwarz' since we can view $(B_1, B_2) \mapsto T(B_1^*B_2)$ as a sort of $\mathcal{A}$-valued (!) 'inner product' on $\mathcal{B}$. Then we can rewrite the inequality above as $T(B^*B)T(1^*1) \geq T(1^*B)^*T(1^*B)$ which looks like $\langle\phi|\phi\rangle\langle\psi|\psi\rangle \geq |\langle\psi|\phi\rangle|^2$.

Remark 2 Note that if a Cauchy-Schwarz *equality* holds for *all* $B \in \mathcal{B}$, then T is a *-algebra homomorphism.

Proof The operator

$$\begin{pmatrix} B^*B & -B^* \\ -B & 1 \end{pmatrix} = \begin{pmatrix} B & -1 \\ 0 & 0 \end{pmatrix}^* \begin{pmatrix} B & -1 \\ 0 & 0 \end{pmatrix}$$

in $M_2(\mathcal{A})$ is positive. Therefore, since T is completely positive, also

$$\begin{pmatrix} T(B^*B) & -T(B)^* \\ -T(B) & 1 \end{pmatrix}$$

is positive. If $\mathcal{H}$ is a representation for $\mathcal{A}$ then $\mathcal{H} \oplus \mathcal{H}$ is a representation for $M_2(\mathcal{A})$ and therefore, for each $\psi \in \mathcal{H}$,

$$\left\langle \begin{pmatrix} \psi \\ T(B)\psi \end{pmatrix} \middle| \begin{pmatrix} T(B^*B) & -T(B)^* \\ -T(B) & 1 \end{pmatrix} \begin{pmatrix} \psi \\ T(B)\psi \end{pmatrix} \right\rangle \geq 0.$$

This means that

$$\left\langle \begin{pmatrix} \psi \\ T(B)\psi \end{pmatrix} \middle| \begin{pmatrix} T(B^*B)\psi - T(B)^*T(B)\psi \\ 0 \end{pmatrix} \right\rangle \geq 0,$$

and therefore

$$\langle\psi|T(B^*B) - T(B)^*T(B)|\psi\rangle \geq 0 \qquad (\psi \in \mathcal{H}),$$

so $T(B^*B) - T(B)^*T(B)$ is a positive operator, i.e., $T(B^*B) \geq T(B)^*T(B)$.

Now assume that $T(B_1^*B_1) = T(B_1)^*T(B_1)$ for some $B_1 \in \mathcal{B}$. Then, for all $B_2 \in \mathcal{B}$ and $\lambda_1, \lambda_2 \in \mathbb{C}$,

$$\begin{aligned}
&|\lambda_1|^2T(B_1^*B_1) + \overline{\lambda}_1\lambda_2T(B_1^*B_2) + \overline{\lambda}_2\lambda_1T(B_2^*B_1) + |\lambda_2|^2T(B_2^*B_2) \\
&\quad = T\big((\lambda_1B_1 + \lambda_2B_2)^*(\lambda_1B_1 + \lambda_2B_2)\big)\big) \\
&\quad \geq |\lambda_1|^2T(B_1)^*T(B_1) + \overline{\lambda}_1\lambda_2T(B_1)^*T(B_2) \\
&\qquad + \overline{\lambda}_2\lambda_1T(B_2)^*T(B_1) + |\lambda_2|^2T(B_2)^*T(B_2).
\end{aligned}$$

Using our assumption that $T(B_1^*B_1) = T(B_1)^*T(B_1)$ this implies that

$$\begin{aligned}
&\overline{\lambda}_1\lambda_2T(B_1^*B_2) + \overline{\lambda}_2\lambda_1T(B_2^*B_1) + |\lambda_2|^2T(B_2^*B_2) \\
&\quad \geq \overline{\lambda}_1\lambda_2T(B_1)^*T(B_2) + \overline{\lambda}_2\lambda_1T(B_2)^*T(B_1) + |\lambda_2|^2T(B_2)^*T(B_2).
\end{aligned}$$

In particular, setting $\lambda_1 = 1$ and $\lambda_2 = \lambda$ where λ is real, we get

$$\begin{aligned}
&\lambda\big(T(B_1^*B_2) + T(B_2^*B_1)\big) + \lambda^2T(B_2^*B_2) \\
&\quad \geq \lambda\big(T(B_1)^*T(B_2) + T(B_2)^*T(B_1)\big) + \lambda^2T(B_2)^*T(B_2).
\end{aligned}$$

This can only be true for all $\lambda \in \mathbb{R}$ (in particular, for λ very small), if

$$T(B_1^*B_2) + T(B_2^*B_1) \geq T(B_1)^*T(B_2) + T(B_2)^*T(B_1).$$

In the same way, setting $\lambda_1 = -i$ and $\lambda_2 = i\lambda$, where λ is real, we get

$$\begin{aligned}
&-\lambda\big(T(B_1^*B_2) + T(B_2^*B_1)\big) + \lambda^2T(B_2^*B_2) \\
&\quad \geq -\lambda\big(T(B_1)^*T(B_2) + T(B_2)^*T(B_1)\big) + \lambda^2T(B_2)^*T(B_2)
\end{aligned}$$

which together with our previous inequality yields

$$T(B_1^*B_2) + T(B_2^*B_1) = T(B_1)^*T(B_2) + T(B_2)^*T(B_1).$$

In the same way, setting $\lambda_1 = 1$ and $\lambda_2 = \pm i\lambda$ gives

$$T(B_1^*B_2) - T(B_2^*B_1) = T(B_1)^*T(B_2) - T(B_2)^*T(B_1).$$

Therefore

$$T(B_1^*B_2) = T(B_1)^*T(B_2) \quad \text{and} \quad T(B_2^*B_1) = T(B_2)^*T(B_1).$$

Note that these two equalities are actually equivalent, since one can be obtained from the other by taking adjoints. ∎

Proof of Theorem 9.1.1 If T' satisfies (9.1) then $T : \mathcal{A} \otimes \mathcal{A} \to \mathcal{A}$ satisfies

$$T(A \otimes 1) = T(1 \otimes A) = A \qquad (A \in \mathcal{A}).$$

It follows that for any $A \in \mathcal{A}$

$$T((1 \otimes A)^*(1 \otimes A)) = T(1 \otimes A^*A) = A^*A = T(1 \otimes A)^*T(1 \otimes A),$$

i.e., a Cauchy-Schwarz equality holds for operators of the form $1 \otimes A$. Therefore, Theorem 9.1.5 tells us that for any $A, B \in \mathcal{A}$

$$\begin{aligned} AB &= T(A \otimes 1)T(1 \otimes B) = T\big((A \otimes 1)(1 \otimes B)\big) \\ &= T\big((1 \otimes B)(A \otimes 1)\big) = T(1 \otimes B)T(A \otimes 1) = BA, \end{aligned}$$

which shows that $\mathcal{A}$ is abelian. ∎

Proof of Theorem 9.1.3 Since cloning is possible for states on abelian algebras, one can show that if states on nonabelian algebras could be completely 'written down' in abelian algebras, then they could also be cloned. In this way, it is possible to derive Theorem 9.1.3 from Theorem 9.1.1. However, we will give a shorter, independent proof.

If T' and S' satisfy (9.2) then

$$T \circ S(A) = A \qquad (A \in \mathcal{A}).$$

Cauchy-Schwarz gives

$$A^*A = T \circ S(A^*A) \geq T\big(S(A)^*S(A)\big) \geq T(S(A)\big)^*T(S(A)) = A^*A \qquad (A \in \mathcal{A}),$$

so we must have two times equality here. In particular, by Theorem 9.1.5, the second equality tells us that $T(S(A)S(B)) = T(S(A))T(S(B))$ for all $A, B \in \mathcal{A}$. Therefore, since $\mathcal{B}$ is abelian,

$$AB = T(S(A))T(S(B)) = T\big(S(A)S(B)\big) = T\big(S(B)S(A)\big) = BA \qquad (A, B \in \mathcal{A}),$$

which shows that $\mathcal{A}$ is abelian. ∎

Proof of Theorem 9.1.4 We take $\mathcal{A} = \mathcal{C} = \mathcal{L}(\mathcal{H})$, where $\mathcal{H}$ is a two-dimensional inner product space with orthonormal basis $\{e_1, e_2\}$. For η, we take the entangled pure state described by the state vector $\chi \in \mathcal{H} \otimes \mathcal{H}$ given by

$$\chi := \tfrac{1}{\sqrt{2}}\, e_1 \otimes e_1 + \tfrac{1}{\sqrt{2}}\, e_2 \otimes e_2.$$

It suffices to prove the theorem for the case that ρ is a pure state. (The general case then follows by linearity.) In that case, the initial state $\rho \otimes \eta \in (\mathcal{A} \otimes \mathcal{C} \otimes \mathcal{C})'_{\text{prob}}$ is described by a state vector in $\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}$ of the form

$$y \otimes \chi = (y_1 e_1 + y_2 e_2) \otimes \chi,$$

where y_1, y_2 are complex numbers (unknown to Alice and Bob) such that $|y_1|^2 + |y_2|^2 = 1$. Now Alice performs an ideal measurement on the joint system consisting of the state she wants to send and her particle from the entangled pair. This ideal measurement is described by the partition of the identity $\{P_1, \ldots, P_4\}$, where

$$P_i = |\psi_i\rangle\langle\psi_i| \qquad (i = 1, \ldots, 4),$$

and

$$\begin{aligned}
\psi_1 &:= \tfrac{1}{\sqrt{2}} e_1 \otimes e_1 + \tfrac{1}{\sqrt{2}} e_2 \otimes e_2, \\
\psi_2 &:= \tfrac{1}{\sqrt{2}} e_1 \otimes e_1 - \tfrac{1}{\sqrt{2}} e_2 \otimes e_2, \\
\psi_3 &:= \tfrac{1}{\sqrt{2}} e_1 \otimes e_2 + \tfrac{1}{\sqrt{2}} e_2 \otimes e_1, \\
\psi_4 &:= \tfrac{1}{\sqrt{2}} e_1 \otimes e_2 - \tfrac{1}{\sqrt{2}} e_2 \otimes e_1.
\end{aligned}$$

We rewrite the state $y \otimes \chi$ as

$$\begin{aligned}
y \otimes \chi &= (y_1 e_1 + y_2 e_2) \otimes \big(\tfrac{1}{\sqrt{2}}\, e_1 \otimes e_1 + \tfrac{1}{\sqrt{2}}\, e_2 \otimes e_2\big) \\
&= \tfrac{1}{\sqrt{2}} y_1 e_1 \otimes e_1 \otimes e_1 + \tfrac{1}{\sqrt{2}} y_1 e_1 \otimes e_2 \otimes e_2 \\
&\quad + \tfrac{1}{\sqrt{2}} y_2 e_2 \otimes e_1 \otimes e_1 + \tfrac{1}{\sqrt{2}} y_2 e_2 \otimes e_2 \otimes e_2 \\
&= y_1 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_1 + \tfrac{1}{\sqrt{2}} e_2 \otimes e_2\big) \otimes e_1 + y_1 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_1 - \tfrac{1}{\sqrt{2}} e_2 \otimes e_2\big) \otimes e_1 \\
&\quad + y_1 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_2 + \tfrac{1}{\sqrt{2}} e_2 \otimes e_1\big) \otimes e_2 + y_1 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_2 - \tfrac{1}{\sqrt{2}} e_2 \otimes e_1\big) \otimes e_2 \\
&\quad + y_2 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_2 + \tfrac{1}{\sqrt{2}} e_2 \otimes e_1\big) \otimes e_1 - y_2 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_2 - \tfrac{1}{\sqrt{2}} e_2 \otimes e_1\big) \otimes e_1 \\
&\quad + y_2 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_1 + \tfrac{1}{\sqrt{2}} e_2 \otimes e_2\big) \otimes e_2 - y_2 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_1 - \tfrac{1}{\sqrt{2}} e_2 \otimes e_2\big) \otimes e_2 \\
&= \tfrac{1}{2}\psi_1 \otimes \big(y_1 e_1 + y_2 e_2\big) + \tfrac{1}{2}\psi_2 \otimes \big(y_1 e_1 - y_2 e_2\big) \\
&\quad + \tfrac{1}{2}\psi_3 \otimes \big(y_1 e_2 + y_2 e_1\big) + \tfrac{1}{2}\psi_4 \otimes \big(y_1 e_2 - y_2 e_1\big).
\end{aligned}$$

The measurement of Alice produces with equal probabilities any of the outcomes $1, \ldots, 4$. If the outcome is 1, then Bob's particle of the entangled pair is in the state $y_1 e_1 + y_2 e_2$. If the outcome of Alice's measurement is 2, then Bob's particle is in the state $y_1 e_1 - y_2 e_2$, and so on.

If Bob learns about the outcome of Alice's experiment, then he perform a deterministic operation on his particle from the entangled pair, so that after this operation, this particle is in the state y that Alice wanted to send. (Note that y is still unknown to Alice and Bob!) If the outcome of Alice's experiment is i, then Bob performs the deterministic operation described by the unitary map U_i, where the matrices of $U_1, \ldots, U_4 \in \mathcal{L}(\mathcal{H})$ with respect to the basis $\{e_1, e_2\}$ are given by

$$U_1 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad U_2 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad U_3 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad U_4 := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

For example, if $i = 4$, then Bob's particle is in the state $(y_1 e_2 - y_2 e_1)$, which under the unitary transformation U_4 becomes

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} -y_2 \\ y_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

It is easy to see that in any of the four cases, Bob ends up with a particle in the pure state $y_1 e_1 + y_2 e_2$, which is the state Alice wanted to send.

More formally, the operations of Alice and Bob can be described as follows.

Alice's operation is a map $T' : \mathcal{L}(\mathcal{H} \otimes \mathcal{H})' \to \mathcal{B}'$. Here $\mathcal{B}$ is of the form $\mathcal{B} \cong \mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C} = \{b = (b_1, \ldots, b_4) : b_i \in \mathbb{C}\}$ and $\mathcal{A} \otimes \mathcal{C} \cong \mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{H}) \cong \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$. Let π_i denote the pure state on $\mathcal{B}$ defined as $\pi_i(b) := b_i$. Then $T' : \mathcal{L}(\mathcal{H} \otimes \mathcal{H})' \to \mathcal{B}'$ maps the pure state described by the state vector ψ_i to the pure state π_i. It is not hard to see that this is achieved by the operator $T : \mathcal{B} \to \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$ given by

$$T(b_1, \ldots, b_4) = \sum_{i=1}^{4} V_i b_i V_i^*,$$

where $V_i : \mathbb{C} \to (\mathcal{H} \otimes \mathcal{H})$ is defined as

$$V_i := |\psi_i\rangle \qquad (i = 1, \ldots, 4).$$

It follows from Stinespring's theorem (Theorem 8.3.1) that T is completely positive. Moreover, $T(1) = \sum_i |\psi_i\rangle\langle\psi_i| = 1$.

Bob's operation is a map $S' : (\mathcal{B} \otimes \mathcal{C})' \to \mathcal{C}'$. Here $\mathcal{C} \cong \mathcal{L}(\mathcal{H})$ and $\mathcal{B} \otimes \mathcal{C} \cong \big(\mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C}\big) \otimes \mathcal{L}(\mathcal{H}) \cong \mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{H})$. It is not hard to see that Bob's operation is achieved by the operator $S : \mathcal{C} \to (\mathcal{B} \otimes \mathcal{C})$ given by

$$S(A) = \big(U_1^* A U_1, \ldots, U_4^* A U_4\big) \qquad (A \in \mathcal{L}(\mathcal{H})),$$

where $U_1, \ldots, U_n$ are the unitary operators defined above. Comparing this with Proposition 8.4.4 we see that S' is a deterministic operation.

Define $(1 \otimes \eta) : \mathcal{A} \otimes \mathcal{C} \otimes \mathcal{C} \to \mathcal{A}$ by

$$\eta(A \otimes B) := \eta(B)A \qquad (A \in \mathcal{A},\ B \in \mathcal{C} \otimes \mathcal{C}).$$

Then the operations T' and S' and the state η satisfy (9.3) for all $\rho \in \mathcal{A}'_{\text{prob}}$ if and only if

$$\rho \circ (1 \otimes \eta) \circ (T \otimes 1) \circ S(A) = \rho(A) \qquad (A \in \mathcal{A},\ \rho \in \mathcal{A}'_{\text{prob}}),$$

which is equivalent to

$$(1 \otimes \eta) \circ (T \otimes 1) \circ S(A) = \rho(A) \qquad (A \in \mathcal{A}).$$

This formula can be checked by straightforward calculation. In fact, we have already seen all the essential ingredients of this calculation in our informal discussion of Alice's and Bob's operations, so we do not go into details. ∎

9.2 Quantum cryptography

In the previous section, we saw that quantum probability leads to some surprising impossibilities: 'no cloning' and 'no coding'. In view of these impossibilities, the possibility of 'quantum teleportation' is surprising, but from the classical point of view, where copying is possible, this is nothing new. In this section, we will see that the peculiarities of quantum probability also open some new possibilities that are not present in classical probability.

Suppose that Alice wants to send a (classical) message to Bob, say, a sequence $(x_1, \ldots, x_n)$ of zeros and ones, while making sure that a third party, called *Eve*, is not eavesdropping. Alice can do this if she has another, random sequence $(y_1, \ldots, y_n)$ of zeros and ones, that is known to her and Bob, but to no one else. With such a sequence, she simply sends Bob the sequence $(z_1, \ldots, z_n)$ given by

$$z_i = x_i + y_i \mod 1.$$

To anyone who does not know the code $(y_1, \ldots, y_n)$, the message $(z_1, \ldots, z_n)$ is just a random sequence of zeros and ones, but Bob, who knows the code, calculates

$$x_i = z_i + y_i \mod 1$$

to get Alice's message $(x_1, \ldots, x_n)$.

Now Alice and Bob may have prepared their code before, when they were together, but since each code may be used only once[1] they may run out of code after a while. This leads to the following question: can Alice send Bob a secret code of independent zeros and ones, while being sure that Eve is not listening somewhere on the telephone line?

In classical probability, this is not possible, since Eve can perform a nonperturbing measurement on the signal passing through the telephone line. But in quantum probability, the situation is quite different. In [J–Z00], a team around professor Zeilinger from Vienna shows how Alice can send a code to Bob while making sure that Eve is not eavesdropping. What is more, they report on an experiment that shows that this form of communication is possible in practice. On April 27, 2004, the *Süddeutsche Zeitung* reported how in Vienna, 3000 euro were transferred from the town hall to the bank account of the university, using this form of quantum cryptography.

This is how it works. Alice prepares pairs entangled photons, of which she keeps one for herself, and sends the other one to Bob through a glass fiber cable. On the photons that Alice keeps for herself, she performs at random, with equal probabilities, either the ideal measurement $\{P_0, P_{\pi/2}\}$ or the ideal measurement $\{P_\gamma, P_{\gamma+\pi/2}\}$, where P_γ is the projection defined on page 98 and γ is an angle that we will choose later. Likewise, Bob performs on his photons with equal probabilities either the measurement $\{P_0, P_{\pi/2}\}$ or $\{P_{-\gamma}, P_{-\gamma+\pi/2}\}$. After sending as many photons as she needs, Alice tells Bob over a regular telephone line which measurements she used for her photons. Bob then tells Alice which measurements he used, and he tells the outcomes of all his measurements in those cases where they did not both perform the measurement $\{P_0, P_{\pi/2}\}$. As we will see in a moment, Alice can see from this information whether Eve was eavesdropping or not. If she sees that Eve was not tapping the phone, Alice sends Bob the message that she wanted to send, using as a code the outcome of those measurements where she and Bob both performed the measurement $\{P_0, P_{\pi/2}\}$.

[1] If the same code is use repeatedly, then the coded messages are no longer sequences of independent random variables, and this dependence may be used to crack the code.

Let ψ be the entangled state from page 98 and let

$$\pi_{\alpha,\beta}(+,+) := \rho_\psi(P_\alpha \otimes P_\beta)$$

denote the probability that Alice and Bob do the observations P_α and P_β, if they perform the ideal measurements $\{P_\alpha, P_{\alpha+\pi/2}\}$ and $\{P_\beta, P_{\beta+\pi/2}\}$, respectively. Likewise, let

$$\begin{aligned}\pi_{\alpha,\beta}(+,-) &:= \rho_\psi(P_\alpha \otimes P_{\beta+\pi/2}),\\ \pi_{\alpha,\beta}(-,+) &:= \rho_\psi(P_{\alpha+\pi/2} \otimes P_\beta),\\ \pi_{\alpha,\beta}(-,-) &:= \rho_\psi(P_{\alpha+\pi/2} \otimes P_{\beta+\pi/2}),\end{aligned}$$

denote the probabilities that one of them, or both, perform the complementary observation. We calculated on page 99 that

$$\begin{aligned}\pi_{\alpha,\beta}(+,+) = \pi_{\alpha,\beta}(-,-) &= \tfrac{1}{2}\cos(\beta-\alpha)^2,\\ \pi_{\alpha,\beta}(+,-) = \pi_{\alpha,\beta}(-,+) &= \tfrac{1}{2}\sin(\beta-\alpha)^2.\end{aligned}$$

In particular, $\pi_{0,0}(+,+) = \pi_{0,0}(-,-) = \frac{1}{2}$ and $\pi_{0,0}(+,-) = \pi_{0,0}(-,+) = 0$, so the pairs of photons on which Alice and Bob both performed the measurement $\{P_0, P_{\pi/2}\}$ can be used as a secret code.

If Eve wants to find out this secret code, she has no other choice than to perform a measurement on all photons that pass through the glass fiber cable on their way to Bob, since she cannot know in advance which photons are going to be used for the secret code. We claim that if Eve fiddles with these photons in any way such that she gets to know the secret code, then she cannot avoid changing the probabilities in such a way that

$$\pi_{\gamma,0}(+,-) + \pi_{0,-\gamma}(+,-) - \pi_{\gamma,-\gamma}(+,-) \geq 0. \tag{9.4}$$

This is called *Wigner's inequality*. It follows from the assumptions of anticorrelations and nonentanglement:

Lemma 9.2.1 (Wigner's inequality). *Assume that probabilities $\pi_{\alpha,\beta}(\pm,\pm)$ satisfy $\pi_{0,0}(+,-) = \pi_{0,0}(-,+) = 0$ and*

$$\pi_{\alpha,\beta}(\sigma,\tau) = \sum_k p_k\, \pi_\alpha^{1,k}(\sigma)\pi_\beta^{2,k}(\tau) \qquad (\sigma,\tau = +,-),$$

where the $\pi_\alpha^{i,k}(\pm)$ are nonnegative numbers such that $\pi_\alpha^{i,k}(+) + \pi_\alpha^{i,k}(-) = 1$ and the p_k are positive numbers summing up to one. Then (9.4) holds.

Proof Write

$$\pi_{\alpha,\beta}(\sigma,\tau) = \sum_k p_k\, \pi^k_{\alpha,\beta}(\sigma,\tau),$$

where $\pi^k_{\alpha,\beta}(\sigma,\tau) = \pi^{1,k}_\alpha(\sigma)\pi^{2,k}_\beta(\tau)$. Since $\pi_{0,0}(+,-) = \pi_{0,0}(-,+) = 0$ we must have that for each k, either $\pi^{1,k}_0(+) = \pi^{2,k}_0(+) = 1$ or $\pi^{1,k}_0(-) = \pi^{2,k}_0(-) = 1$. In the first case,

$$\begin{aligned}
&\pi^k_{\gamma,0}(+,-) + \pi^k_{0,-\gamma}(+,-) - \pi^k_{\gamma,-\gamma}(+,-) \\
&\quad = \pi^{1,k}_\gamma(+)\cdot 0 + 1\cdot \pi^{2,k}_{-\gamma}(-) - \pi^{1,k}_\gamma(+)\pi^{2,k}_{-\gamma}(-) \\
&\quad = \big(1 - \pi^{1,k}_\gamma(+)\big)\pi^{2,k}_{-\gamma}(-) \geq 0,
\end{aligned}$$

while in the second case

$$\pi^k_{\gamma,0}(+,-) + \pi^k_{0,-\gamma}(+,-) - \pi^k_{\gamma,-\gamma}(+,-) = \pi^{1,k}_\gamma(+)\big(1 - \pi^{2,k}_{-\gamma}(-)\big) \geq 0.$$

Summing up over k we arrive at (9.4). ∎

We claim that Eve's measurement of the secret code necessarily destroys the entanglement of the two photons. Indeed, if Eve wants to be sure that she gets the same code as Alice, she has no other option than to perform the ideal measurement $\{P_0, P_{\pi/2}\}$ on the photon that is on its way to Bob. If she also wants Bob to recieve the secret code, she must send the photon on to Bob after she has performed her measurement, or she must send another photon that is polarized in the direction that she measured. In any case, in doing so, she will have changed the system from the pure state ρ_ψ with

$$\psi = \tfrac{1}{\sqrt{2}}\,\phi_1 \otimes \phi_1 + \tfrac{1}{\sqrt{2}}\,\phi_2 \otimes \phi_2$$

to the mixed state

$$\tfrac{1}{2}\rho_{\phi_1\otimes\phi_1} + \tfrac{1}{2}\rho_{\phi_2\otimes\phi_2},$$

which is not entangled, and therefore satisfies Wigner's inequality.

We now know how Alice can find out from the information that Bob sends her whether Eve has been eavesdropping. If Eve has not interfered with the signal, the quantity in Wigner's inequality is

$$\tfrac{1}{2}\sin(\gamma)^2 + \tfrac{1}{2}\sin(\gamma)^2 - \tfrac{1}{2}\sin(2\gamma)^2,$$

which reaches a minimal value of $-\frac{1}{8}$ at $\gamma = \frac{1}{6}\pi$. From the relative frequencies of outcomes in those measurements where she and Bob did not both perform the measurement $\{P_0, P_{\pi/2}\}$, Alice can check whether Wigner's inequality is violated.

Indeed, Wigner's inequality holds also with + and − reversed, so for greater statistical precision, Alice checks whether the quantity

$$\begin{aligned}&\pi_{\gamma,0}(+,-)+\pi_{0,-\gamma}(+,-)-\pi_{\gamma,-\gamma}(+,-)\\&\quad+\pi_{\gamma,0}(-,+)+\pi_{0,-\gamma}(-,+)-\pi_{\gamma,-\gamma}(-,+)\end{aligned}$$

is close enough to $-\frac{1}{4}$. If she is satisfied with the answer, she knows that the outcomes of the experiments where she and Bob both performed the measurement $\{P_0, P_{\pi/2}\}$ are not known to Eve, and she uses these as a secret code to send her message to Bob.

Exercise 9.2.2 Eve decides to use the following tactic: She cuts the glass fiber between Alice and Bob, blocking all direct communication between them. Instead, Eve communicates now with both Alice and Bob, pretending to be Bob when she communicates with Alice, and pretending to be Alice when she communicates with Bob. In this way, she builds up a secret code with Alice and another secret code with Bob. When Alice sends the coded signal, Eve decodes it using the code she shares with Alice and then codes it with the code she shares with Bob, before passing it on to Bob. Can Alice and Bob do anything to detect this kind of eavesdropping?

Chapter 10

Solutions of chosen exercises

Solution of Exercise 1.2.15 By definition, an operator A is hermitian if and only if $A^* = A$. By the definition of A^*, this is equivalent to

$$\begin{aligned} &\langle\phi|A\psi\rangle = \langle\phi|A^*\psi\rangle \quad \forall\phi,\psi\in\mathcal{H} \\ \Leftrightarrow \quad &\langle\phi|A\psi\rangle = \langle A\phi|\psi\rangle \quad \forall\phi,\psi\in\mathcal{H} \\ \Leftrightarrow \quad &\langle\phi|A|\psi\rangle = \langle\psi|A|\phi\rangle^* \quad \forall\phi,\psi\in\mathcal{H}. \end{aligned}$$

In particular, putting $\psi = \phi$, this implies that

$$\langle\phi|A|\phi\rangle = \langle\phi|A|\phi\rangle^* \quad \forall\phi\in\mathcal{H},$$

which shows that $\langle\phi|A|\phi\rangle \in \mathbb{R}$ for all $\phi\in\mathcal{H}$.
Conversely, if A is a *normal* operator, then by Theorem 1.2.5 there exists an orthonormal basis $\{e(1),\ldots,e(n)\}$ of $\mathcal{H}$ that diagonalizes A. Now $\langle\phi|A|\phi\rangle\in\mathbb{R}$ for all $\phi\in\mathcal{H}$ implies that

$$A_{ii} = \langle e(i)|A|e(i)\rangle \in \mathbb{R}$$

for each $i = 1,\ldots,n$, which by the fact that A is diagonal w.r.t. $\{e(1),\ldots,e(n)\}$ implies that $A_{ji} = (A_{ij})^*$ for each i,j and hence A is hermitian.
Unfortunately, we do not know a priori that A is normal, so we have to find a different proof. For general $\phi,\psi\in\mathcal{H}$ and $\lambda\in\mathbb{C}$, we observe that

$$\underbrace{\langle\phi+\lambda\psi|A|\phi+\lambda\psi\rangle}_{\in\mathbb{R}} = \underbrace{\langle\phi|A|\phi\rangle + \langle\lambda\psi|A|\lambda\psi\rangle}_{\in\mathbb{R}} + \langle\phi|A|\lambda\psi\rangle + \langle\lambda\psi|A|\phi\rangle,$$

so we obtain that

$$\lambda\langle\phi|A|\psi\rangle + \lambda^*\langle\psi|A|\phi\rangle \in \mathbb{R} \quad \forall\phi,\psi\in\mathcal{H},\ \lambda\in\mathbb{C}.$$

We rewrite this expression as

$$\lambda\big[\langle\phi|A|\psi\rangle - \langle\psi|A|\phi\rangle^*\big] + \underbrace{\lambda\langle\psi|A|\phi\rangle^* + \lambda^*\langle\psi|A|\phi\rangle}_{\in\mathbb{R}},$$

where we have used that for any two complex numbers λ, μ, one has

$$\lambda\mu^* + \lambda^*\mu = \lambda\mu^* + (\lambda\mu^*)^* \in \mathbb{R}.$$

It follows that

$$\lambda\big[\langle\phi|A|\psi\rangle - \langle\psi|A|\phi\rangle^*\big] \in \mathbb{R} \quad \forall\phi,\psi\in\mathcal{H},\ \lambda\in\mathbb{C}.$$

It is easy to see that this implies

$$\langle\phi|A|\psi\rangle = \langle\psi|A|\phi\rangle^* \quad \forall\phi,\psi\in\mathcal{H},$$

which by our initial remarks is equivalent to $A = A^*$. ∎

Solution of Exercise 1.2.16 By definition, an operator A is positive if and only if there exists an orthonormal basis $\{e(1),\ldots,e(n)\}$ and nonnegative numbers λ_i such that $A = \sum_i \lambda_i |e(i)\rangle\langle e(i)|$. It follows that for any $\phi\in\mathcal{H}$

$$\langle\phi|A|\phi\rangle = \sum_i \lambda_i \langle\phi|e(i)\rangle\langle e(i)|\phi\rangle = \sum_i \lambda_i \big|\langle\phi|e(i)\rangle\big|^2 \geq 0,$$

proving the implication (1)⇒(2). Conversely, by Exercise 1.2.15, (2) implies that A is a hermitian operator, so there exist an orthonormal basis $\{e(1),\ldots,e(n)\}$ and real numbers λ_i such that $A = \sum_i \lambda_i |e(i)\rangle\langle e(i)|$. If $\lambda_j < 0$ for some j, then

$$\langle e(j)|A|e(j)\rangle = \sum_i \lambda_i \langle e(j)|e(i)\rangle\langle e(i)|e(j)\rangle = \lambda_j < 0,$$

contradicting (2), so we conclude that (2)⇒(1).
For any $B\in\mathcal{L}(\mathcal{H})$ and $\phi\in\mathcal{H}$, we have

$$\langle\phi|B^*B|\phi\rangle = \langle B\phi|B\phi\rangle \geq 0,$$

so (3)⇒(2). To conclude the proof, we will show that (1)⇒(3). Since A is a positive operator, there exists an orthonormal basis $\{e(1),\ldots,e(n)\}$ and nonnegative numbers λ_i such that $A = \sum_i \lambda_i |e(i)\rangle\langle e(i)|$. This allows us to define $\sqrt{A}$ using the functional calculus for normal operators, i.e., we set

$$\sqrt{A} := \sum_i \sqrt{\lambda_i}\, |e(i)\rangle\langle e(i)|.$$

Then $A = \sqrt{A}\sqrt{A}$ and $(\sqrt{A})^* = \sqrt{A}$, so setting $B := \sqrt{A}$ we see that $A = B^*B$ for some $B \in \mathcal{L}(\mathcal{H})$. ∎

Solution of Exercise 1.2.17 If $\mathcal{F} \subset \mathcal{G}$, then we can find an orthonormal basis $\{e(1), \ldots, e(n)\}$ of $\mathcal{H}$ and $0 \leq k \leq m \leq n$ such that $e(1), \ldots, e(k)$ span $\mathcal{F}$ while $e(1), \ldots, e(m)$ span $\mathcal{G}$. Now

$$P_{\mathcal{G}} - P_{\mathcal{F}} = \sum_{i=1}^{m} |e(i)\rangle\langle e(i)| - \sum_{i=1}^{k} |e(i)\rangle\langle e(i)| = \sum_{i=k+1}^{m} |e(i)\rangle\langle e(i)|,$$

which is clearly a positive operator. Conversely, if $P_{\mathcal{F}} \leq P_{\mathcal{G}}$, then for any $\psi \in \mathcal{F}$ we have

$$0 \leq \langle\psi|P_{\mathcal{G}} - P_{\mathcal{F}}|\psi\rangle = \langle\psi|P_{\mathcal{G}}|\psi\rangle - \langle\psi|\psi\rangle = \|P_{\mathcal{G}}\psi\|^2 - \|\psi\|^2.$$

By the hint, this implies $\psi \in \mathcal{G}$.
For completeness, let us also prove the hint. Choosing an orthonormal basis as before, we have

$$\|P_{\mathcal{G}}\psi\|^2 = \sum_{i=1}^{m} |\psi_i|^2 \leq \sum_{i=1}^{n} |\psi_i|^2 = \|\psi\|^2,$$

with equality if and only if $\psi_i = 0$ for all $i = m+1, \ldots, n$, which says that $\psi \in \mathcal{G}$. ∎

Solution of Exercise 1.2.18 Let $\mathcal{F}$ and $\mathcal{G}$ be the spaces that P and Q project on, i.e., $P = P_{\mathcal{F}}$ and $Q = P_{\mathcal{G}}$. Then

$$\begin{aligned}(1-P)Q = 0 \quad &\Leftrightarrow \quad Q = PQ \quad \Leftrightarrow \quad Q\psi = PQ\psi \quad \forall\psi \in \mathcal{H} \\ &\Leftrightarrow \quad \psi = P\psi \quad \forall\psi \in \mathcal{G} \quad \Leftrightarrow \quad \psi \in \mathcal{F} \quad \forall\psi \in \mathcal{G}.\end{aligned}$$

By Exercise 1.2.17, this is in turn equivalent to $P \leq Q$. ∎

Solution of Exercise 2.1.8 It is straightforward to check that

$$X^2 = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad X^3 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

It follows that

$$\begin{pmatrix} a & -c & -b \\ b & a & -c \\ c & b & a \end{pmatrix} = a1 + bX + cX^2.$$

In other words,

$$\mathcal{A} = \{a1 + bX + cX^2 : a, b, c \in \mathbb{C}\} \subset \mathcal{L}(\mathbb{C}^3).$$

Since $1 \in \mathcal{A}$ and

$$\begin{aligned}&(a_1 1 + b_1 X + c_1 X^2)(a_2 1 + b_2 X + c_2 X^2)\\ &\quad = (a_1 a_2 - b_1 c_2 - c_1 b_2)1 + (a_1 b_2 + b_1 a_2 - c_1 c_2)X + (a_1 c_2 + b_1 b_2 + c_1 a_2)X^2,\end{aligned}$$

we see that $\mathcal{A}$ is a sub-algebra of $\mathcal{L}(\mathbb{C}^3)$. In particular, this proves that $\mathcal{A}$ is an algebra. Since $1, X$, and X^2 all commute with each other, it is easy to see that A is in fact abelian. By definition

$$(a1 + bX + cX^2)^* := a^* 1 + b^* X + c^* X^2.$$

This clearly satisfies axioms (vi) and (vii) of an adjoint operation. Let

$$A_1 := a_1 1 + b_1 X + c_1 X^2 \quad \text{and} \quad A_2 := a_2 1 + b_2 X + c_2 X^2.$$

Then

$$\begin{aligned}(A_1 A_2)^* &= \big(a_1 a_2 - b_1 c_2 - c_1 b_2\big)^* 1 + \big(a_1 b_2 + b_1 a_2 - c_1 c_2\big)^* X\\ &\quad + \big(a_1 c_2 + b_1 b_2 + c_1 a_2\big)^* X^2\\ &= \big(a_1^* a_2^* - b_1^* c_2^* - c_1^* b_2^*\big)1 + \big(a_1^* b_2^* + b_1^* a_2^* - c_1^* c_2^*\big)X\\ &\quad + \big(a_1^* c_2^* + b_1^* b_2^* + c_1^* a_2^*\big)X^2\\ &= (a_1^* 1 + b_1^* X + c_1^* X^2)(a_2^* 1 + b_2^* X + c_2^* X^2) = A_1^* A_2^* = A_2^* A_1^*,\end{aligned}$$

which shows that our adjoint operation also satisfies (vii). (Alternatively, one can check (vii) first for the case that A, B are the basis elements $1, X, X^2$ and then use property (vi) (colinearity) to conclude that (vii) holds generally.)
We observe that the representation of the elements of our algebra as matrices is a (faithful) representation of $\mathcal{A}$ *as an algebra*. However, this is not a representation of $\mathcal{A}$ *as a* $*$*-algebra*, since the definition of the adjoint does not coincide with the usual definition of the adjoint of a matrix. In view of this, we start to suspect that perhaps, $\mathcal{A}$ does not have a faithful representation as a $*$-algebra, and hence, by Theorem 2.1.5, the adjoint operation does not satisfy property (viii) (positivity). This is not a proof, however, since the fact that our representation of $\mathcal{A}$ is not a representation of $\mathcal{A}$ as a $*$-algebra does not prove that there cannot exist other representations of $\mathcal{A}$ that have this property.
Nevertheless, we are on the right track. We observe that $X = X^*$ and $X^3 = -1$. Imagine that $\mathcal{A}$ has a faithful representation as a $*$-algebra. Then in this representation, X must be a hermitian operator with the property that $X^3 = -1$. Since X is a hermitian operator, it can be diagonalized w.r.t. an orthonormal basis, and all its eigenvalues are real. But $X^3 = -1$ means that every eigenvalue λ of X

must satisfy $\lambda^3 = -1$. The only real solution of this equation is $\lambda = -1$, so we find that $X = -1$. But this contradicts the assumption that our reprentation is faithful. This proves that $\mathcal{A}$ does not have a faithful representation as a $*$-algebra, and hence, by Theorem 2.1.5, the adjoint operation cannot satisfy property (viii). It is not so easy to find an explicit example of an element $A \in \mathcal{A}$ for which property (viii) does not hold. Nevertheless, a bit of trial and error yields the following. Let $\lambda := e^{i2\pi/6}$ and set

$$A := \lambda 1 + X + \lambda^* X^2.$$

Then, using the facts that $\lambda^*\lambda = 1$, $\lambda^2 = -\lambda^*$, and $(\lambda^*)^2 = -\lambda$, we find that

$$\begin{aligned} A^*A &= (\lambda^* 1 + X + \lambda X^2)(\lambda 1 + X + \lambda^* X^2) \\ &= (1 - \lambda^* - \lambda)1 + (\lambda^* + \lambda - 1)X + (1 - \lambda - \lambda^*)X^2 = 0. \end{aligned}$$

We have seen that $\mathcal{A}$ is an abelian $*$-algebra whose adjoint operation is not positive, and that $\mathcal{A}$ does not have a faithful representation as a $*$-algebra. ∎

Solution of Exercise 2.3.2 For any $A \in \mathcal{L}(\mathcal{H})$, we have

$$\begin{aligned} P_\alpha A P_\alpha &= |\eta(\alpha)\rangle\langle\eta(\alpha)|A|\eta(\alpha)\rangle\langle\eta(\alpha)| \\ &= \langle\eta(\alpha)|A|\eta(\alpha)\rangle|\eta(\alpha)\rangle\langle\eta(\alpha)| = \langle\eta(\alpha)|A|\eta(\alpha)\rangle P_\alpha. \end{aligned}$$

By the linearity of ρ, this implies that

$$\begin{aligned} \rho_\alpha(A) &= \frac{\rho(P_\alpha A P_\alpha)}{\rho(P_\alpha)} = \frac{\rho(\langle\eta(\alpha)|A|\eta(\alpha)\rangle P_\alpha)}{\rho(P_\alpha)} \\ &= \langle\eta(\alpha)|A|\eta(\alpha)\rangle \frac{\rho(P_\alpha)}{\rho(P_\alpha)} = \langle\eta(\alpha)|A|\eta(\alpha)\rangle \end{aligned}$$

whenever $\rho(P_\alpha) \neq 0$. In particular,

$$\rho_\alpha(P_{\alpha'}) = \langle\eta(\alpha)|\eta(\alpha')\rangle\langle\eta(\alpha')|\eta(\alpha)\rangle = \langle\eta(\alpha)|\eta(\alpha')\rangle^*\langle\eta(\alpha)|\eta(\alpha')\rangle = \big|\langle\eta(\alpha)|\eta(\alpha')\rangle\big|^2.$$

To calculate the inner product $\langle\eta(\alpha)|\eta(\alpha')\rangle$, there are several ways to proceed. With respect to the basis $\{e(1), e(2)\}$, the coordinates of $\eta(\alpha)$ are given by

$$\eta_1(\alpha) = \cos(\alpha) \quad \text{and} \quad \eta_2(\alpha) = \sin(\alpha).$$

It follows that

$$\langle\eta(\alpha)|\eta(\alpha')\rangle = \eta_1(\alpha)^*\eta_1(\alpha') + \eta_2(\alpha)^*\eta_2(\alpha') = \cos(\alpha)\cos(\alpha') + \sin(\alpha)\sin(\alpha').$$

We can rewrite this as

$$\begin{aligned}\cos(\alpha)\cos(\alpha') + \sin(\alpha)\sin(\alpha') &= \Re(e^{i\alpha})\Re(e^{-i\alpha'}) - \Im(e^{i\alpha})\Im(e^{-i\alpha'}) \\ &= \Re(e^{i\alpha}e^{-i\alpha'}) = \cos(\alpha - \alpha').\end{aligned}$$

A simpler way to arrive at the same answer is as follows. We observe that $\eta(\alpha)$ and $\eta(\alpha')$ lie in the two-dimensional real space $\{ae(1) + be(2) : a, b \in \mathbb{R}\}$ and a little drawing shows that they are vectors of length one which make an angle $\alpha - \alpha'$ with each other, so their inner product should be $\langle \eta(\alpha)|\eta(\alpha')\rangle = 1 \cdot 1 \cdot \cos(\alpha - \alpha')$. Continuing this last line of argument, we observe that for each α, the vectors $e'(1) := \eta(\alpha)$ and $e'(2) := \eta(\alpha + \pi/2)$ form an orthonormal basis for $\mathcal{H}$. With respect to this new basis, the vectors $\eta(\alpha)$ are given by

$$\eta(\alpha) = 1e'(1) + 0e'(2) \quad \text{and} \quad \eta(\alpha') = \cos(\alpha' - \alpha)e'(1) + \sin(\alpha' - \alpha)e'(2).$$

Thus, the inner product we are looking for is

$$\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \middle| \begin{pmatrix} \cos(\alpha' - \alpha) \\ \sin(\alpha' - \alpha) \end{pmatrix} \right\rangle = \cos(\alpha - \alpha').$$

In any case, we find that

$$\rho_\alpha(P_{\alpha'}) = \left|\langle \eta(\alpha)|\eta(\alpha')\rangle\right|^2 = \cos(\alpha' - \alpha)^2.$$

The final question we need to resolve is when P_α and P_β commute. There are several ways to resolve this.
Let $\mathcal{F}_\alpha := \{c\eta(\alpha) : c \in \mathbb{C}\}$ denote the one-dimensional subspace of $\mathcal{H}$ spanned by $\eta(\alpha)$. Then for any $\psi \in \mathcal{H}$, we have that $P_\alpha P_\beta \psi \in \mathcal{F}_\alpha$ and $P_\beta P_\alpha \psi \in \mathcal{F}_\beta$, so a necessary condition for P_α and P_β to commute is that

$$P_\alpha P_\beta \psi \in \mathcal{F}_\alpha \cap \mathcal{F}_\beta \quad \forall \psi \in \mathcal{H}.$$

If $\alpha - \beta = k\pi$ for some $k \in \mathbb{Z}$, then $\mathcal{F}_\alpha = \mathcal{F}_\beta$ and hence $P_\alpha = P_\beta$ clearly commute. In all other cases, $\mathcal{F}_\alpha \cap \mathcal{F}_\beta = \{0\}$ so we need that $P_\alpha P_\beta \psi = 0 = P_\beta P_\alpha \psi$ for all $\psi \in \mathcal{H}$, which holds if and only if $\alpha - \beta = \pi/2 + k\pi$ for some $k \in \mathbb{Z}$. Concluding, we see that P_α and P_β commute if and only if $\alpha - \beta = k\pi/2$ for some $k \in \mathbb{Z}$
Another way to approach the same problem is to calculate the matrices of P_α and P_β with respect to a suitably chosen basis. By using the basis $\{e'(1), e'(2)\}$ mentioned earlier, we see that whether P_α and P_β commute depends only on the difference $\beta - \alpha$, so we can equivalently check whether P_0 and $P_{\beta-\alpha} =: P_\gamma$ commute.

To do this, we will express P_0 and P_γ in the basis $\{e(1), e(2)\}$. With respect to this basis, we identify an operator A with the matrix given by

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{with} \quad A_{ij} = \langle e(i)|A|e(j)\rangle.$$

In particular

$$\begin{aligned}(P_\gamma)_{ij} &= \langle e(i)|\eta(\gamma)\rangle\langle\eta(\gamma)|e(j)\rangle \\ &= \langle e(i)|\cos(\gamma)e(1) + \sin(\gamma)e(2)\rangle\langle\cos(\gamma)e(1) + \sin(\gamma)e(2)|e(j)\rangle,\end{aligned}$$

which gives

$$P_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad P_\gamma = \begin{pmatrix} \cos(\gamma)^2 & \cos(\gamma)\sin(\gamma) \\ \cos(\gamma)\sin(\gamma) & \sin(\gamma)^2 \end{pmatrix}.$$

It follows that

$$P_0P_\gamma = \begin{pmatrix} \cos(\gamma)^2 & \cos(\gamma)\sin(\gamma) \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad P_\gamma P_0 = \begin{pmatrix} \cos(\gamma)^2 & 0 \\ \cos(\gamma)\sin(\gamma) & 0 \end{pmatrix}.$$

We see from this that

$$[P_0, P_\gamma] = 0 \quad \Leftrightarrow \quad \cos(\gamma)\sin(\gamma) = 0 \quad \Leftrightarrow \quad \gamma = k\pi/2 \text{ for some } k \in \mathbb{Z}.$$

∎

English-Czech glossary

abelian	abelovský
ace	eso
addition	operace sčítání
adjoint	adjoint / hermitovský združení
angular momentum	moment hybnosti
block-diagonal form	
bounded	omezený
closure	uzávěr
clover	kříže
complete	úplný
completion	zúplnění
complex conjugate	
composition	skladání
conditional probability ... given	podmíněná pravděpodobnost ... za podmínky
conditioning	podmíňování
coordinate	souřadnice
density	hustota
diagonalizable	diagonalizovatelný
diamonds	káry
direct sum	direktní suma, příma suma
division ring	okruh dělitelnosti
eigenvector	vlastní vektor
eigenvalue	vlastní číslo
entanglement	entanglement, propletení
entry (of a matrix)	prvek
event	jev, událost
expectation	střední hodnota, očekávání
faithful representation	věrná representace
field (in algebra)	těleso
functional calculus	funkcionální počet, funkcionální kalkulus
hermitian	hermitovský
identity	identita, jednotkový operátor, jednotkový prvek
indicator function	indikátor
inner automorphism	vnitřní isomorphismus
inner product	skalární součin, vnitřní součin
intersection	průnik
jack	svršek
kernel	jádro
matrix	matice
measure	míra
measurement	měření
metric space	metrický prostor
mixed state	smíšený stav

momentum	hybnost
multiplication with scalars	násobení skaláry
multiplicity	násobnost
normed space	normovaný prostor
observable	pozorovatelná
observation	pozorování
origin	počátek, nulový vektor
orthogonal complement	ortogonální doplňek
partition	rozklad
physical quantity	fyzikální veličina
probability law	pravděpodobnostní rozdělení
probability space	pravděpodobnostní prostor
projection operator	projektor
proper subspace	vlastní podprostor
pure state	čistý stav
quantum mechanics	kvantová mechanika
quotient space	kvocientní prostor, zlomkový prostor
random variable	náhodná proměnná
range	obor hodnot, dosah, obraz
reducible	reducibilní
relative frequencies	relativní četnosti
reversible	reversibilní, vratný
root	kořen
self-adjoint	samozdružený
semisimple	poloprostý
separable	separovatelný
set of all subsets of Ω	potence množiny Ω
set operation	množinová operace
simple algebra	prostá algebra
simultaneous measurement	simultání měření
spades	piky
span / to span	lineární obal / lineárně pokrývat
spectral decomposition	spektrální rozklad
state	stav (elementární jev)
state space	stavový prostor
super selection rule	super vyběrové pravidlo
supremum norm	supremová norma

tensor product	tensorový součin
tensor calculus	tensorový počet
time evolution	časový vývoj
trace	stopa
uncertainty relation	principa neurčitosti
union	sjednocení
unit element	jednotkový prvek
wave function	vlnová funkce

Bibliography

[Bel64] J.S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics* 1, 195–200.

[B–W93] C.H. Bennett, G.Brassard, C. Crépeau, 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.

[B–Z99] D. Bouwmeester, J.-W. Pan, M. Daniell, H. Weinfurter, and A. Zeilinger. Observation of Three-Photon Greenberger-Horne-Zeilinger Entanglement. *Phys. Rev. Lett.* 82, 1345–1349, 1999.

[Cir80] B.S. Cirel'son. Quantum generalizations of Bell's inequality. *Lett. Math. Phys.* 4(2), 93-100, 1980. See also: http://www.math.tau.ac.il/~tsirel/Research/mybound/main.html

[CS78] J.F. Clauser and A. Shimony. Bell's theorem: experimental tests and implications. *Reports on Progress in Physics* 41, 1881-1927, 1978.

[Con90] J.B. Conway. *A Course in Functional Analysis. 2nd ed.* Springer, 1990.

[Dav96] K.R. Davidson. *C∗-algebras by example.* Fields Institute monographs 6. AMS, Providence, 1996.

[Dir58] P.A.M. Dirac. *The Principles of Quantum Mechanics,* 4th. edn. Clarendon Press, Oxford, 1958.

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* 47, 777, 1935.

[Gil06] R.D. Gill. Better Bell inequalities (passion at a distance). Preprint, 2006. ArXiv:math.ST/0610115v1

[GHZ89] D.M. Greenberger, M. Horne, and A. Zeilinger. Going beyond Bells theorem. Pages 73–76 in: *Bells Theorem, Quantum Theory, and Conceptions of the Universe.* (M. Kafatos, ed.) Kluwer, Dordrecht, 1989. Available on ArXiv:0712.0921

[G–Z90] D.M. Greenberger, M. Horne, A. Shimony, and A. Zeilinger. Bells theorem without inequalities. *American Journal of Physics* 58(12), 1131–1143, 1990.

[Hun74] T.W. Hungerford. *Algebra.* Springer, New York, 1974.

[Jac51] N. Jacobson. *Lectures in Abstract Algebra.* Part I: Basic Concepts. Van Nostrand, New York, 1951. (Reprint: Springer, New York, 1975.)

[Jac53] N. Jacobson. *Lectures in Abstract Algebra.* Part II: Linear Algebra. Van Nostrand, New York, 1951. (Reprint: Springer, New York, 1975.)

[J–Z00] Quantum Cryptography with Entangled Photons. T. Jennewein, C. Simon, G. Weihs, H. Weinfurter, and A. Zeilinger. *Phys. Rev. Lett.* 84, 4729–4732, 2000.

[KS67] S. Kochen and E. Specker. The problem of hiden variables in quantum mechanics. *J. Math. Mechanics* 17, 59–87, 1967.

[Lan71] S. Lang. *Algebra.* Revised edition. Addison-Wesley, Reading, 1971.

[GHJ89] F.M. Goodman, P. de la Harpe, and V.F.R. Jones. *Coxeter Graphs and Towers of Algebras.* (Mathematical Sciences Research Institute Publications 14.) Springer, New York, 1989.

[Kol33] A. Kolmogorov. Grundbegriffe der Wahrscheinlichkeitsrechnung. *Ergebnisse der Mathematik*, 1933. New edition published by Chelsea Publishing Company, New York, 1946. Translated as *Foundations of the Theory of Probability.* Chelsea Publishing Company, New York, 1956.

[Red87] M. Redhead. *Incompleteness, Nonlocality and Realism; a Prolegogomenon to the Philosophy of Quantum Mechanics.* Clarendon Press, Oxford, 1987.

[Ren70] A. Rényi. *Foundations of Probability.* Holden-Day, San Francisco, 1970.

[Sch35] E. Schrödinger. Die gegenwärtige Situation in der Quantenmechanik, *Die Naturwissenschaften* 23, 807–812, 823–828, and 844–849, 1935.

[Sti55] W.F. Stinespring. Positive functions on C^*-algebras. *Proc. Am. Math. Soc.* 6, 211–216, 1955.

[Swa04] J.M. Swart. *Introduction to Quantum Probability.* Lecture Notes (2004) available from http://staff.utia.cas.cz/swart/

[Tak79] M. Takesaki. *Theory of Operator Algebras I.* Springer, New York, 1979.

Index