Introduction to Quantum Probability

J.M. Swart

May 25, 2007

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], the basis of probability theory is provided by 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 an axiomatic basic 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, and its difficulties. 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

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

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]

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

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

[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

$$\big(\, A_{ij} \,\big)_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$$

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.*

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* $A_{ij} = \delta_{ij}\lambda_i$, where λ_i is the eigenvalue corresponding to the eigenvector ϕ_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

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

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 respective bases $\{e(1), \ldots, e(n)\}$ and $\{f(1), \ldots, f(m)\}$, 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 $\langle\phi|A^* = \langle A\psi|$ 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., $A_{ij} = \delta_{ij}\lambda_i$. 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) By Proposition 1.1.3, each $A \in \mathcal{L}(\mathcal{H})$ has at least one eigenvector ϕ. Using the fact that A is normal, one can show that A maps the space $\{\phi\}^\perp$ into itself. Thus, again by Proposition 1.1.3, 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.

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

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

Exercise 1.2.12 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 1.2.13 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 $A = B^*B$.

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

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

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 a 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. ∎

It is obvious from Proposition 1.3.8 that

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

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 $\{f(1),\ldots,f(nm)\}$ 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 1.3.10 (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 1.3.10 Consider the properties (i)–(v) from Propositions 1.3.8 and 1.3.10. 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 $k \in \mathcal{V}_i'$ and $l \in \mathcal{V}_i'$, 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. ∎

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]

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

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

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

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:

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

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

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

[1] The existence of a unit element is not always included in the definition of an algebra. Actually, depending on the methematical 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 4.5. 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.

Let $\mathcal{A}$ be an algebra. By definition, a *left inverse* of an element $A \in \mathcal{A}$ is an algebra element $B \in \mathcal{A}$ such that $BA = 1$. A *right inverse* of A is a $B' \in R$ such that $AB' = 1$.

Exercise 2.1.6 Let $\mathcal{A}$ be an algebra. Show that if $A \in \mathcal{A}$ has both a left inverse B and a right inverse B', then $B = B'$.

By Exercise 2.1.6, if A has both a left and a right inverse, then the left and right inverse coincide and are necessarily unique. In this case we say that A is *invertible* and we call its unique left and right inverse the *inverse* of A, denoted by A^{-1}. The next lemma, which will be proved in Section 4.3, shows that an element of a Q-algebra is invertible as an algebra element if and only if it is invertible as an operator in some, and hence every representation. In other words: being invertible is a property that is representation independent.

Lemma 2.1.7 (Invertible algebra elements) *Let $\mathcal{H}$ be an inner product space, let $\mathcal{A}$ be a sub-$*$-algebra of $\mathcal{L}(\mathcal{H})$, and let $A \in \mathcal{A}$. Then the following statements are equivalent:*

$$\begin{array}{ll} (1) & A \text{ has a left inverse } B \in \mathcal{L}(\mathcal{H}), \\ (2) & A \text{ has a right inverse } B' \in \mathcal{L}(\mathcal{H}), \\ (3) & A \text{ has an inverse } A^{-1} \in \mathcal{A}. \end{array}$$

It follows that for an element of a Q-algebra, being a unitary operator is a representation independent property. The same is true for being normal, hermitian, or a projection. By Exercise 1.1.2, the spectrum $\sigma(A)$ of an element A of a Q-algebra also does not depend on the representation. By Theorem 1.2.10, every normal operator A can uniquely be written as

$$A = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda, \tag{2.1}$$

where $\{P_\lambda : \lambda \in \sigma(A)\}$ is a partition of the identity. By our previous remarks, this *spectral decomposition* of a normal operator is also representation independent.

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:

$$\begin{array}{ll} \text{(a)} & \mu(\Omega) = 1, \\ \text{(b)} & A, B \subset \Omega,\ A \cap B = \emptyset \Rightarrow \mu(A \cup B) = \mu(A) + \mu(B). \end{array}$$

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, $A^{\rm c}$ 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

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

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:

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

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 1.2.13, 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 we know that someone performs the ideal measurement $\{P_1, \ldots, P_n\}$ on the system, then our knowledge about the system changes. We must describe our changed knowledge 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 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.

On the level of the individual *photons* (light particles), this amounts to performing an ideal measurement, along a prechosen direction, the outcome of which is either that the photon is polarized in that direction, or in the perpendicular direction. Let us denote directions in which polarization can be measured by an angle α. Let $\mathcal{H}$ be a two-dimensional inner product space with orthonormal basis $\{e(1), e(2)\}$. Then our knowledge of the polarization of a single photon can be described by a probability law on the Q-algebra

$$\mathcal{A} = \mathcal{L}(\mathcal{H}).$$

The observation that a photon is polarized in the direction α is described by the projection operator

$$P_\alpha := |\eta_\alpha\rangle\langle\eta_\alpha|,$$

where

$$\eta_\alpha := \cos(\alpha)e(1) + \sin(\alpha)e(2).$$

Note that $P_\alpha = P_{\alpha+\pi}$. An ideal measurement of the polarization of the photon in the direction α is described by the ideal measurement

$$\{P_\alpha, P_{\alpha+\pi/2}\},$$

the outcome of which can be either that the photon is polarized in the direction α, or in the perpendicular direction $\alpha + \pi/2$.

Exercise 2.3.2 Show that the projections P_α and P_β in different directions α and β in general do not commute. Show that the conditional probability of the ideal observation P_α, given that before we have done the observation P_β, is given by $\cos(\beta - \alpha)^2$.

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:

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

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

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

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’). 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 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*

$$\rho(f) := \int f \,\mathrm{d}\mu \tag{2.2}$$

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 4.2.

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

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

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.

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:

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

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

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

Such a function is called a *measure*. If

$$\text{(b)} \quad \mu(\Omega) = 1,$$

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

$$\begin{array}{lll} \text{(a)} & d(x,y) = d(y,x) & (x, y \in E), \\ \text{(b)} & d(x,z) \leq d(x,y) + d(y,z) & (x, y, z \in E), \\ \text{(c)} & d(x,y) = 0 \text{ if and only if } x = y & (x, y \in E). \end{array}$$

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.,

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

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 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

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

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,[1]

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

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.

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 $\mathcal{C}(E)$.*

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

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

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

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

Algebras

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 determine the general structure of Q-algebra's and their representations. In particular, we will prove Lemma 2.1.7 and Theorem 2.4.1. 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 4.5.

4.1 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 $\mathcal{A}$. It is easy to see that $\mathcal{A}^{\rm c}$ is a subalgebra of $\mathcal{L}(\mathcal{H})$. Moreover, 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 4.1.1 (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 a preparatory lemma.

Lemma 4.1.2 *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 $B \in (\mathcal{A}^{\rm c})^{\rm c}$, 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., $P_\mathcal{F} \in \mathcal{A}^{\rm c}$. Hence, if $B \in (\mathcal{A}^{\rm c})^{\rm c}$, 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 4.1.1 Lemma 4.1.2 says that for each $B \in (\mathcal{A}^{\rm c})^{\rm c}$ 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 $B \in (\mathcal{A}^{\rm c})^{\rm c}$ 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 desciribe the commutant $(A^{(n)})^{\rm c}$. With respect to an obvious orthonormal basis for $\mathcal{H}_1, \ldots, \mathcal{H}_n$, 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 $C \in \mathcal{L}(\mathcal{H}_1, \ldots, \mathcal{H}_n)$ 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 4.1.2 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. ∎

4.2 Abelian algebras

In this section, we look at abelian algebras. In particular, we will prove Theorem 2.4.1. Unlike in the previous section, the results in the present section are true only for algebras over the complex numbers.

Theorem 4.2.1 (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*

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

Proof Pick any element $A \in \mathcal{A}$. Obviously, A is a normal operator, so by Theorem 1.2.10 there exists a partition of the identity $\{P_1, \ldots, P_n\}$ such that $A = \sum_{i=1}^{n} a_i P_i$, where the a_i are all different. We claim that each operator of the form

$$B = \sum_{i=1}^{n} b_i P_i \qquad (b_1, \ldots, b_n \in \mathbb{C})$$

is also an element of $\mathcal{A}$. To prove this this claim, we will show that we can find $\lambda_0, \ldots, \lambda_{n-1} \in \mathbb{C}$ such that $\sum_{k=0}^{n-1} \lambda_k A^k = B$, where $A^0 := 1$. Indeed, since

$$\sum_{k=0}^{n-1} \lambda_k A^k = \sum_{i=1}^{n} \Big(\sum_{k=0}^{n-1} \lambda_k a_i^k \Big) P_i,$$

this will be true provided that

$$\sum_{k=0}^{n-1} \lambda_k a_i^k = b_i \qquad (i = 1, \ldots, n).$$

This means that we are looking for a polynomial of degree $n-1$ that passes through all of the points (a_i, b_i). Since the complex numbers are algebraically complete, such a polynomial exists.
Now let $A' \in \mathcal{A}$ be another element of $\mathcal{A}$, different from $A = \sum_{i=1}^n a_i P_i$. By what we have just proved, $P_i \in \mathcal{A}$ for each i, so by the fact that $\mathcal{A}$ is abelian, A' commutes with each P_i. Let $\mathcal{F}_1, \ldots, \mathcal{F}_n$ be orthogonal subspaces of $\mathcal{H}$ such that P_i projects on $\mathcal{F}_i$. Since A' commutes with each P_i, it must respect these subspaces, i.e., $\phi \in \mathcal{F}_i$ implies $A'\phi \in \mathcal{F}_i$. Since A' is a normal operator on $\mathcal{F}_i$, we can find an orthonormal basis of $\mathcal{F}_i$ that diagonalizes A' on $\mathcal{F}_i$. Thus, we can find projection operators $P_{i1}, \ldots, P_{ik(i)}$ such that $P_{ij}P_{ij'} = \delta_{j,j'}$, $\sum_{j=1}^{k(i)} P_{ik(i)} = P_i$, and complex numbers $a'_{i1}, \ldots, a'_{ik(i)}$, all different, such that $A'P_i = \sum_{j=1}^{k(i)} a'_{ij}P_{ij}$. Since $A'P_i \in \mathcal{A}$, by our previous arguments, the projections $P_{i1}, \ldots, P_{ik(i)}$ and all their linear combinations are elements of $\mathcal{A}$. We observe that $A' = \sum_{i=1}^n AP_i = \sum_{i=1}^n \sum_{j=1}^{k(i)} a'_{ij}P_{ij}$.
Continuing this process, we see that we can step by step find partitions of the identity $\{P_1, \ldots, P_n\}$ such that at each moment in our construction, all linear combinations of the $P_1, \ldots, P_n$ are in $\mathcal{A}$, and whevener $\mathcal{A}$ still contains an element A' that is not a linear combinations of the $P_1, \ldots, P_n$, we can refine our partition so that A' is a linear combinations of the new $P_1, \ldots, P_n$. By finite dimensionality, we cannot keep refining partitions ad infinitum, so at some point we are done. ∎

Theorem 4.2.1 has a useful corollary.

Theorem 4.2.2 (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 cannow also easily give the:

Proof of Theorem 2.4.1 By Theorem 4.2.1, 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 4.2.3 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?

4.3 Structure of Q-algebras

After we have resolved the structure of abelian Q-algebras in the previous section, we are now ready to tackle the general, nonabelian case. Let $\mathcal{A}$ be a 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.,

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

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

(d) $\rho(A^*A) = 0$ implies $A = 0$.

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

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

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

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

Lemma 4.3.1 (Existence of a pseudotrace) *On every Q-algebra there exists a faithful pseudotrace.*

Proof By Theorem 2.1.5, $\mathcal{A}$ has a faithful representation. Now the usual trace has all the desired properties. ∎

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

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

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 4.3.2 (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*

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

Proof Imagine that $\mathcal{A}$ has a proper ideal $\mathcal{I}$. By Lemma 4.3.1, we can choose a faithful pseudotrace τ on $\mathcal{A}$. Let $\mathcal{I}$ be the orthogonal complement of $\mathcal{I}$ with respect to the inner product $\langle\cdot|\cdot\rangle_\tau$, 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

$$(A_1 + A_2)(B_1 + B_2) = (A_1B_1 + A_2B_2) \qquad (A_1, B_1 \in \mathcal{I},\ A_2, B_2 \in \mathcal{I}^\perp) \tag{4.1}$$

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 $\mathcal{A}_1$ and $\mathcal{A}_2$ 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 $\mathcal{I}$ and $\mathcal{I}^\perp$ are Q-algebras and that $\mathcal{A}$ is the direct sum of $\mathcal{A}_1$ and $\mathcal{A}_2$ in the sense of $*$-algebras. By symmetry, it suffices to show that $\mathcal{I}$ is a $*$-ideal.

We claim that for any $A \in \mathcal{A}$,

$$A \in \mathcal{I} \text{ if and only if } \langle B|AC\rangle_\tau = 0 \quad \text{for all } B, C \in \mathcal{I}^\perp. \tag{4.2}$$

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 (4.1), 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 (4.2), $\langle B|AC\rangle_\tau = \tau(B^*AC) = \tau((A^*B)^*C) = \langle A^*B|C\rangle_\tau = 0$ for all $B, C \in \mathcal{I}^\perp$, which shows that $A^* \in \mathcal{I}$. ∎

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 4.3.3 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)). 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 $\mathcal{A}$. If $\mathcal{H}$ is a representation of a $*$-algebra and $\psi \in \mathcal{F}^\perp$, then we moreover see that $\langle A\psi|\phi\rangle = \langle\psi|A^*\phi\rangle = 0$ for all $\phi \in \mathcal{F}$, hence $\mathcal{F}^\perp$ is also an invariant subspace. It follows that $\mathcal{H} \cong \mathcal{F} \oplus \mathcal{F}^\perp$. 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 4.3.4 (Decomposition of representations) *Every representation of a Q-algebra can be written as a direct sum of irreducible representations.*

Proof Obvious from the discussion above. ∎

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 4.3.5 (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 Excercise 2.1.2, it suffices to show that each self-adjoint element $B \in \mathcal{A}^{\mathrm{c}}$ is a multiple of the identity. Let $\mathcal{B}$ be the sub-$*$-algebra of $\mathcal{A}$ generated by B, i.e., $\mathcal{B} = \{\sum_{k=1}^{n} b_k B^k : n \geq 0,\ b_k \in \mathbb{C}\}$ where $B^0 := 1$. If B is not a multiple of the identity, then by Theorem 4.2.1, there exists a projection $0, 1 \neq P \in \mathcal{A}^{\mathrm{c}}$. Let $\mathcal{F}$ be the space that P projects upon. Then $\mathcal{F}$ is a proper invariant subspace of $\mathcal{H}$, contradicting our assumptions.
(3)⇒(4): It suffices to show that $\mathcal{L}(\mathcal{H})$ has a trivial center. If the center is not trivial, then by Theorem 4.2.1 there exists a projection $0, 1 \neq P \in \mathcal{L}(\mathcal{H})$ that commutes with all elements of $\mathcal{H}$. In particular, this means that the space that P projects upon is a proper invariant subspace of $\mathcal{L}(\mathcal{H})$. It is easy to see that $\mathcal{L}(\mathcal{H})$ has no proper invariant subspaces.
(4)⇒(1): If $\mathcal{A}$ is not a factor algebra, then by Proposition 4.3.2, 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 4.3.5 has a useful corollary.

Corollary 4.3.6 (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 4.3.5. ∎

Exercise 4.3.7 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 $|e(i)\rangle\langle e(j)| = |e(i)\rangle\langle\phi|\phi\rangle\langle e(j)|$.

4.4 Structure of representations

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 4.4.1 (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 $\mathcal{I}$ is arbitrary, this proves that all irreducible representations of $\mathcal{A}$ are equivalent.

Fix $0 \neq C \in \mathcal{I}$. By Corollary 4.3.6, $\mathcal{H}$ 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

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

By Corollary 4.3.6, 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 4.4.2 (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

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

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 4.4.2 By Proposition 4.3.2, $\mathcal{A}$ 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 4.3.4, 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 4.3.6, 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 4.4.1, 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. ∎

As an application of the results in this section, we can give the:

Proof of Lemma 2.1.7 Let $\mathcal{A}$ be a sub-$*$-algebra of $\mathcal{L}(\mathcal{H})$. We start by observing that an element $A \in \mathcal{L}(\mathcal{H})$ has a left inverse if and only if $\mathrm{Ker}(A) = 0$ and a right inverse if and only if $\mathrm{Ran}(A) = \mathcal{H}$. Therefore, by finite dimensionality, the following statements are equivalent:

(1) A has a left inverse,
(2) A has a right inverse,
(3) A has an inverse.

Therefore, we are done if we can show that whenever $A \in \mathcal{A}$ has an inverse $A^{-1} \in \mathcal{L}(\mathcal{H})$, we have $A^{-1} \in \mathcal{A}$. By Theorem 4.4.2 we can find an orthonormal basis for $\mathcal{H}$ such that a general element $A \in \mathcal{A}$ has a block-diagonal form of the type

$$A = \begin{pmatrix} A_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & A_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & A_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & A_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & A_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & A_3 \end{pmatrix}.$$

(In this example, $\mathcal{A}$ has three irreducible representations with multiplicities 2,1, and 3, respectively.) Now A has an inverse in $\mathcal{L}(\mathcal{H})$ if and only if A_1, A_2, and A_3 all have inverses, in which case

$$A^{-1} = \begin{pmatrix} A_1^{-1} & 0 & 0 & 0 & 0 & 0 \\ 0 & A_1^{-1} & 0 & 0 & 0 & 0 \\ 0 & 0 & A_2^{-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & A_3^{-1} & 0 & 0 \\ 0 & 0 & 0 & 0 & A_3^{-1} & 0 \\ 0 & 0 & 0 & 0 & 0 & A_3^{-1} \end{pmatrix},$$

which lies in $\mathcal{A}$. ∎

4.5 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 4.3.2 says that every Q-algebra is semisimple. Unfortunately, our proof of Proposition 4.3.2 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 4.4.1 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,

$$\langle A + \mathcal{N}, B + \mathcal{N} \rangle := \rho(A^*B) \tag{4.3}$$

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

$$A(B + \mathcal{N}) := AB + \mathcal{N} \qquad (A, B \in \mathcal{A}) \tag{4.4}$$

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 5

States and independence

5.1 States

The next proposition says that if $\mathcal{A}$ is a Q-algebra and τ is a faithful pseudotrace on $\mathcal{A}$, then every probability law ρ has a *density* (or *density operator*) R with respect to τ.

Proposition 5.1.1 (Density operator) *Let $\mathcal{A}$ be a Q-algebra and let τ be a faithful pseudotrace on $\mathcal{A}$. Let $R \in \mathcal{A}$ be positive hermitian such that $\tau(R) = 1$. Then the formula*

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

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

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

$$\rho(A) = \langle R|A\rangle_\tau = \tau(R^*A) \qquad (A \in \mathcal{A}).$$

Since ρ is real,

$$\tau(R^*A^*) = \rho(A^*) = \rho(A)^* = \tau(R^*A)^* = \tau(A^*R) = \tau(RA^*).$$

Since this holds for all $A \in \mathcal{A}$, we must have $R^* = R$, i.e., R is hermitian. Write $R = \sum_i \lambda_i P_i$; assume that one of the eigenvalues λ_i is strictly negative. Then $\rho(P_i) = \tau((\sum_j \lambda_j P_j)P_i) = \lambda_j \tau(P_i^2) < 0)$, which gives a contradiction. Thus R must be positive. ∎

Exercise 5.1.2 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 5.1.3 Show that the pure states on a Q-algebra $\mathcal{A}$ span the space of all linear forms on $\mathcal{A}$.

Let $\mathcal{A}$ be a Q-algebra. In quantum mechanics, it is a (bad) tradition to call a probability ρ on $\mathcal{A}$ a *state*. Note that the set of all probabilities 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 on $\mathcal{A}$. By definition, a *pure state* is a probability ρ that is not a nontrivial convex combination of other states, i.e., 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 that is not a pure state is called a *mixed state.*

Lemma 5.1.4 (Pure states on factor algebras) *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*

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

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 It is easy to see that $\rho(A) := \langle \psi | A | \psi \rangle$ defines a state if $\psi \in \mathcal{H}$ satisfies $\|\psi\| = 1$. Now let ρ be any state and let R be its density with respect to the usual trace on $\mathcal{L}(\mathcal{H})$. Since R is a positive operator, there exists an orthonormal basis $\{e(1), \ldots, e(n)\}$ and nonnegative numbers $p_1, \ldots, p_n$ such that

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

Now $\rho(A) = \mathrm{tr}(RA) = \sum_i p_i \langle e(i) | A | e(i) \rangle$. Since $\mathrm{tr}(R) = 1$, the p_i sum up to one.

It remains to show that states of the form $\rho_\psi(A) := \langle\psi|A|\psi\rangle$ are pure. Imagine that $\rho_\psi = p\rho_1 + (1-p)\rho_2$ for some $0 < p < 1$. By what we have just shown, there exists an orthonormal basis $\{e(1), \ldots, e(n)\}$ and nonnegative numbers $p_1, \ldots, p_n$ such that $\rho_1(A) = \sum_i p_i \langle e(i)|A|e(i)\rangle$. For each ϕ such that $\langle\psi|\phi\rangle = 0$, we have

$$p \sum_i p_i |\langle e(i)|\phi\rangle|^2 = p\rho_1(|\phi\rangle\langle\phi|) \leq \rho_\psi(|\phi\rangle\langle\phi|) = 0,$$

hence for each i such that $p_i > 0$ we must have $\langle e(i)|\phi\rangle = 0$ for each ϕ that is orthogonal to ψ. It follows that there exists one i such that $p_i = 1$ and $e(i) = \lambda\psi$ for some $|\lambda| = 1$. In particular, $\rho_1 = \rho_\psi$. By the same argument, also $\rho_2 = \rho_\psi$ so ρ_ψ is not a nontrivial convex combination of other states. ∎

Let $\mathcal{A}$ be a Q-algebra. By definition, a *minimal projection* 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.

Lemma 5.1.5 (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, for 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.

Proof If $\mathcal{A} = \mathcal{L}(\mathcal{H})$ is a factor algebra, then minimal projections are of the form $P = |\psi\rangle\langle\psi|$ where $\psi \in \mathcal{H}$ satisfies $\|\psi\| = 1$, hence the statement follows from Lemma 5.1.4. The general case follows by writing $\mathcal{A}$ as a direct sum of factor algebras. ∎

Lemma 5.1.5 says, among other things, that every state can be written as a convex combination of pure states. This decomposition is in general not unique! In the special case that our Q-algebra is a factor algebra $\mathcal{L}(\mathcal{H})$, Lemma 5.1.4 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 Excersice 5.1.6 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.

Exercise 5.1.6 Show that the projections in a Q-algebra $\mathcal{A}$ span the whole algebra $\mathcal{A}$. (Hint: Excercise 2.1.2.)

Exercise 5.1.7 Let $\mathcal{A}$ be a Q-algebra and let ρ_1, ρ_2 be states on $\mathcal{A}$. Show that $\rho_1(P) = \rho_2(P)$ 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 excercise shows that in the nonabelian case, the situation is quite different.

Exercise 5.1.8 (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 $0 < \rho(P) < 1$.

5.2 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

$$\alpha(\mathcal{D}) := \mathrm{span}(\{1\} \cup \{D_1 \cdots D_n : n \geq 1,\ D_i \in \mathcal{D} \text{ or } D_i^* \in \mathcal{D}\ \forall i = 1, \ldots, n\}),$$

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 5.2.1 (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) $\sum_{i=1}^{n} \rho(P_{2,i} P_1 P_{2,i}) = \rho(P_1)$ $\quad \forall P_1 \in \mathcal{B}_1$ *projection,* $\{P_{2,1}, \ldots, P_{2,n}\} \subset \mathcal{B}_2$ *partition of the identity,* ρ *state on* $\mathcal{A}$,

(ii) $P_1 P_2 = P_2 P_1$ $\quad \forall P_1 \in \mathcal{B}_1,\ P_2 \in \mathcal{B}_2,\ P_1, P_2$ *projections,*

(iii) $B_1 B_2 = B_2 B_1$ $\quad \forall B_1 \in \mathcal{B}_1,\ B_2 \in \mathcal{B}_2.$

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 $\mathcal{A}$

$$\begin{array}{rl} & \rho(P_2P_1P_2) + \rho((1-P_2)P_1(1-P_2)) = \rho(P_1) \\ \Leftrightarrow & \rho(P_2P_1P_2) + \rho(P_1) + \rho(P_2P_1P_2) - \rho(P_1P_2) - \rho(P_2P_1) = \rho(P_1) \\ \Leftrightarrow & 2\rho(P_2P_1P_2) = \rho(P_1P_2) + \rho(P_2P_1). \end{array}$$

By Excersice 5.1.3, this holds for every state ρ if and only if it holds for every linear form ρ. Hence, this holds if and only if

$$2P_2P_1P_2 = P_1P_2 + P_2P_1. \tag{5.1}$$

Represent $\mathcal{A}$ on some inner product space $\mathcal{H}$ and let $\mathcal{H}_1, \mathcal{H}_2$ be the subspaces that P_1, P_2 project on. By (5.1),

$$P_1P_2\psi = 2P_2P_1P_2\psi - P_2P_1\psi \in \mathcal{H}_2$$

for each ψ, hence $P_1P_2 = P_2P_1P_2$, which together with (5.1) implies that $P_1P_2 = P_2P_1$ for all $P_1 \in \mathcal{B}_1$, $P_2 \in \mathcal{B}_2$.

(ii)⇒(iii): This follows from Excersice 5.1.6.

(iii)⇒(i): Obvious, since

$$\sum_{i=1}^n \rho(P_{2,i}B_1P_{2,i}) = \sum_{i=1}^n \rho(P_{2,i}P_{2,i}B_1) = \sum_{i=1}^n \rho(P_{2,i}B_1) = \rho(1B_1) = \rho(B_1)$$

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

$$\{P_iQ_j : 1 \leq i \leq n,\ 1 \leq j \leq m\}$$

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).

5.3 Independence

By Lemma 5.2.1, performing a measurement on a sub-$*$-algebras $\mathcal{B}_1$ does not have any effect on a sub-$*$-algebras $\mathcal{B}_2$ 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 5.3.1 (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) $\rho(B_1 B_2) = \rho(B_1)\rho(B_2)$ $\forall B_1 \in \mathcal{B}_1,\ B_2 \in \mathcal{B}_2$.

Proof Since $\mathcal{B}_1$ and $\mathcal{B}_2$ commute, $\rho(P_1|P_2) = \rho(P_2 P_1 P_2) = \rho(P_1 P_2 P_2) = \rho(P_1 P_2)$, so (i) is equivalent to

$$\rho(P_1 P_2) = \rho(P_1)\rho(P_2) \tag{5.2}$$

for all projections $P_1 \in \mathcal{B}_1$, $P_2 \in \mathcal{B}_2$ with $\rho(P_2) \neq 0$. In fact, (5.2) is automatically satisfied if $\rho(P_2) = 0$; to see this, note that since $\mathcal{B}_1$ and $\mathcal{B}_2$ commute, $P_1 P_2$ is a projection. Now $P_1 P_2 \leq P_2$, hence $\rho(P_1 P_2) \leq \rho(P_) = 0$. Thus, (i) holds if and only if (5.2) 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 5.1.6), 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 5.3.1 (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 5.3.2 (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)$\Rightarrow$(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 5.3.2. 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

$$(A_1 \otimes A_2)(B_1 \otimes B_2) := (A_1B_1) \otimes (A_2B_2) \qquad (A_1, B_1 \in \mathcal{A}_1,\ A_2, B_2 \in \mathcal{A}_2)$$

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

$$(A_1 \otimes A_2)(\phi_1 \otimes \phi_2) := (A_1\phi_1) \otimes (A_2\phi_2) \qquad (A_1 \in \mathcal{A}_1,\ A_2 \in \mathcal{A}_2,\ \phi_1 \in \mathcal{H}_1,\ \phi_2 \in \mathcal{H}_2) \tag{5.3}$$

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 5.3.3 (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 5.3.2 and Excercise 5.1.3, 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 1.3.10 (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.) Product of three and more Q-algebras and states are defined analoguously.

Chapter 6

Quantum paradoxes

6.1 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 5.1.8, 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.

6.2 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 6.2.1 (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 6.2.1 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 6.1 hold.

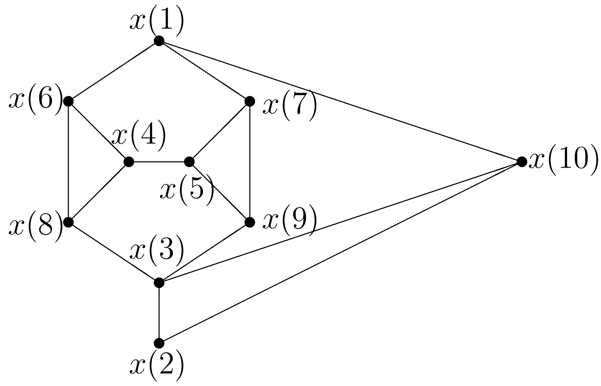

Figure 6.1: Kochen-Specker diagram

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 6.1 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 6.1 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 6.1 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.) ∎

6.3 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 6.3.1 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

$$\psi := \tfrac{1}{\sqrt{2}}\, e \otimes f + \tfrac{1}{\sqrt{2}}\, e' \otimes f', \tag{6.1}$$

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

$$\begin{aligned}
\rho_1(A_1) &= \langle \psi | A_1 \otimes 1 | \psi \rangle \\
&= \tfrac{1}{2}\langle e \otimes f | A_1 \otimes 1 | e \otimes f \rangle + \tfrac{1}{2}\langle e' \otimes f' | A_1 \otimes 1 | e' \otimes f' \rangle \\
&= \tfrac{1}{2}\langle e | A_1 | e \rangle \langle f | f \rangle + \tfrac{1}{2}\langle e' | A_1 | e' \rangle \langle f' | f' \rangle \\
&= \tfrac{1}{2}\langle e | A_1 | e \rangle + \tfrac{1}{2}\langle e' | A_1 | e' \rangle \qquad (A_1 \in \mathcal{L}(\mathcal{H}_1)),
\end{aligned}$$

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 6.3.2 (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*

$$|c_\rho(P_1, P_2) + c_\rho(P_1', P_2) + c_\rho(P_1, P_2') - c_\rho(P_1', P_2')| \leq 2. \tag{6.2}$$

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 (6.2) 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 (6.2) 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

$$
\begin{aligned}
&|c_\rho(P_1, P_2) + c_\rho(P_1', P_2) + c_\rho(P_1, P_2') - c_\rho(P_1', P_2')| \\
&\quad \leq \sum_k p_k \big| c_{\rho_k}(P_1, P_2) + c_{\rho_k}(P_1', P_2) + c_{\rho_k}(P_1, P_2') - c_{\rho_k}(P_1', P_2') \big| \leq 2
\end{aligned}
$$

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 (6.1), which we interpret interms 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,

$$
\psi = \tfrac{1}{\sqrt{2}}\, \eta_\gamma \otimes \zeta_\gamma + \tfrac{1}{\sqrt{2}}\, \eta_{\gamma+\pi/2} \otimes \zeta_{\gamma+\pi/2}. \tag{6.3}
$$

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 (6.3).
The probability of finding one photon polarized in the direction γ and the other photon in the direction $\tilde{\gamma}$ is given by

$$\begin{aligned}\rho_\psi(P_\gamma \otimes P_{\tilde{\gamma}}) &= \rho_\psi(P_0 \otimes P_{\tilde{\gamma}-\gamma}) \\ &= \tfrac{1}{2}\langle e(1) \otimes f(1)|P_0 \otimes P_{\tilde{\gamma}-\gamma}|e(1) \otimes f(1)\rangle \\ &\quad + \tfrac{1}{2}\langle e(2) \otimes f(2)|P_0 \otimes P_{\tilde{\gamma}-\gamma}|e(2) \otimes f(2)\rangle \\ &= \tfrac{1}{2}\langle e(1)|e(1)\rangle\langle f(1)|\zeta_{\tilde{\gamma}-\gamma}\rangle\langle\zeta_{\tilde{\gamma}-\gamma}|f(1)\rangle \\ &= \tfrac{1}{2}\cos(\tilde{\gamma}-\gamma)^2.\end{aligned}$$

(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 (6.2). 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 (6.2) 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

$$\begin{aligned} &\tfrac{1}{2i}\left(e^{i2\alpha} - e^{-i2\alpha}\right) = \tfrac{1}{2i}\left(e^{i6\alpha} - e^{-i6\alpha}\right) \\ \Leftrightarrow \quad & z - z^{-1} = z^3 - z^{-3} \\ \Leftrightarrow \quad & z^6 - z^4 + z^2 - 1 = 0. \end{aligned}$$

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 (6.2) 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].

Bell versus Tsirelson

We have seen that in classical probability theory, the quantity in (6.2) 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]. In fact, there exist several Bell inequalities; the one in (6.2) is just the simplest one. These inequalities have quantum mechanical analogues, the Tsirelson inequalities.

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 (6.2) considers the case $s = m = n = 2$. Numbering the systems, measurements, and outcomes in some arbitrary way, we are interested in $(mn)^s$ conditional probabilities, say

$$p(a_1, \ldots, a_s | b_1 \ldots, b_s),$$

that experiments $b_1,\dots,b_s \in \{1,\dots,m\}$ yield outcomes $a_1,\dots,a_1 \in \{1,\dots,n\}$. 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 $\mathcal{P}_{\rm nosignal}$ of functions p that assign probabilities $p(a_1,\dots,a_s|b_1\dots,b_s)$ to the outcomes of different measurements.
It turns out that not all these probability functions p can arise from classical probability. More precisely, clasically, we imagine that there are certain 'hidden variables' that deterministically predict the outcome of each measurement. Thus, we imagine that

$$p(a_1,\dots,a_s|b_1\dots,b_s)=\sum_h P(h)p_h(a_1,\dots,a_s|b_1\dots,b_s) \tag{6.4}$$

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, $p_h(a_1,\dots,a_s|b_1\dots,b_s)$ is either 0 or 1 for each choice of $a_1,\dots,a_s,b_1\dots,b_s$. Since there are only finitely many such functions, the collection of functions p of the form (6.4) is a convex set $\mathcal{P}_{\rm clasical}$ with finitely many extreme points, which are the functions p_h. It turns out that $\mathcal{P}_{\rm clasical}$ is strictly smaller than $\mathcal{P}_{\rm nosignal}$. 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 $\mathcal{P}_{\rm clasical}$ correspond to inequalities that are not satisfied by general elements of $\mathcal{P}_{\rm nosignal}$. In fact, the Tsirelson inequalities show that $\mathcal{P}_{\rm quantum}$, the quantum analogue of $\mathcal{P}_{\rm clasical}$, is also not equal to $\mathcal{P}_{\rm nosignal}$. 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.

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
eigenvector	vlastní vektor
eigenvalue	vlastní číslo
entanglement	entanglement, propletení
event	jev, událost
expectation	střední hodnota, očekávání
faithful representation	věrná representace
functional calculus	funcionální počet, funcionální kalkulus
hermitian	hermitovský
identity	identita, jednotový operátor, jednotový 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	smíšení stav

momentum	hybnost
multiplication with scalars	násobení skalary
multiplicity	násobnost
normed space	normovaný prostor
observable	pozorovatelná
observation	pozorování
origin	počátek, nulová 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
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	jednotový 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.

[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

[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