Introduction to Quantum Probability

J.M. Swart

July 20, 2004

Contents

Chapter 1

Two kinds of probability

1.1 Introduction

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. The 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 $(\Omega, \mathcal{F}, P)$, where Ω is a set, called the state space, $\mathcal{F}$ is a σ-algebra on Ω, and P 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. Instead of a probability space $(\Omega, \mathcal{F}, P)$, physicists use a C^*-algebra $\mathcal{A}$ and a continuous positive linear form ρ on $\mathcal{A}$ to describe incomplete knowledge about a quantum physical system. Here $\mathcal{A}$ takes the place of the σ-algebra $\mathcal{F}$ and ρ replaces the probability measure P, while the state space Ω has no good quantum mechanical counterpart.

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, and so on.

I could never have written these notes as they are without the help of Roberto Conti, who showed me the proof of Theorem 1.3.7 in [GHJ89, Appendix II.a] and helped me to understand it. The lecture notes of Hans Maassen were also of a great help. I also got some information from the homepage of Reinhard Werner.

1.2 Probability spaces

For any set Ω, we write $\mathcal{P}(\Omega) := \{A : A \subset \Omega\}$ to denote the set of all subsets of Ω. Throughout these notes, we use the symbol $\subset$ in the meaning 'subset of'. Thus, in particular, $\Omega \in \mathcal{P}(\Omega)$. We do not use the symbol $\subseteq$. If we want to indicate that A is a proper subset of Ω, we write: $A \subset \Omega$ and $A \neq \Omega$. On $\mathcal{P}(\Omega)$ there 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 measure*.

Exercise 1.2.1 Show that every probability measure on a finite set Ω is of the form

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

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

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

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

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

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

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

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 measure $\tilde{\mu}$ 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 that 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 measure. 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 measures has no justification, but the point of view taken here is that nature is as it is and does not need justification.

1.3 Quantum probability spaces

Let $\mathcal{A}$ be a linear space over the complex numbers. By definition, in these notes, $\mathcal{A}$ is called an *algebra* (over the complex numbers) if $\mathcal{A}$ is equipped with a multiplication $(A, B) \mapsto AB$ from $\mathcal{A} \times \mathcal{A}$ into $\mathcal{A}$ that is associative, distributive, and has a unit element, i.e.,

$$\begin{array}{rll}
\text{(i)} & (AB)C = A(BC) & (A, B, C \in \mathcal{A}), \\
\text{(ii)} & A(\lambda_1 B_1 + \lambda_2 B_2) = \lambda_1 AB_1 + \lambda_2 AB_2 & (A, B_1, B_2 \in \mathcal{A},\ \lambda_1, \lambda_2 \in \mathbb{C}), \\
\text{(iii)} & (\lambda_1 A_1 + \lambda_2 A_2)B = \lambda_1 A_1 B + \lambda_2 A_2 B & (A_1, A_2, B \in \mathcal{A},\ \lambda_1, \lambda_2 \in \mathbb{C}) \\
\text{(iv)} & \exists 1 \in \mathcal{A} \text{ such that } 1 \neq 0,\ 1A = A = A1 & (A \in \mathcal{A}).
\end{array}$$

Exercise 1.3.1 Prove that each algebra contains at most one element 1 such that $1A = A = A1$ for all $A \in \mathcal{A}$. Such an element is called the *unit element* or the *identity*.

Algebras over the real numbers are defined in the same way, with $\mathbb{C}$ replaced by $\mathbb{R}$. We say that an algebra $\mathcal{A}$ is *abelian* if the multiplication is commutative, i.e.,

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

Let $\mathcal{A}$ be an algebra over the complex numbers. 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:

$$\begin{array}{rll}
\text{(v)} & (A^*)^* = A & (A \in \mathcal{A}), \\
\text{(vi)} & (\lambda_1 A_1 + \lambda_2 A_2)^* = \overline{\lambda}_1 A_1^* + \overline{\lambda}_2 A_2^* & (A_1, A_2 \in \mathcal{A},\ \lambda_1, \lambda_2 \in \mathbb{C}), \\
\text{(vii)} & (AB)^* = B^* A^* & (A, B \in \mathcal{A}).
\end{array}$$

Here $\overline{\lambda}$ denotes the complex conjugate of a complex number λ. An adjoint operation is called *positive* if

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

By definition, a *-*algebra* (pronounce: star-algebra) is an algebra $\mathcal{A}$ over the complex numbers that is equipped with an adjoint operation. We say that $\mathcal{A}$ is a *positive* *-*algebra* if the adjoint operation is positive.

Exercise 1.3.2 Let $\mathcal{H}$ be a finite dimensional linear space over the complex numbers that is equipped with an inner product $\langle \phi | \psi \rangle$. Show that the space $\mathcal{L}(\mathcal{H})$ of linear operators $A : \mathcal{H} \to \mathcal{H}$, equipped with operator multiplication and adjugation (defined as $\langle A^* \phi | \psi \rangle := \langle \phi | A \psi \rangle$), is a positive *-algebra.

Exercise 1.3.3 Show that the properties (i),(ii), and (v)–(vii) imply property (iii). Derive from the properties (i)–(vii) that

(ix) $0A = 0 = A0 \qquad (A \in \mathcal{A})$,
(x) $0^* = 0$,
(xi) $1^* = 1$.

Exercise 1.3.4 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 $\mathrm{Re}(A), \mathrm{Im}(A) \in \mathcal{A}_r$. Show that $A^*A = \mathrm{Re}(A)^2 + \mathrm{Im}(A)^2$ $(A \in \mathcal{A})$.

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

(a) $l(\lambda_1 A_1 + \lambda_2 A_2) = \lambda_1 l(A_1) + \lambda_2 l(A_2) \qquad (A_1, A_2 \in \mathcal{A},\ \lambda_1, \lambda_2 \in \mathcal{B})$,
(b) $l(AB) = l(A)l(B) \qquad (A, B \in \mathcal{A})$,
(c) $l(A^*) = l(A)^* \qquad (A \in \mathcal{A})$.

If l is a bijection then one can check that also l^{-1} is a *-algebra homomorphism. In this case we call l a *-algebra isomorphism and we say that $\mathcal{A}$ and $\mathcal{B}$ are isomorphic (as *-algebras). By definition, a *sub-*-algebra* of a *-algebra $\mathcal{A}$ is a linear subspace $\mathcal{A}' \subset \mathcal{A}$ that is closed under multiplication and adjugation and that contains an element $1'$ such that $1'A = A = A1'$ for all $A \in \mathcal{A}'$. Warning: in general it need not be the case that $1 = 1'$! If $\mathcal{A}'$ is a sub-*-algebra of $\mathcal{A}$, then $\mathcal{A}'$, equipped with the multiplication and adjoint operation from $\mathcal{A}$, is itself a *-algebra.

Exercise 1.3.5 Assume that $\mathcal{A}, \mathcal{B}$ are *-algebras. Show that a *-algebra isomorphism $\psi : \mathcal{A} \to \mathcal{B}$ preserves the identity: $\psi(1) = 1$. Does a *-algebra homomorphism in general preserve the identity?

Exercise 1.3.6 Let $\mathcal{A}'$ be a sub-*-algebra of a positive *-algebra $\mathcal{A}$. Show that $\mathcal{A}'$ is a positive *-algebra.

The next theorem says that all finite dimensional positive *-algebras are sub-*-algebras of some $\mathcal{L}(\mathcal{H})$ as in Exercise 1.3.2.

Theorem 1.3.7 (Representation of positive *-algebras) *Let $\mathcal{A}$ be a finite dimensional positive *-algebra. Then there exists a finite dimensional complex inner product space $\mathcal{H}$ and a sub-*-algebra $\mathcal{A}'$ of $\mathcal{L}(\mathcal{H})$ such that $\mathcal{A}$ is isomorphic to $\mathcal{A}'$. It is possible to choose $\mathcal{H}$ and $\mathcal{A}'$ in such a way that the identity in $\mathcal{A}'$ is the identity in $\mathcal{L}(\mathcal{H})$.*

We postpone the proof of Theorem 1.3.7 till Section 3.6.

Exercise 1.3.8 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

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

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 $a1 + bX + cX^2$.)

Let $\mathcal{A}$ be a finite-dimensional positive *-algebra. By definition, an element $P \in \mathcal{A}$ is called a *projection* if

$$P = P^* \quad \text{and} \quad P^2 = P.$$

Exercise 1.3.9 Show that the identity 1 is a projection. Show that if P is a projection, then $1 - P$ is a projection.

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

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

By definition, a *probability* on $\mathcal{A}$ is a map $\rho : \mathcal{A} \to \mathbb{C}$ with the following properties

(a) $\rho(\lambda_1 A_1 + \lambda_2 A_2) = \lambda_1 \rho(A_1) + \lambda_2 \rho(A_2)$ $\qquad (A_1, A_2 \in \mathcal{A},\ \lambda_1, \lambda_2 \in \mathbb{C})$,
(b) $\rho(A^*) = \overline{\rho(A)}$ $\qquad (A \in \mathcal{A})$,
(c) $\rho(A^* A) \geq 0$ $\qquad (A \in \mathcal{A})$,
(d) $\rho(1) = 1$.

By definition, a finite dimensional *quantum probability space* is a tuple $(\mathcal{A}, \rho)$ where $\mathcal{A}$ is a finite dimensional positive *-algebra and ρ is a probability on $\mathcal{A}$. The interpretation of quantum mechanics is notoriously difficult. For the moment, we will be satisfied with the following interpretation of finite dimensional quantum probability spaces.

1° A finite dimensional 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 an observation. We interpret the projection $1-P$ as the observation that P does not occur. 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 ρ assigns to each projection $P \in \mathcal{A}$ a probability $\rho(P)$. The probability $\rho(P)$ measures how likely we judge it to be that an ideal measurement on the system that can yield the observation P as one of its possible outcomes will actually yield P, if we perform the measurement. If $\rho(P) = 1$ then P occurs surely in any ideal measurement that can yield it. If $\rho(P) > \rho(Q)$ then P occurs more likely, in any ideal measurement that can yield it, than the observation Q. If $\rho(P) = \rho(Q)$ then P is occurs equally likely as Q.

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 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 $\tilde{\rho}$ 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 1.3.10 If ρ is a probability 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 probabilities on $\mathcal{A}$.

A disadvantage of the interpretation of a quantum probability space we have just given is the central role played by the so-called 'ideal measurements'. We will later see that there are also measurements that are not ideal, and this will force us to extend our interpretation.

A lot has been written about the interpretation of quantum mechanics. As a consequence, there are innumerably many different interpretations around, and 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 ρ a 'mixed state', even though it is conceptually something very different from the states ω of 'classical' probability.

Two awkward differences between quantum probability and classical probability strike us immediately.

1° The states ω that play such an important role in classical probability have completely disappeared from the picture.

2° 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 ρ'.

We will see in the next section that if the algebra $\mathcal{A}$ is abelian, 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.

1.4 (Non)commutative probability

Although the quantum probability spaces and their interpretation from Section 1.3 seem rather different from the 'classical' probability spaces from Section 1.2, we will see here that the latter are actually a special case of the former. More precisely, we will show that a finite dimensional 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

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

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.,

$$\begin{array}{ll}(\lambda_1 f_1 + \lambda_2 f_2)(i) := \lambda_1 f_1(i) + \lambda_2 f_2(i) & (f_1, f_2 \in \mathbb{C}^\Omega,\ \lambda_1, \lambda_2 \in \mathbb{C},\ i \in \Omega),\\ (fg)(i) := f(i)g(i) & (f, g \in \mathbb{C}^\Omega,\ i \in \Omega),\\ f^*(i) := \overline{f(i)} & (f \in \mathbb{C}^\Omega,\ i \in \Omega).\end{array}$$

It is clear from the second relation that $\mathbb{C}^\Omega$ is abelian. Note that $\mathbb{C}^\Omega$ satisfies property (viii) from the last section, i.e., $\mathbb{C}^\Omega$ is a positive *-algebra. It turns out that every finite dimensional abelian positive *-algebra is of the form $\mathbb{C}^\Omega$:

Theorem 1.4.1 (Abelian positive *-algebras) *Let $\mathcal{A}$ be a finite dimensional abelian positive *-algebra. Then there exists a finite set Ω such that $\mathcal{A}$ is isomorphic to $\mathbb{C}^\Omega$.*

We defer the proof of Theorem 1.4.1 to Section 3.6.

Exercise 1.4.2 Give an example of a finite dimensional abelian *-algebra that is not isomorphic to a *-algebra of the form $\mathbb{C}^\Omega$ with Ω a finite set.

Let $(\mathcal{A}, \rho)$ be a finite dimensional quantum probability space and assume that $\mathcal{A}$ is abelian. Then by Theorem 1.4.1, we may assume that $\mathcal{A} = \mathbb{C}^\Omega$ for some finite set Ω. We now go through a couple of exercises that will show us that $(\mathbb{C}^\Omega, \rho)$ is equivalent to a classical probability space $(\Omega, \mathcal{P}(\Omega), \mu)$.

Exercise 1.4.3 Show that in the algebra $\mathbb{C}^\Omega$ the projections are given by

$$\mathcal{P}(\mathbb{C}^\Omega) = \{1_A : A \in \mathcal{P}(\Omega)\},$$

where for any subset $A \subset \Omega$ the *indicator function* $1_A \in \mathbb{C}^\Omega$ is defined as

$$1_A(i) := 1 \quad \text{if } i \in A \qquad \text{and} \qquad 1_A(i) := 0 \quad \text{if } i \notin A.$$

Exercise 1.4.4 Show that if ρ is a probability on $\mathbb{C}^\Omega$, then

$$\mu(A) := \rho(1_A) \qquad (A \in \mathcal{P}(\Omega)) \tag{1.1}$$

defines a probability measure μ on Ω, and conversely, for every probability measure μ on Ω there exists a unique probability ρ on $\mathbb{C}^\Omega$ such that (1.1) holds.

The next exercise brings the states ω from classical probability back into the picture. By definition, in these notes, we call a probability ρ on a positive *-algebra $\mathcal{A}$ a *precise state* if

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

Exercise 1.4.5 Show that there are as many precise states on $\mathbb{C}^\Omega$ as Ω has elements. Show that every probability on $\mathbb{C}^\Omega$ can in a unique way be written as a convex combination of these precise states.

We will later see that on a nonabelian positive *-algebra, not every probability can be written as a convex combination of precise states. We conclude this section by showing that conditioning in abelian quantum probability spaces is equivalent to conditioning in classical probability spaces.

Exercise 1.4.6 Let ρ be a probability on $\mathbb{C}^\Omega$ and let μ be the probability measure on Ω defined in (1.1). Show that $B \in \mathcal{P}(\Omega)$ satisfies $\mu(B) > 0$ if and only if $1_B \in \mathcal{P}(\mathbb{C}^\Omega)$ satisfies $\rho(1_B) > 0$. Define $\tilde{\mu}(A) := \mu(A \cap B)/\mu(B)$ $(A \in \mathcal{P}(\Omega))$ and $\tilde{\rho}(f) := \rho(1_B f 1_B)$ $(f \in \mathcal{C}^\Omega)$. Show that $\tilde{\mu}$ and $\tilde{\rho}$ are related as in (1.1).

Exercise 1.4.7 Let $(\mathcal{A}, \rho)$ be a finite dimensional quantum probability space. Let $\{P_1, \ldots, P_n\}$ be a partition of the identity and define $\rho'(A) := \sum_{i=1}^n \rho(P_i A P_i)$. Show that $\rho' = \rho$ if $\mathcal{A}$ is abelian.

We have seen that if the positive *-algebra $\mathcal{A}$ in a quantum probability space is abelian, then we are back at classical probability theory. The next section makes a small digression to infinite dimensional probability spaces. In the Chapter 2, we return to finite dimensions when we recapitulate some basic facts about linear spaces. In Chapter 3 we will study positive *-algebras in more detail from the mathematical point of view, and prove Theorems 1.3.7 and 1.4.1 on the way. After that, we will come back to issues of interpretation, and see that the noncommutative quantum reality is very different from the reality that we are used to.

1.5 Infinite dimensional spaces

In practice, one often needs probability spaces that are not finite dimensional. One reason is that very large but finite spaces are difficult to study directly, so that one is led naturally to infinite spaces as a good approximation. It also seems that many systems in the physical reality are inherently infinite dimensional, although this is hard to say for sure, since one can never exclude the possibility that their true dimension is finite but so large that we don't notice the difference.

Whatever our reasons for studying infinite systems may be, it seems that all systems of interest arise mathematically in some way as a limit of finite systems. This limit may be taken in different ways giving rise to different infinite dimensional structures. We present here two formalisms for infinite dimensional probability that have proved to be very useful; one for classical probability spaces and one for quantum probability spaces.

By definition, a *probability 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:

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

Such a $\mathcal{F}$ is called a *σ-algebra* or *σ-field.*[1] 3° $\mu : \mathcal{F} \to [0, 1]$ is a function with the following properties:

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

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

Probability spaces $(\Omega, \mathcal{F}, \mu)$ of the type we have just described are the standard framework for doing classical probability theory. It should be noted, however, that in practice one usually needs a bit more structure. It seems that in all cases of interest Ω can be equipped with a topology that is generated by a complete metric, and $\mathcal{F}$ is the σ-algebra generated by the open sets of this topology.

The most usual generalization of quantum probability spaces to infinite dimensions is the formalism of C^*-algebras. Before defining these, we recall the definition of a Banach space and a Hilbert space. Let $\mathcal{V}$ be a linear space (possibly infinite dimensional). Recall that a *norm* is a map $\phi \mapsto |\phi|$ from $\mathcal{V}$ into $\mathbb{R}$ such that

$$\begin{array}{rl} \text{(i)} & |\lambda\phi| = |\lambda||\phi|, \\ \text{(ii)} & |\phi + \psi| \leq |\phi| + |\psi|, \\ \text{(iii)} & |\phi| \geq 0, \\ \text{(iv)} & |\phi| = 0 \;\Rightarrow\; \phi = 0. \end{array}$$

A sequence $\phi_n \in \mathcal{V}$ is said to *converge* to a limit ϕ in the norm $|\cdot|$ if

$$\lim_{n\to\infty} |\phi_n - \phi| = 0.$$

[1]This terminology dates from a time that the meaning of the words 'algebra' and 'field' was not as fixed as it is today. Actually, according to the definitions in these notes, $\mathcal{F}$ is neither a field nor an algebra.

By definition, a *Cauchy sequence* is a sequence ϕ_n such that

$$\forall \varepsilon > 0 \; \exists n \text{ such that } |\phi_k - \phi_l| \leq \varepsilon \; \forall k, l \geq n.$$

A linear space $\mathcal{V}$ is called *complete* in the norm $|\cdot|$ if every Cauchy sequence in $\mathcal{V}$ has a limit in $\mathcal{V}$. A *Banach space* is a linear space $\mathcal{V}$ equipped with a norm, such that $\mathcal{V}$ is complete in this norm. A *Hilbert space* is a linear space $\mathcal{H}$ equipped with an inner product $\langle \phi | \psi \rangle$ such that $\mathcal{H}$ is complete in the norm

$$\|\phi\| := \sqrt{\langle \phi | \phi \rangle}.$$

If $\mathcal{H}$ is a normed space and $A : \mathcal{H} \to \mathcal{H}$ is a linear operator, then we define the *operator norm* $\|A\|_\infty$ of A by

$$\|A\|_\infty := \sup_{\phi \in \mathcal{H}} \frac{\|A\phi\|}{\|\phi\|}.$$

The space of continuous linear operators $A : \mathcal{H} \to \mathcal{H}$ is

$$\mathcal{L}(\mathcal{H}) := \{A : \|A\| < \infty\}.$$

By definition, a *C^*-algebra* is a *-algebra $\mathcal{A}$ (possibly infinite dimensional) on which a norm $A \mapsto \|A\|$ is defined with the following properties:

(a) $\mathcal{A}$ is complete in the norm $A \mapsto \|A\|$,
(b) $\|AB\| \leq \|A\| \|B\|$,
(c) $\|A^* A\| = \|A\|^2$.

Note that property (c) implies that the adjoint operation is positive. If $\mathcal{H}$ is a Hilbert space then $\mathcal{L}(\mathcal{H})$, equipped with operator multiplication, the usual adjoint, and the operator norm, is a C^*-algebra. In analogy with Theorem 1.3.7 one has the following theorem about C^*-algebras, that we will not prove in these notes. For a proof, see any book on C^*-algebras, for example [Dav96, Theorem I.9.12].

Theorem 1.5.1 (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 $\mathcal{A}'$ and $\mathcal{A}'$ contains the identity in $\mathcal{L}(\mathcal{H})$.*

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

Quantum probability spaces $(\mathcal{A}, \rho)$ of the type we have just described are a widely used framework for doing quantum probability theory, but some variations are possible. In practice one usually needs a bit more structure, in particular separability. On the other hand, the assumption that a C^*-algebra should contain an identity is often dropped.

There is now also an infinite dimensional analogue of Theorem 1.4.1. (See, for example, [Dav96, Theorem I.3.1].)

Theorem 1.5.2 (Abelian C^*-algebras) *Every abelian C^*-algebra is isomorphic to a space of the form*

$$\mathcal{C}(X) := \{f : X \to \mathbb{C} \text{ continuous}\}$$

where X is a compact Hausdorff space, and $\mathcal{C}(X)$ is equipped with the norm $\|f\|_\infty := \sup_{x\in X} |f(x)|$.

In the situation of Theorem 1.5.2, it can moreover be shown that probabilities on $\mathcal{C}(X)$ correspond to probability measures on X, equipped with the σ-algebra generated by the open sets (see [Dav96, Exercise I.29]). Thus, every abelian quantum probability space $(\mathcal{A}, \rho)$ corresponds to a classical quantum probability space $(\Omega, \mathcal{F}, \mu)$. The converse of this statement is not true, however. This is a result of the fact that in going to infinite dimensions we have made different continuity assumptions in the classical and the quantum case.

Chapter 2

Linear spaces

2.1 Fields

By definition, a *ring* is a set R that is equipped with an addition $(x, y) \mapsto x + y$ and a multiplication $(x, y) \mapsto xy$. It is supposed that the addition is associative, abelian, has a neutral element 0 called zero, and that each element x has an inverse element (with respect to addition) denoted by $-x$. The multiplication is associative, has a neutral element $1 \neq 0$ called identity, and the multiplication is distributive with respect to the addition. In formulas:

$$\begin{array}{rl}
\text{(i)} & (x+y)+z = x+(y+z), \\
\text{(ii)} & x+y = y+x, \\
\text{(iii)} & x+0 = x, \\
\text{(iv)} & x+(-x) = 0, \\
\text{(v)} & (xy)z = x(yz), \\
\text{(vi)} & x1 = x = 1x, \\
\text{(vii)} & x(y+z) = xy+xz \quad \text{and} \quad (x+y)z = xz+yz.
\end{array}$$

We write $x^n := x \cdots x$ (the product of n times x). Note that a ring is almost the same as an algebra, except that multiplication with scalars is not defined. In fact, much of the theory of algebras that we will meet in these electure notes applies more generally to rings. Good references are [Lan71, Hun74].

Let R be a ring. By definition, a *left inverse* of an element $x \in R$ is a $y \in R$ such that $yx = 1$. A *right inverse* of x is a $y' \in R$ such that $xy' = 1$.

Exercise 2.1.1 Show that if $x \in R$ has both a left inverse y and a right inverse y', then $y = y'$.

By Exercise 2.1.1, if x 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 x is *invertible* and we call its unique left and right inverse the *inverse* of x, denoted by x^{-1}.

A ring is called a *division ring* if every nonzero element has an inverse:

$$\text{(viii)} \quad x^{-1}x = 1 = xx^{-1} \quad (x \neq 0).$$

A division ring is called a *field* if the multiplication is abelian:

$$\text{(ix)} \quad xy = yx.$$

With an occasional exception, the only fields that we will meet in these notes are $\mathbb{R}$ and $\mathbb{C}$.

By definition, a *subfield* R' of a field R is a subset $R' \subset R$ such that R' contains 0 and 1, R' is closed under addition and multiplication, and R', equipped with the addition and multiplication from R, is itself a field.

Exercise 2.1.2 Let $R' \subset R$ be a subfield. Show that if $0 \neq x \in R'$ then $-x \in R'$ and $x^{-1} \in R'$. (Hint: use the uniqueness of inverses.)

By definition, an *extension* of a field R is a field $\overline{R}$ such that R is a subfield of $\overline{R}$. We call $\overline{R}$ an *algebraic extension* if every element x of $\overline{R}$ is a solution of a nontrivial polynomial equation in R, i.e., there exist $a_0, \ldots, a_n \in R$, not all zero, such that

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = 0.$$

We will need the following theorem from algebra. For a proof, see [Hun74, Theorem V.3.3].

Theorem 2.1.3 (Algebraically closed fields) *Let R be a field. Then the following statements are equivalent.*

(i) *Every equation of the form $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = 0$ with $n \geq 1$ and $a_0, \ldots, a_n \in R$ fixed has a solution $x \in R$.*

(ii) *Every algebraic extension $\overline{R}$ of R satisfies $\overline{R} = R$.*

If a field R satisfies the equivalent conditions (i) and (ii) from Theorem 2.1.3 then R is called *algebraically closed*. The following fact is well-known:

Theorem 2.1.4 (Complex numbers are algebraically closed) *The field $\mathbb{C}$ of the complex numbers is algebraically closed.*

2.2 Linear spaces

By definition, a *linear space* (or *vector space*) over a division ring R is a set $\mathcal{V}$ on which an addition $(\phi, \psi) \mapsto \phi + \psi$ and multiplication with scalars $(x, \phi) \mapsto x\phi$ are defined, such that

$$\begin{array}{rl}
\text{(i)} & (\phi + \psi) + \chi = \phi + (\psi + \chi), \\
\text{(ii)} & \phi + \psi = \psi + \phi, \\
\text{(iii)} & \phi + 0 = \phi, \\
\text{(iv)} & \phi + (-\phi) = 0, \\
\text{(v)} & (xy)\phi = x(y\phi), \\
\text{(vi)} & 1\phi = 1, \\
\text{(vii)} & x(\phi + \psi) = x\phi + x\psi \quad \text{and} \quad (x + y)\phi = x\phi + y\phi.
\end{array}$$

Here $\phi, \psi, \chi \in \mathcal{V}$ and $x, y \in R$. In general R does not have to be a field. However, the linear spaces that we will meet in these notes are (with an occasional exception) linear spaces over $R = \mathbb{R}$ or $R = \mathbb{C}$.

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

$$\mathrm{span}(\mathcal{W}) := \{x_1\phi_1 + \cdots + x_n\phi_n : \phi_1, , \ldots, \phi_n \in \mathcal{W}\}$$

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 $\mathcal{V} = \mathrm{span}(\mathcal{W})$.

A finite collection $\{\phi_1, \ldots, \phi_n\}$ of elements of a linear space $\mathcal{V}$ is called *linearly independent* if the equation

$$x_1\phi_1 + \cdots + x_n\phi_n = 0$$

has no other solutions than $x_1 = x_2 = \cdots = x_n = 0$. If moreover $\{\phi_1, \ldots, \phi_n\}$ spans $\mathcal{V}$ then we call $\{\phi_1, \ldots, \phi_n\}$ a *basis* for $\mathcal{V}$. If $\{\phi_1, \ldots, \phi_n\}$ is a basis for $\mathcal{V}$, then every $\psi \in \mathcal{V}$ can in a unique way be written as

$$\psi = x_1\phi_1 + \cdots + x_n\phi_n.$$

Proposition 2.2.1 (Existence of a basis) *Every finite dimensional linear space has a basis.*

Proof By definition of finite dimensionality, there exists a finite set $\mathcal{W}$ such that $\mathrm{span}(\mathcal{W}) = \mathcal{V}$. We claim that we can find a finite subset $\mathcal{W}'$ of $\mathcal{W}$ that is a basis for $\mathcal{V}$. If $\mathcal{W}$ contains no nonzero elements then $\mathcal{V} = \{0\}$ and we can choose

$\mathcal{W}' = \emptyset$. Now we proceed by induction. Assume that $\{\phi_1, \ldots, \phi_n\} \subset \mathcal{W}$ are linearly independent, but that they do not span $\mathcal{V}$ yet. Then there must be a $\phi_{n+1} \in \mathcal{W}$ such that $\phi_{n+1} \not\in \operatorname{span}\{\phi_1, \ldots, \phi_n\}$. We claim that $\{\phi_1, \ldots, \phi_{n+1}\}$ are linearly independent. For assume that

$$x_1\phi_1 + \cdots + x_{n+1}\phi_{n+1} = 0$$

has a solution with $x_1, \ldots, x_{n+1}$ not all zero. Since $\{\phi_1, \ldots, \phi_n\}$ is linearly independent, this is possible only of $x_{n+1} \neq 0$. But then

$$\phi_{n+1} = (-x_{n+1}^{-1}x_1)\phi_1 + \cdots + (-x_{n+1}^{-1}x_n)\phi_n,$$

contradicting the assumption that $\phi_{n+1} \not\in \operatorname{span}\{\phi_1, \ldots, \phi_n\}$. ∎

Note that in the proof above we have used the fact that every nonzero element of R has an inverse but not that R is abelian. This is why linear spaces can be defined not only over fields but more generally over division rings.

If $\mathcal{V}$ is finite dimensional, then one can check that all bases have the same number of elements n. This number is called the *dimension* $\dim(\mathcal{V})$ of $\mathcal{V}$.

2.3 Finite dimensional linear spaces

In this section we recapitulate some basic facts about linear spaces. Unless stated otherwise, all linear spaces are finite dimensional linear spaces over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.

Let $\mathcal{V}_1, \mathcal{V}_2$ be linear spaces. By definition, a map $l : \mathcal{V}_1 \to \mathcal{V}_2$ is called *linear* if

$$l(\lambda_1\phi_1 + \lambda_2\phi_2) = \lambda_1 l(\phi_1) + \lambda_2 l(\phi_2) \qquad (\lambda_1, \lambda_2 \in \mathbb{K},\ \phi_1, \phi_2 \in \mathcal{V}_1).$$

We denote the space of all linear maps from $\mathcal{V}_1$ into $\mathcal{V}_2$ by $\mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$. In an obvious way $\mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$ is itself a linear space. If $l \in \mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$, $\{\phi_1, \ldots, \phi_n\}$ is a basis for $\mathcal{V}_1$, and $\{\psi_1, \ldots, \psi_m\}$ is a basis for $\mathcal{V}_2$, then

$$l\Big(\sum_{j=1}^n \lambda_j\phi_j\Big) = \sum_{i=1}^m \sum_{j=1}^n l_{ij}\lambda_j\psi_i,$$

where

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

is called the *matrix* of l with respect to the bases $\{\phi_1, \ldots, \phi_n\}$ and $\{\psi_1, \ldots, \psi_m\}$.

Exercise 2.3.1 If $l \in \mathcal{L}(\mathcal{V}_2, \mathcal{V}_3)$ and $m \in \mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$, then show that

$$(lm)_{ij} = \sum_k l_{ik} m_{kj}.$$

The *kernel* and *range* of a linear operator $l \in \mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$ are defined by

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

One has

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

If a linear map $l : \mathcal{V}_1 \to \mathcal{V}_2$ is a bijection then one can check that its inverse l^{-1} is also linear. In this case we call l a *linear isomorphism* and we say that $\mathcal{V}_1$ and $\mathcal{V}_2$ are isomorphic (as linear spaces). Two finite dimensional linear spaces are isomorphic if and only if their dimensions agree. A linear map is a bijection if and only if $\mathrm{Ker}(l) = \{0\}$ and $\mathrm{Ran}(l) = \mathcal{V}_2$. This is equivalent to $\mathrm{Ker}(l) = \{0\}$ and $\dim(\mathcal{V}_1) = \dim(\mathcal{V}_2)$.

Let $\mathcal{V}$ be a linear space and let $\mathcal{W}$ be a linear subspace of $\mathcal{V}$. For any $\phi \in \mathcal{V}$ write $\phi + \mathcal{W} := \{\phi + \psi : \psi \in \mathcal{W}\}$. 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

$$\lambda_1(\phi_1 + \mathcal{W}) + \lambda_2(\phi_2 + \mathcal{W}) := (\lambda_1\phi_2 + \lambda_2\phi_2) + \mathcal{W} \qquad (\lambda_1, \lambda_2 \in \mathbb{K},\ \phi_1, \phi_2 \in \mathcal{V}).$$

Exercise 2.3.2 Show that linear combinations in $\mathcal{V}/\mathcal{W}$ are well-defined, i.e., if $\phi_1 + \mathcal{W} = \tilde{\phi}_1 + \mathcal{W}$ and $\phi_2 + \mathcal{W} = \tilde{\phi}_2 + \mathcal{W}$, then $(\lambda_1\phi_2 + \lambda_2\phi_2) + \mathcal{W} = (\lambda_1\tilde{\phi}_2 + \lambda_2\tilde{\phi}_2) + \mathcal{W}$.

Exercise 2.3.3 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 2.3.4 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 2.3.5 Let $\mathcal{V}_3 \subset \mathcal{V}_2 \subset \mathcal{V}_1$ be linear spaces. Show that there exists a natural linear isomorphism

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

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

$$\alpha(\phi_1, \ldots, \phi_n) + \beta(\psi_1, \ldots, \psi_n) := (\alpha\phi_1 + \beta\psi_1, \ldots, \alpha\phi_n + \beta\psi_n).$$

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 $\mathcal{V}_1, \ldots, \mathcal{V}_n$. One has

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

If $\mathcal{V}_1, \mathcal{V}_2$ are linear subspaces of $\mathcal{V}$ such that $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$, then the *projection* on $\mathcal{V}_1$ with respect to this decomposition is the map $P : \mathcal{V} \to \mathcal{V}_1$ defined by

$$P(\phi_1 + \phi_2) := \phi_1 \qquad (\phi_1 \in \mathcal{V}_1,\ \phi_2 \in \mathcal{V}_2).$$

Note that this is a good definition since every $\psi \in \mathcal{V}$ can in a unique way be written as $\psi = \phi_1 + \phi_2$ with $\phi_1 \in \mathcal{V}_1$ and $\phi_2 \in \mathcal{V}_2$. Warning: the definition of P depends not only on $\mathcal{V}_1$ but also on the choice of $\mathcal{V}_2$!

Exercise 2.3.6 Show that

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

Here $\cong$ means that the two spaces are in a natural way isomorphic. 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?

By definition,

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

is the *dual* of the linear space $\mathcal{V}$. The elements of $\mathcal{V}'$ are called *linear forms* on $\mathcal{V}$. The dual space $\mathcal{V}'$ has the same dimension as $\mathcal{V}$. If $\{\phi_1, \ldots, \phi_n\}$ is a basis for $\mathcal{V}$ then the linear forms $\{l_1, \ldots, l_n\}$ given by

$$l_i(\phi_j) := \delta_{ij}$$

form a basis of $\mathcal{V}'$, called the *dual basis* with respect to $\{\phi_1, \ldots, \phi_n\}$. Here δ_{ij} is Dirac's δ-notation

$$\delta_{ii} := 1, \quad \text{and} \quad \delta_{ij} := 0 \quad (i \neq j).$$

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 $L \in \mathcal{L}_\phi(\mathcal{V}', \mathbb{K})$ 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 $l \in \mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$, then by definition its *adjoint* is the linear map $l' \in \mathcal{L}(\mathcal{V}_2', \mathcal{V}_1')$ defined by

$$l'(l_2) := l_2 \circ l \qquad (l_2 \in \mathcal{V}_2'),$$

where $\circ$ denotes composition.

If $\mathcal{V}$ is a linear space, then 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.*

Exercise 2.3.7 Show that $\mathcal{L}(\mathcal{V})$ with the usual operator multiplication is an algebra.

If $\mathcal{V}_1, \ldots, \mathcal{V}_n$ are linear spaces and $A_1 \in \mathcal{L}(\mathcal{V}_1), \ldots, A_n \in \mathcal{L}(\mathcal{V}_n)$ are linear operators, then we let $A_1 \oplus \cdots \oplus A_n$ denote the linear operator in $\mathcal{L}(\mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_n)$ given by

$$(A_1 \oplus \cdots \oplus A_n)(\phi_1, \ldots, \phi_n) := (A_1\phi_1, \ldots, A_n\phi_n).$$

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; the right-hand side is independent of the choice of the basis $\{\phi_1, \ldots, \phi_n\}$ of $\mathcal{V}$. The trace is linear and satisfies

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

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

The following proposition holds only for linear spaces over the complex numbers. Since we will not need it in what follows, we skip the proof.

Proposition 2.3.9 (Nonempty spectrum) *Let $\mathcal{V}$ 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 for $\mathcal{V}$ consisting of eigenvectors. 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.

2.4 Inner product spaces

Let $\mathcal{H}$ be a linear space over $\mathbb{K} = \mathbb{R}$ of $\mathbb{C}$. We continue to assume that all linear spaces are finite dimensional.

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

$$\begin{array}{rll}
\text{(i)} & \langle\phi|\lambda_1\psi_1 + \lambda_2\psi_2\rangle = \lambda_1\langle\phi|\lambda_1\psi_1\rangle + \lambda_2\langle\phi|\lambda_1\psi_2\rangle & (\phi, \psi_1, \psi_2 \in \mathcal{H},\ \lambda_1, \lambda_2 \in \mathbb{C}), \\
\text{(ii)} & \langle\phi|\psi\rangle = \overline{\langle\psi|\phi\rangle} & (\phi, \psi \in \mathcal{H}), \\
\text{(iii)} & \langle\phi|\phi\rangle \geq 0 & (\phi \in \mathcal{H}), \\
\text{(iv)} & \langle\phi|\phi\rangle = 0 \ \Rightarrow\ \phi = 0. &
\end{array}$$

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 $\langle\psi|\phi\rangle$.

Exercise 2.4.1 Let $\mathcal{H}_1, \mathcal{H}_2$ be inner product spaces and let $l : \mathcal{H}_1 \to \mathcal{H}_2$ be a linear map such that $\langle l(\phi)|l(\psi)\rangle_2 = \langle \phi|\psi\rangle_1$ for all $\phi, \psi \in \mathcal{H}_1$, 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 $\mathcal{H}_2$. Show that $\mathrm{Ker}(l) = \{0\}$.

As the previous exercise shows, any linear map $l : \mathcal{H}_1 \to \mathcal{H}_2$ between two inner product spaces preserving the inner product is one-to-one. If moreover $\mathrm{Ran}(l) = \mathcal{H}_2$ then l is a bijection and also $l^{-1} : \mathcal{H}_2 \to \mathcal{H}_1$ preserves the inner product. In this case we say that l is a *unitary* map.

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 $\langle\psi| \in \mathcal{H}' = \mathcal{L}(\mathcal{H}, \mathbb{K})$ and $|\psi\rangle \in \mathcal{L}(\mathbb{K}, \mathcal{H}) \cong \mathcal{H}$ by

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

Then for any $\phi, \psi \in \mathcal{H}$, the composition $\langle\phi|\,|\psi\rangle$ is an operator in $\mathcal{L}(\mathbb{K}, \mathbb{K})$ that can be associated with the number $\langle\phi|\psi\rangle \in \mathbb{K}$. One has the important fact:

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

and $\phi \mapsto \langle\phi|$ is a colinear (i.e., $\langle\lambda_1\phi_1 + \lambda_2\phi_2| = \overline{\lambda}_1\langle\phi_1| + \overline{\lambda}_2\langle\phi_2|$) bijection from $\mathcal{H}$ into the dual space $\mathcal{H}'$.[1]

We equip the dual space $\mathcal{H}'$ with the *dual inner product*

$$\big\langle\langle\phi|\big|\langle\psi|\big\rangle := \langle\psi|\phi\rangle.$$

A basis $\{\phi_1, \ldots, \phi_n\}$ of an inner product space is called *orthonormal* if

$$\langle\phi_i|\phi_j\rangle = \delta_{ij}.$$

Exercise 2.4.2 If $\{\phi_1, \ldots, \phi_n\}$ is an orthonormal basis for $\mathcal{H}'$, then what is the corresponding dual basis for $\mathcal{H}'$? Is it orthonormal?

[1]By definition, the complex conjugate of a complex linear space $\mathcal{V}$ is tuple $(\mathcal{W}, \gamma)$ where $\mathcal{W}$ is a complex linear space and $\gamma : \mathcal{V} \to \mathcal{W}$ is a colinear bijection. If $(\mathcal{W}', \gamma')$ is another such tuple, then $l := \gamma' \circ \gamma^{-1}$ is a linear bijection $l : \mathcal{W} \to \mathcal{W}'$ such that $l(\gamma(\phi)) = \gamma'(\phi)$ for all $\phi \in \mathcal{V}$. This shows that $(\mathcal{W}, \gamma)$ and $(\mathcal{W}', \gamma')$ are in a natural way isomorphic. Therefore, it makes sense to talk about *the* complex conjugate of a complex linear space $\mathcal{V}$. One usually writes $\overline{\mathcal{V}}$ instead of $\mathcal{W}$ and $\overline{\phi}$ instead of $\gamma(\phi)$. It is easy to see that there exists a natural isomorphism $\mathcal{V} \cong \overline{\overline{\mathcal{V}}}$. If $\mathcal{H}$ is a complex inner product space then $\overline{\phi} \mapsto \langle\phi|$ is a natural linear isomorphism between $\overline{\mathcal{H}}$ and $\mathcal{H}'$.

If $\{\phi_1, \ldots, \phi_n\}$ is an orthonormal basis of $\mathcal{H}$, then one has the important relation

$$\sum_i |\phi_i\rangle\langle\phi_i| = 1,$$

where 1 denotes the identity in $\mathcal{L}(\mathcal{H})$. Note that $|\phi_i\rangle\langle\phi_i|$ is the composition of an operator in $\mathcal{L}(\mathcal{H}, \mathbb{K})$ and an operator in $\mathcal{L}(\mathbb{K}, \mathcal{H})$, hence an operator in $\mathcal{L}(\mathcal{H}, \mathcal{H}) = \mathcal{L}(\mathcal{H})$.

Exercise 2.4.3 Let $\mathcal{H}_1, \mathcal{H}_2$ be inner product spaces, let $\{\phi_1, \ldots, \phi_n\}$ be an orthonormal basis of $\mathcal{H}_1$, and let $\{\psi_1, \ldots, \psi_n\}$ be an orthonormal basis of $\mathcal{H}_2$. Show that the matrix of a linear map $l \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ with respect to this bases is given by

$$l_{ij} = \langle\psi_i|l\phi_j\rangle = \langle\psi_i|l|\phi_j\rangle.$$

Show that l can be written as

$$l = \sum_{ij} l_{ij}|\psi_i\rangle\langle\phi_j|.$$

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

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

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 $\mathcal{H}_2$. Obviously,

$$(l^*)^* = l.$$

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

$$(lm)^* = m^*l^*.$$

Exercise 2.4.4 Describe the relation between the adjoint l' of a linear map $l \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ as defined on page 25 and the adjoint l^* defined here.

Exercise 2.4.5 Let $\mathcal{H}_1, \mathcal{H}_2$ be inner product spaces and let $l : \mathcal{H}_1 \to \mathcal{H}_2$ be linear. Show that l is unitary if and only if l is a bijection and $l^{-1} = l^*$.

Exercise 2.4.6 Let $\mathcal{H}_1, \mathcal{H}_2$ be inner product spaces and let $l : \mathcal{H}_1 \to \mathcal{H}_2$ be linear. Show that $\langle\psi|l^* = \langle l\psi|$ for all $\psi \in \mathcal{H}_1$.

Exercise 2.4.7 Equip $\mathbb{K}$ with the inner product $\langle\lambda|\mu\rangle := \overline{\lambda}\mu$. Check that the adjoint of an element $\lambda \in \mathbb{K}$ is the complex conjugate of that element. Show that for any inner product space $\mathcal{H}$,

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

and

$$(|\phi\rangle\langle\psi|)^* = |\psi\rangle\langle\phi| \qquad (\phi, \psi \in \mathcal{H}).$$

Exercise 2.4.8 Show that the matrix, with respect to some orthonormal basis $\{\phi_1, \ldots, \phi_n\}$, of the adjoint A^* of a linear operator $A \in \mathcal{L}(\mathcal{H})$ is given by

$$A^*_{ij} = \overline{A}_{ji}.$$

If $\mathcal{H}_1, \mathcal{H}_2$ are inner product spaces with inner products $\langle\cdot|\cdot\rangle_1$ and $\langle\cdot|\cdot\rangle_2$, then we equip their direct sum $\mathcal{H}_1 \oplus \mathcal{H}_2$ with the inner product

$$\langle\phi_1 + \phi_1|\psi_1 + \psi_2\rangle := \langle\phi_1|\psi_1\rangle_1 + \langle\phi_2|\psi_2\rangle_2 \qquad (\phi_1, \psi_1 \in \mathcal{H}_1,\ \phi_2, \psi_2 \in \mathcal{H}_2).$$

Note that $\langle\phi_1|\phi_2\rangle = 0$ for all $\phi_1 \in \mathcal{H}_1$ and $\phi_2 \in \mathcal{H}_2$, i.e., $\mathcal{H}_1$ and $\mathcal{H}_2$ are *orthogonal* subspaces of $\mathcal{H}$. If $\mathcal{H}$ is an inner product space and $\mathcal{F}$ is a linear subspace, then the *orthogonal complement* $\mathcal{F}^\perp$ of $\mathcal{F}$ is defined as

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

Exercise 2.4.9 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 2.4.10 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.

If $\mathcal{H}$ is an inner product space and $\mathcal{F}$ is a linear subspace, then the *orthogonal projection* on $\mathcal{F}$ is defined by

$$P(\phi_1 + \phi_2) := \phi_1 \qquad (\phi_1 \in \mathcal{F},\ \phi_2 \in \mathcal{F}^\perp).$$

One has

$$P^* = P \quad \text{and} \quad P^2 = P,$$

and conversely, every linear operator P with these properties is the orthogonal projection on some subspace $\mathcal{F} \subset \mathcal{H}$ (possibly $\mathcal{F} = \{0\}$ in which case $P = 0$).

Let $\mathcal{H} = \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$ (and therefore, by the definition of the inner product on $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$, the $\mathcal{H}_i$ are orthogonal subspaces of $\mathcal{H}$) and let $P_1, \ldots, P_n$ be the orthogonal projections on $\mathcal{H}_1, \ldots, \mathcal{H}_n$, respectively. Then

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

Conversely, for every set of projections $\{P_1, \ldots, P_n\}$ with these properties there exist orthogonal subspaces $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$ such that $\mathcal{H} = \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$ and P_i projects on $\mathcal{H}_i$.

Exercise 2.4.11 Show that

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

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

Exercise 2.4.12 Show that $\mathcal{L}(\mathcal{H})$, equipped with the usual operator multiplication and adjoint operation, is a positive *-algebra.

By definition, 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 is well-known.

Theorem 2.4.13 (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 $\{\phi_1, \ldots, \phi_n\}$ and complex numbers $\lambda_1, \ldots, \lambda_n$ such that*

$$A = \sum_i \lambda_i |\phi_i\rangle\langle\phi|. \tag{2.1}$$

Note that (2.1) says that the matrix of A with respect to $\{\phi_1, \ldots, \phi_n\}$ is diagonal, i.e., $A_{ij} = \delta_{ij}\lambda_i$.

It is easy to check that an opeator of the form (2.1) commutes with its adjoint. The deeper converse can be proved in an elementary way, but we will postpone the proof till Chapter 4, when we have the powerful tools available from the theory of positive *-algebras that we will prove in the next chapter.

Chapter 3

Algebras

3.1 Algebras

The definition of an algebra (over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) was given in Section 1.3. In this chapter, we are going to investigate the properties of algebras in more detail. Unless stated otherwise, all algebras are assumed to be finite dimensional. While developing the abstract theory of algebras, it is good to keep an explicit example in mind.

Exercise 3.1.1 Let $\mathcal{E}$ be the real linear space

$$\mathcal{E} := \{(a, b, c) : a, b \in \mathbb{C},\ c \in \mathbb{R}\},$$

equipped with the multiplication

$$(a_1, b_1, c_1)(a_2, b_2, c_2) := (a_1 a_2, a_1 b_2 + b_1 c_2, c_1 c_2).$$

Show that $\mathcal{E}$ is an algebra over the real numbers.

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

$$\begin{array}{lll} \text{(a)} & l(\lambda_1 A_1 + \lambda_2 A_2) = \lambda_1 l(A_1) + \lambda_2 l(A_2) & (A_1, A_2 \in \mathcal{A},\ \lambda_1, \lambda_2 \in \mathcal{B}), \\ \text{(b)} & l(AB) = l(A) l(B) & (A, B \in \mathcal{A}). \end{array}$$

If l is a bijection then one can check that also l^{-1} is an algebra homomorphism. In this case we call l an algebra isomorphism and we say that $\mathcal{A}$ and $\mathcal{B}$ are isomorphic (as algebras). By definition, a *sub-algebra* of an algebra $\mathcal{A}$ is a linear subspace $\mathcal{A}' \subset \mathcal{A}$ that is closed under multiplication and that contains an element $1'$ such that $1'A = A = A1'$ for all $A \in \mathcal{A}'$. (In general not $1 = 1'$.) If $\mathcal{A}'$ is a sub-algebra of $\mathcal{A}$, then $\mathcal{A}'$, equipped with the multiplication from $\mathcal{A}$, is itself an algebra.

Exercise 3.1.2 Let $\mathcal{E}$ be the algebra from Exercise 3.1.1 and $\mathcal{E}_r := \{(a,b,c) \in \mathcal{E} : a,b \in \mathbb{R}\}$. Show that $\mathcal{E}_r$ is a subalgebra of $\mathcal{E}$.

By definition, a *left ideal* of an algebra $\mathcal{A}$ is a linear subspace $\mathcal{I}$ such that

$$AB \in \mathcal{I} \quad \forall A \in \mathcal{A},\ B \in \mathcal{I}.$$

A *right ideal* is a linear subspace $\mathcal{I}$ such that $BA \in \mathcal{I}$ $\forall A \in \mathcal{A},\ B \in \mathcal{I}$. An *ideal* is a subspace that is both a left and a righ ideal.

Exercise 3.1.3 Let $\mathcal{E}$ be the algebra in Exercise 3.1.1. Consider the real subspaces $\mathcal{I}_1 := \{(a,b,0) : a,b \in \mathbb{C}\}$, $\mathcal{I}_2 := \{(0,b,c) : b \in \mathbb{C},\ c \in \mathbb{R}\}$, $\mathcal{I}_3 := \{(a,0,0) : a \in \mathbb{C}\}$, $\mathcal{I}_4 := \{(0,b,0) : b \in \mathbb{C}\}$, $\mathcal{I}_5 := \{(0,0,c) : c \in \mathbb{R}\}$, and $\mathcal{I}_6 := \{(a,0,c) : a,c \in \mathbb{R}\}$. Which of these spaces are sub-algebras, left ideals, and/or right ideals of $\mathcal{E}$?

If $\mathcal{I}$ is an ideal of an algebra $\mathcal{A}$, then one can equip the quotient space $\mathcal{A}/\mathcal{I}$ with the structure of an algebra by putting

$$(A+\mathcal{I})(B+\mathcal{I}) := AB+\mathcal{I} \qquad (A,B \in \mathcal{A}).$$

To see that this is a good definition, note that if $A-\tilde{A} \in \mathcal{I}$ and $B-\tilde{B} \in \mathcal{I}$, then $AB-\tilde{A}\tilde{B} = A(B-\tilde{B})+(A-\tilde{A})\tilde{B} \in \mathcal{I}$. If 1 is the identity in $\mathcal{A}$ then obviously $1+\mathcal{I}$ is an identity in $\mathcal{A}/\mathcal{I}$.

Exercise 3.1.4 Let $\mathcal{A},\mathcal{B}$ be algebras and let $l : \mathcal{A} \to \mathcal{B}$ be an algebra homomorphism. Show that the kernel $\mathrm{Ker}(l) := \{A \in \mathcal{A} : l(A) = 0\}$ is an ideal of $\mathcal{A}$ and that $\mathrm{Ran}(l)$ is a subalgebra of $\mathcal{B}$ or $\mathrm{Ran}(l) = \{0\}$. Show that

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

where $\cong$ denotes isomorphism of algebras.

Exercise 3.1.5 Let $\mathcal{A}$ be an algebra and $\mathcal{I} \subset \mathcal{A}$. Show that $\mathcal{I}$ is an ideal if and only if there exists an algebra $\mathcal{B}$ and a surjective algebra homomorphism $l : \mathcal{A} \to \mathcal{B}$ such that $\mathrm{Ker}(l) = \mathcal{I}$. (Hint: take $\mathcal{B} = \mathcal{A}/\mathcal{I}$.)

Exercise 3.1.6 Let $\mathcal{A}$ be an algebra and let $\mathcal{I}_2 \subset \mathcal{I}_1 \subset \mathcal{A}$ be ideals of $\mathcal{A}$. Show that $\mathcal{I}_1/\mathcal{I}_2 = \{B+\mathcal{I}_2 : B \in \mathcal{I}_2\}$ is an ideal of $\mathcal{A}/\mathcal{I}_2 = \{A+\mathcal{I}_2 : A \in \mathcal{A}\}$. Show that

$$(\mathcal{A}/\mathcal{I}_1) \cong (\mathcal{A}/\mathcal{I}_2)/(\mathcal{I}_1/\mathcal{I}_2).$$

(Compare exercise 2.3.5.)

Exercise 3.1.7 Let $\{\mathcal{I}_\gamma\}_{\gamma\in\Gamma}$ be a collection of (left) ideals in an algebra $\mathcal{A}$. Show that $\bigcap_{\gamma\in\Gamma}\mathcal{I}_\gamma$ is a (left) ideal of $\mathcal{A}$.

For arbitrary subsets $\mathcal{B}_1,\ldots,\mathcal{B}_n\subset\mathcal{A}$ we write

$$\begin{aligned}\mathcal{B}_1+\cdots+\mathcal{B}_n &:= \{B_1+\cdots+B_n : B_1\in\mathcal{B}_1,\ldots B_n\in\mathcal{B}_n\},\\ \mathcal{B}_1\cdots\mathcal{B}_n &:= \mathrm{span}\{B_1\cdots B_n : B_1\in\mathcal{B}_1,\ldots B_n\in\mathcal{B}_n\}.\end{aligned}$$

Exercise 3.1.8 Let $\mathcal{A},\mathcal{B}$, and $\mathcal{C}$ be linear subspaces of $\mathcal{A}$. Prove that

$$\begin{array}{rl}\text{(i)} & (\mathcal{A}+\mathcal{B})+\mathcal{C}=\mathcal{A}+\mathcal{B}+\mathcal{C}=\mathcal{A}+(\mathcal{B}+\mathcal{C}),\\ \text{(ii)} & (\mathcal{A}\mathcal{B})\mathcal{C}=\mathcal{A}\mathcal{B}\mathcal{C}=\mathcal{A}(\mathcal{B}\mathcal{C}),\\ \text{(iii)} & \mathcal{A}(\mathcal{B}+\mathcal{C})=\mathcal{A}\mathcal{B}+\mathcal{A}\mathcal{C}\quad\text{and}\quad(\mathcal{A}+\mathcal{B})\mathcal{C}=\mathcal{A}\mathcal{C}+\mathcal{B}\mathcal{C}.\end{array}$$

Exercise 3.1.9 Let $\mathcal{I}_1,\mathcal{I}_2$ be ideals of an algebra $\mathcal{A}$. Show that $\mathcal{I}_1+\mathcal{I}_2$ and $\mathcal{I}_1\mathcal{I}_2$ are ideals, and $\mathcal{I}_1\mathcal{I}_2\subset\mathcal{I}_1\cap\mathcal{I}_2$.

Let $\mathcal{A}_1,\mathcal{A}_2$ be algebras and let $\mathcal{A}_1\oplus\mathcal{A}_2$ be their direct sum (in the sense of linear spaces). We make $\mathcal{A}_1\oplus\mathcal{A}_2$ into an algebra by equipping it with the multiplication

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

(Note that this is a good definition since every element in $\mathcal{A}_1\oplus\mathcal{A}_2$ can in a unique way be written as A_1+A_2 with $A_1\in\mathcal{A}_1$ and $A_2\in\mathcal{A}_2$.) The direct sum of three and more algebras is defined in the obvious way.

Exercise 3.1.10 Check that $\mathcal{A}_1\oplus\mathcal{A}_2$ has an identity. Show that $\mathcal{A}_1$ (viewed as a linear subspace of $\mathcal{A}_1\oplus\mathcal{A}_2$) is a subalgebra of $\mathcal{A}_1\oplus\mathcal{A}_2$ and that the identity in $\mathcal{A}_1$ is not the identity in $\mathcal{A}_1\oplus\mathcal{A}_2$.

Exercise 3.1.11 Check that $\mathcal{A}_2$ is an ideal of $\mathcal{A}_1\oplus\mathcal{A}_2$ and that

$$(\mathcal{A}_1\oplus\mathcal{A}_2)/\mathcal{A}_2\cong\mathcal{A}_1.$$

Exercise 3.1.12 Let $\mathcal{A}_1:=\mathcal{L}(\mathbb{K}^3)$ and $\mathcal{A}=\mathcal{L}(\mathbb{K}^2)$. Show that the direct sum $\mathcal{A}_1\oplus\mathcal{A}_2$ is isomorphic to the space of all matrices of the form

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

with $a_{11},\ldots,b_{22}\in\mathbb{K}$.

Exercise 3.1.13 Generalizing the example in Exercise 3.1.12, let $\mathcal{V}$ be a linear space and let $\mathcal{V}_1, \ldots, \mathcal{V}_n$ be linear subspaces of $\mathcal{V}$ such that $\mathcal{V} = \mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_n$, i.e., every $\phi \in \mathcal{V}$ can in a unique way be written as $\phi = \phi_1 + \cdots + \phi_n$ with $\phi_i \in \mathcal{V}_i$. Let P_i denote the projection on $\mathcal{V}_i$ with respect to the decomposition $\mathcal{V} = \mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_n$, i.e.,

$$P_i(\phi_1 + \cdots + \phi_n) := \phi_i \qquad (i = 1, \ldots, n).$$

Show that $\mathcal{L}(\mathcal{V}_1) \oplus \cdots \oplus \mathcal{L}(\mathcal{V}_n) \cong \{\sum_{i=1}^n P_i A P_i : A \in \mathcal{L}(\mathcal{V})\}$.

Theorem 3.1.14 (Chinese remainder theorem for algebras) *Let $\mathcal{A}$ be an algebra, let $\mathcal{I}_1, \ldots, \mathcal{I}_n$ be ideals in $\mathcal{A}$, and assume that $\mathcal{I}_i + \mathcal{I}_j = \mathcal{A}$ for all $i \neq j$. Then there exists an isomorphism*

$$\mathcal{A}/(\mathcal{I}_1 \cap \cdots \cap \mathcal{I}_n) \cong \mathcal{A}/\mathcal{I}_1 \oplus \cdots \oplus \mathcal{A}/\mathcal{I}_n.$$

Proof We are going to prove that

$$\mathcal{A} = \mathcal{I}_1 + (\mathcal{I}_2 \cap \cdots \cap \mathcal{I}_n). \tag{3.1}$$

For $n = 2$ the statement follows immediately from our assumptions. Since $\mathcal{A}$ has an identity, $\mathcal{A}^2 = \mathcal{A}$ and therefore, using Exercise 3.1.9,

$$\mathcal{A} = \mathcal{A}^2 = (\mathcal{I}_1 + \mathcal{I}_2)(\mathcal{I}_1 + \mathcal{I}_3) = \mathcal{I}_1^2 + \mathcal{I}_1\mathcal{I}_2 + \mathcal{I}_1\mathcal{I}_3 + \mathcal{I}_2\mathcal{I}_3 \subset \mathcal{I}_1 + \mathcal{I}_2 \cap \mathcal{I}_3 \subset \mathcal{A},$$

which proves the statement for $n = 3$. Now assume that (3.1) holds for n ideals. Then, arguing as before,

$$\mathcal{A} = (\mathcal{I}_1 + (\mathcal{I}_2 \cap \cdots \cap \mathcal{I}_n))(\mathcal{I}_1 + \mathcal{I}_{n+1}) \subset \mathcal{A},$$

and therefore (3.1) holds for $n + 1$ ideals. Thus (3.1) follows by induction.

Now define $l : \mathcal{A} \to \mathcal{A}/\mathcal{I}_1 \oplus \cdots \oplus \mathcal{A}/\mathcal{I}_n$ by

$$l(A) := (A + \mathcal{I}_1, \cdots, A + \mathcal{I}_n) \qquad (A \in \mathcal{A}).$$

Then $\mathrm{Ker}(l) = \mathcal{I}_1 \cap \cdots \cap \mathcal{I}_n$. We claim that $\mathrm{Ran}(l) = \mathcal{A}/\mathcal{I}_1 \oplus \cdots \oplus \mathcal{A}/\mathcal{I}_n$, so that

$$\mathcal{A}/(\mathcal{I}_1 \cap \cdots \cap \mathcal{I}_n) = \mathcal{A}/\mathrm{Ker}(l) \cong \mathrm{Ran}(l) = \mathcal{A}/\mathcal{I}_1 \oplus \cdots \oplus \mathcal{A}/\mathcal{I}_n.$$

Let $A_1 + \mathcal{I}_1, \ldots, A_n + \mathcal{I}_n$ be elements of $\mathcal{A}/\mathcal{I}_1, \ldots, \mathcal{A}/\mathcal{I}_n$, respectively. Then we must show that there exists an $A \in \mathcal{A}$ such that

$$\begin{aligned} A + \mathcal{I}_1 &= A_1 + \mathcal{I}_1, \\ &\vdots \\ A + \mathcal{I}_n &= A_n + \mathcal{I}_n. \end{aligned}$$

By (3.1), there exists an $B_1 \in \mathcal{I}_1$ and $C_1 \in \mathcal{I}_2 \cap \cdots \cap \mathcal{I}_n$ such that $B_1 + C_1 = A_1$. Since there is nothing special about $\mathcal{I}_1$, for each $i = 1, \ldots, n$ we can find $B_i \in \mathcal{I}_i$ and $C_i \in \bigcap_{j \neq i} \mathcal{I}_j$ such that $B_i + C_i = A_i$. Thus,

$$C_i + \mathcal{I}_i = A_i + \mathcal{I}_i \quad \text{and} \quad C_i + \mathcal{I}_j = 0 + \mathcal{I}_j \qquad (j \neq i).$$

It is now not hard to see that $A := C_1 + \cdots + C_n$ does the trick. ∎

By definition, an algebra $\mathcal{A}$ is called a *factor algebra* (or a *simple* algebra) if $\mathcal{A}$ has no proper ideals, i.e., its only ideals are $\mathcal{A}$ and $\{0\}$. We note the following simple fact.

Lemma 3.1.15 *Let $\mathcal{A}$ be a factor algebra, $\mathcal{B}$ an algebra, and let $l : \mathcal{A} \to \mathcal{B}$ be an algebra homomorphism. Then either $l = 0$ or $l : \mathcal{A} \to \mathrm{Ran}(l)$ is an algebra isomorphism.*

Proof $\mathrm{Ker}(l)$ is an ideal of $\mathcal{A}$ and therefore either $\mathrm{Ker}(l) = \mathcal{A}$ or $\mathrm{Ker}(l) = \{0\}$. ∎

In the next sections we will investigate factor algebras in more detail.

3.2 Representations of algebras

By definition, a *representation* of an algebra $\mathcal{A}$ is a linear space $\mathcal{V}$ together with a map $(A, \phi) \mapsto A\phi$ from $\mathcal{A} \times \mathcal{V}$ into $\mathcal{V}$ such that

$$\begin{array}{lrcll}
\text{(a)} & A(\lambda_1\phi_1 + \lambda_2\phi_2) &=& \lambda_1 A\phi_1 + \lambda_2 A\phi_2 & (\lambda_1, \lambda_2 \in \mathbb{K},\ A \in \mathcal{A},\ \phi_1, \phi_2 \in \mathcal{V}),\\
\text{(b)} & (\lambda_1 A_1 + \lambda_2 A_2)\phi &=& \lambda_1 A_1\phi + \lambda_2 A_2\phi & (\lambda_1, \lambda_2 \in \mathbb{K},\ A_1, A_2 \in \mathcal{A},\ \phi \in \mathcal{V}),\\
\text{(c)} & (AB)\phi &=& A(B\phi) & (A, B \in \mathcal{A},\ \phi \in \mathcal{V}),\\
\text{(d)} & 1\phi &=& \phi & (\phi \in \mathcal{V}).
\end{array}$$

If $\mathcal{V} \neq \{0\}$ is a representation of $\mathcal{A}$ then the map $l_{\mathcal{V}} : \mathcal{A} \to \mathcal{L}(\mathcal{V})$ given by

$$l_{\mathcal{V}}(A)(\phi) := A\phi$$

is an algebra homomorphism such that $l_{\mathcal{V}}(1) = 1$ (where we equip $\mathcal{L}(\mathcal{V})$ with the structure of an algebra in the obvious way), and vice versa, if $l : \mathcal{A} \to \mathcal{L}(\mathcal{V})$ is an algebra homomorphism with $l(1) = 1$, then $A\phi := l(A)\phi$ makes $\mathcal{V}$ into a representation of $\mathcal{A}$. A representation $\mathcal{V}$ is called *faithful* if $l_{\mathcal{V}} : \mathcal{A} \to \mathrm{Ran}(l_{\mathcal{V}})$ is an algebra isomorphism.

Exercise 3.2.1 Let $\mathcal{E}$ be the algebra from Exercise 3.1.1 and let $\mathcal{V}$ be the real linear space $\mathcal{V} := \{\binom{x}{y} : x \in \mathbb{C},\ y \in \mathbb{R}\}$, equipped with the map from $\mathcal{E} \times \mathcal{V}$ to $\mathcal{V}$ defined by

$$(a,b,c)\binom{x}{y} := \binom{ax+by}{cy}.$$

Show that $\mathcal{V}$ is a faithful representation of $\mathcal{E}$. Note that $\{(1,0),(i,0),(0,1)\}$ is a basis for $\mathcal{V}$, considered as a real linear space. What does the matrix of $l_{\mathcal{V}}(a,b,c)$ look like with respect to this basis?

If $\mathcal{I}$ is a left ideal of and algebra $\mathcal{A}$, then one can equip the quotient space $\mathcal{A}/\mathcal{I}$ with the structure of a representation by putting

$$A(B+\mathcal{I}) := AB + \mathcal{I} \qquad (A, B \in \mathcal{A}).$$

To see that this is a good definition, note that if $B - \tilde{B} \in \mathcal{I}$, then $AB - A\tilde{B} \in \mathcal{I}$ for all $A \in \mathcal{A}$.

Exercise 3.2.2 Let $\mathcal{A}$ be an algebra and $\mathcal{I} \subset \mathcal{A}$. Show that $\mathcal{I}$ is a left ideal if and only if there exists a representation $\mathcal{V}$ of $\mathcal{A}$ and $\phi \in \mathcal{V}$ such that $\mathcal{I} = \{A \in \mathcal{A} : A\phi = 0\}$. (Hint: take $\mathcal{V} := \mathcal{A}/\mathcal{I}$.)

The following lemma is a sort of analogue of Theorem 1.3.7 for algebras. Unlike Theorem 1.3.7, its proof is almost trivial.

Lemma 3.2.3 (Faithful representations) *Every algebra $\mathcal{A}$ has a faithful representation $\mathcal{V}$.*

Proof Every algebra becomes a representation of itself with the map $(A,B) \mapsto AB$. In other words, the map $l_{\mathcal{A}} : \mathcal{A} \to \mathcal{L}(\mathcal{A})$ defined by $l_{\mathcal{A}}(A)B := AB$ is an algebra homomorphism. Since $\mathcal{A}$ contains an identity, $l_{\mathcal{A}}(1)B = B$ for all $B \in \mathcal{A}$, and therefore $l_{\mathcal{A}}(1) = 1$ (the identity in $\mathcal{L}(\mathcal{A})$). Now $\mathrm{Ker}(l_{\mathcal{A}}) = \{0\}$ since $A1 \neq 0$ for all $A \neq 0$, and therefore this representation is faithful. ∎

Using Lemma 3.2.3 we can prove a nice fact about inverses of elements of an algebra.

Lemma 3.2.4 (Inverses) *Let $\mathcal{A}$ be an algebra and $A \in \mathcal{A}$. Then every left inverse of A is a right inverse and every right inverse is a left inverse.*

Proof Assume that $A \in \mathcal{A}$ has a left inverse $B \in \mathcal{A}$. By Lemma 3.2.3, we may view $\mathcal{A}$ as a subalgebra of some $\mathcal{L}(\mathcal{V})$ such that the identity in $\mathcal{A}$ is the identity in $\mathcal{L}(\mathcal{V})$. Then $\dim(\mathcal{V}) = \dim(\mathrm{Ran}(1)) = \dim(\mathrm{Ran}(BA)) \leq \dim(\mathrm{Ran}(A))$ so $A : \mathcal{V} \to \mathcal{V}$ is a bijection. It follows that A has a unique inverse $A^{-1} \in \mathcal{L}(\mathcal{V})$ and every left inverse is also a right inverse (see Exercise 2.1.1). In particular $B = A^{-1}$ and B is also a right inverse. The same argument works when B is a right inverse. ∎

Let $\mathcal{V}, \mathcal{U}$ be two representations of the same algebra $\mathcal{A}$. By definition, a *representation homomorphism* is a linear map $U : \mathcal{V} \to \mathcal{U}$ such that

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

Obviously, the composition of two representation homomorphisms is again a representation homomorphism. If U is a bijection then it is not hard to check that also U^{-1} is a representation homomorphism. In this case we call U an *representation isomorphism* and we say that $\mathcal{V}$ and $\mathcal{U}$ are *equivalent* representations of $\mathcal{A}$.

Exercise 3.2.5 Show that two representations $\mathcal{V}_1, \mathcal{V}_2$ of an algebra $\mathcal{A}$ are equivalent if and only if there exists a linear bijection $U : \mathcal{V}_1 \to \mathcal{V}_2$ such that

$$l_{\mathcal{V}_2}(A) = U l_{\mathcal{V}_1}(A) U^{-1} \qquad (A \in \mathcal{A}).$$

Warning: as the next exercise shows, being equivalent as representations is a stronger statement than being isomorphic as linear spaces!

Exercise 3.2.6 Let $\mathcal{E}_r$ be the algebra from Exercise 3.1.2. Let $\mathcal{V}_1 := \mathbb{R}$, equipped with the map $(a, b, c)x := ax$, and let $\mathcal{V}_2 := \mathbb{R}$, equipped with the map $(a, b, c)x := cx$. Show that $\mathcal{V}_1$ and $\mathcal{V}_2$ are representations of $\mathcal{E}_r$ that are *not* equivalent.

Let $\mathcal{A}$ be an algebra and let $\mathcal{V}$ be a representation of $\mathcal{A}$. By definition, an *invariant subspace* of $\mathcal{V}$ is a linear subspace $\mathcal{U} \subset \mathcal{V}$ such that

$$A\phi \in \mathcal{U} \qquad \forall A \in \mathcal{A},\ \phi \in \mathcal{U}.$$

Obviously, $\mathcal{U}$, together with the map $(A, \phi) \mapsto A\phi$, is itself a representation of $\mathcal{A}$.

Exercise 3.2.7 Let $\mathcal{E}$ be the algebra from Exercise 3.1.1 and let $\mathcal{V}$ be the representation from Exercise 3.2.1. Show that $\mathcal{U} := \{\binom{x}{0} : x \in \mathbb{C}\}$ is an invariant subspace of $\mathcal{V}$. Can you find any other invariant subspaces of $\mathcal{V}$?

Exercise 3.2.8 Let $\mathcal{V}, \mathcal{U}$ be representations of an algebra $\mathcal{A}$ and let $U : \mathcal{V} \to \mathcal{U}$ be a representation homomorphism. Check that $\mathrm{Ker}(U)$ is an invariant subspace of $\mathcal{V}$ and $\mathrm{Ran}(U)$ is an invariant subspace of $\mathcal{U}$.

If $\mathcal{V}$ is a representation of an algebra $\mathcal{A}$ and $\mathcal{U}$ is an invariant subspace of $\mathcal{V}$, then we equip the quotient space $\mathcal{V}/\mathcal{U}$ with the structure of a representation of $\mathcal{A}$ by putting

$$A(\phi+\mathcal{U}) := A\phi+\mathcal{U}.$$

Note that this is a good definition since $(\phi-\psi)\in\mathcal{U} \;\Rightarrow\; (A\phi-A\psi)\in\mathcal{U}$. We call $\mathcal{V}/\mathcal{U}$ the quotient representation of $\mathcal{V}$ and $\mathcal{U}$.

Let $\mathcal{A}$ be an algebra, let $\mathcal{V}_1,\mathcal{V}_2$ be representations of $\mathcal{A}$ and let $\mathcal{V}_1\oplus\mathcal{V}_2$ be the direct sum of $\mathcal{V}_1$ and $\mathcal{V}_2$ (in the sense of linear spaces). We make $\mathcal{V}_1\oplus\mathcal{V}_2$ into a representatation of $\mathcal{A}$ by putting

$$A(\phi_1+\phi_2) := A\phi_1+A\phi_2 \qquad (\phi_1\in\mathcal{V}_1,\ \phi_2\in\mathcal{V}_2).$$

Note that $\mathcal{V}_1$ and $\mathcal{V}_2$ are invariant subspaces of $\mathcal{V}_1\oplus\mathcal{V}_2$. Warning: if $\mathcal{V}$ is a representation of $\mathcal{A}$ and $\mathcal{U}$ is an invariant subspace of $\mathcal{V}$, then it is not always possible to find another invariant subspace $\mathcal{W}$ such that $\mathcal{V}=\mathcal{U}\oplus\mathcal{W}$. (In fact, this is the case in Exercise 3.2.7.)

Let $\mathcal{V}$ be a representation of an algebra $\mathcal{A}$. We say that $\mathcal{U}\subset\mathcal{V}$ is a *proper* invariant subspace if $\mathcal{U}$ is neither $\{0\}$ nor $\mathcal{V}$. A representation $\mathcal{V}$ of an algebra $\mathcal{A}$ is called *irreducible* if $\mathcal{V}\neq\{0\}$ and $\mathcal{V}$ has no proper invariant subspaces. A representation that is not irreducible is called reducible.

Exercise 3.2.9 Show that the representation $\mathcal{U}$ from Exercise 3.2.7 is irreducible.

Note that if $\mathcal{A}$ is an algebra and $\mathcal{V}\neq\{0\}$ is a representation of $\mathcal{A}$, then $\mathcal{V}$ always contains an invariant subspace $\mathcal{U}$ such that $\mathcal{U}$ is an irreducible representation of $\mathcal{A}$. Indeed, $\mathcal{V}$ may contain some proper invariant subspace $\mathcal{V}'$ and $\mathcal{V}'$ may contain a further proper invariant subspace $\mathcal{V}''$, but by finite dimensionality this process stops after a while, and in this way we arrive at an irreducible representation of $\mathcal{A}$.

We will see that there is a close connection between irreducible representations and factor algebras. As a start, we prove:

Proposition 3.2.10 (Irreducible representations of factor algebras) *Every irreducible representation of a factor algebra is faithful. All irreducible representations of a factor algebra are equivalent.*

Proof Let $\mathcal{A}$ be a factor algebra and let $\mathcal{V}$ be an irreducible representation of $\mathcal{A}$. Choose $\phi\in\mathcal{V}$, $\phi\neq 0$. Then $1\phi=\phi\neq 0$ which shows that $\mathrm{Ker}(l_{\mathcal{V}})\neq\mathcal{A}$. Since $\mathrm{Ker}(l_{\mathcal{V}})$ is an ideal of $\mathcal{A}$ and $\mathcal{A}$ is a factor algebra, this implies that $\mathrm{Ker}(l_{\mathcal{V}})=\{0\}$, i.e., $\mathcal{V}$ is faithful.

Let $\mathcal{A}$ be any algebra and consider $\mathcal{A}$ (together with the map $(A,B) \mapsto AB$) as a representation of itself. Then $\mathcal{I} \subset \mathcal{A}$ is an invariant subspace if and only if $\mathcal{I}$ is a left ideal. $\mathcal{I}$ is an irreducible representation if and only if $\mathcal{I} \neq \{0\}$ and $\mathcal{I}$ contains no proper invariant subspaces, which is the same as saying that the only left ideals $\mathcal{I}' \subset \mathcal{I}$ are $\mathcal{I}' = \{0\}$ and $\mathcal{I}' = \mathcal{I}$. In this case we say that $\mathcal{I}$ is a *minimal left ideal*.

Now let $\mathcal{A}$ be a factor algebra and let $\mathcal{V}$ be any irreducible representation of $\mathcal{A}$. Then $\mathcal{A}$ contains some minimal left ideal $\mathcal{I}$, and $\mathcal{I}$ (together with the map $(A,B) \mapsto AB$) is another irreducible representation of $\mathcal{A}$. We will show that $\mathcal{V}$ is equivalent to $\mathcal{I}$. Since $\mathcal{V}$ was arbitrary, this proves that all irreducible representations of $\mathcal{A}$ are equivalent.

We have just seen that $\mathcal{V}$ is faithful. It follows that for every $A \in \mathcal{A}$ with $A \neq 0$ there exists a $\phi \in \mathcal{V}$ such that $A\phi \neq 0$. In particular, we can choose $\phi \in \mathcal{V}$ and $B \in \mathcal{I}$, $B \neq 0$, such that $B\phi \neq 0$. Define $U : \mathcal{I} \to \mathcal{V}$ 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 $\mathcal{V}$ and $\mathrm{Ker}(U)$ is an invariant subspace of $\mathcal{I}$. We have just seen that $B\phi \neq 0$ for some $0 \neq B \in \mathcal{I}$, so $\mathrm{Ran}(U) \neq \{0\}$ and $\mathrm{Ker}(U) \neq \mathcal{I}$. Since $\mathcal{V}$ and $\mathcal{I}$ are irreducible it follows that $\mathrm{Ran}(U) = \mathcal{V}$ and $\mathrm{Ker}(U) = \{0\}$. This shows that U is a representation isomorphism. ∎

3.3 Factor algebras

Let $\mathcal{V}$ be a linear space and let $\mathcal{B}$ be any subset of $\mathcal{L}(\mathcal{V})$. The *commutant* of $\mathcal{B}$ is defined as

$$\mathcal{B}^{\mathrm{C}} := \{A \in \mathcal{L}(\mathcal{V}) : AB = BA \ \forall B \in \mathcal{B}\}.$$

It is easy to see that $\mathcal{B}^{\mathrm{C}}$ is a subalgebra of $\mathcal{L}(\mathcal{V})$.

Lemma 3.3.1 (Commutants and invariant subspaces) *Let $\mathcal{V}$ be a linear space and assume that $\mathcal{B} \subset \mathcal{L}(\mathcal{V})$ has no proper invariant subspaces. Then every nonzero element of $\mathcal{B}^{\mathrm{C}}$ has an inverse in $\mathcal{B}^{\mathrm{C}}$.*

Proof Assume that $A \in \mathcal{B}^{\mathrm{C}}$. We will show that $\mathrm{Ker}(A)$ is an invariant subspace under $\mathcal{B}$. Indeed, if $\phi \in \mathrm{Ker}(A)$ and $B \in \mathcal{B}$ then $A(B\phi) = BA\phi = B0 = 0 \ \Rightarrow B\phi \in \mathrm{Ker}(A)$. Since $\mathcal{B}$ has no proper invariant subspaces, either $\mathrm{Ker}(A) = \mathcal{V}$, in which case $A = 0$, or $\mathrm{Ker}(A) = \{0\}$, in which case A is invertible. To see that

$A^{-1} \in \mathcal{B}^{\mathrm{C}}$, note that $AB = BA \Rightarrow A^{-1}ABA^{-1} = A^{-1}BAA^{-1} \Rightarrow BA^{-1} = A^{-1}B$ for all $B \in \mathcal{B}$. ∎

If $\mathcal{V}$ is a linear space over the complex numbers, then a considerably stronger statement holds.

Lemma 3.3.2 (Commutants and invariant subspaces, complex case) *Let $\mathcal{V}$ be a linear space over $\mathbb{C}$ and assume that $\mathcal{B} \subset \mathcal{L}(\mathcal{V})$ has no proper invariant subspaces. Then $\mathcal{B}^{\mathrm{C}} = \{\lambda 1 : \lambda \in \mathbb{C}\}$.*

Proof Assume that $A \in \mathcal{B}^{\mathrm{C}}$. By Lemma 3.3.1, every nonzero element of $\mathcal{B}^{\mathrm{C}}$ has an inverse. Let $\mathcal{C}$ be smallest subspace of $\mathcal{L}(\mathcal{V})$ that contains A and I and that is closed under taking of inverses.

We claim that $\mathcal{C}$ is abelian. We may construct $\mathcal{C}$ in a finite number of steps: we start with the space $\mathcal{C}_0 := \mathrm{span}\{1, A\}$. Now if $\mathcal{C}_0$ contains two elements whose product is not in $\mathcal{C}_0$ then we add this product, take the span, and call this $\mathcal{C}_1$. If $\mathcal{C}_1$ contains some nonzero element whose inverse is not in $\mathcal{C}_1$, then we add this inverse, take the span, and call this $\mathcal{C}_2$. By finite dimensionality, this process stops somewhere, and in this way we arrive at $\mathcal{C}$. (It is not hard to see that one always arrives at the same set $\mathcal{C}$, but this is not relevant for our proof.) Obviously, $\mathcal{C}_0$ is abelian. Now if $\mathcal{C}_n$ is abelian and $A, B \in \mathcal{C}_n$, then every element in $\mathcal{C}_n$ commutes with AB, so if we add AB to $\mathcal{C}_n$, take the span, and call this $\mathcal{C}_{n+1}$, then $\mathcal{C}_{n+1}$ is abelian. Likewise, if $\mathcal{C}_n$ is abelian and $0 \neq A \in \mathcal{C}_n$, then every element in $\mathcal{C}_n$ commutes with A^{-1} (check this!), so if we add A^{-1} to $\mathcal{C}_n$, take the span, and call this $\mathcal{C}_{n+1}$, then $\mathcal{C}_{n+1}$ is abelian. Therefore, by induction, $\mathcal{C}$ is abelian.

Now $\mathcal{C}$ is a finite dimensional complex abelian algebra in which every nonzero element has an inverse. In particular, $\mathcal{C}$ is a field. Moreover, $\mathcal{C}$ contains $\{\lambda 1 : \lambda \in \mathbb{C}\} \cong \mathbb{C}$ as a subfield. Assume that $C \in \mathcal{C}$. Since $\mathcal{C}$ is finite dimensional, there exists an n such that $C^{n+1} = a_0 1 + a_1 C + \cdots + a_n C^n$ with $a_0, \ldots, a_n \in \mathbb{C}$. This shows that $\mathcal{C}$ is an algebraic extension of the field $\mathbb{C}$. Since $\mathbb{C}$ is algebraically closed, by Theorem 2.1.3, this implies that $\mathcal{C} = \{\lambda 1 : \lambda \in \mathbb{C}\}$. In particular, $A = \lambda 1$ for some $\lambda \in \mathbb{C}$. ∎

Exercise 3.3.3 Let $\mathcal{B} \subset \mathcal{L}(\mathbb{R}^2)$ be the space of all linear operators with a matrix of the form

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

Show that $\mathcal{B}$ is an algebra with no proper invariant subspaces. Show that each nonzero element in $\mathcal{B}$ has an inverse. Give a formula for this inverse. Does $\mathcal{B}$ remind you of some algebra you know?

The next theorem is a special case of Jacobson's density theorem, see [Lan71, Theorem XVII.3.1] or [Hun74, Theorem IX.1.12]. We call the commutant of the commutant of an algebra $\mathcal{B}$ the *bicommutant* of $\mathcal{B}$.

Theorem 3.3.4 (Bicommutants and invariant subspaces) *Let $\mathcal{V} \neq \{0\}$ be a linear space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and assume that $\mathcal{B} \subset \mathcal{L}(\mathcal{V})$ is an algebra with no proper invariant subspaces. Then $(\mathcal{B}^{\mathrm{C}})^{\mathrm{C}} = \mathcal{B}$.*

Proof Let $\mathcal{V}^{(n)} := \mathcal{V} \oplus \cdots \oplus \mathcal{V}$ be the direct sum of n copies of $\mathcal{V}$. For any $A \in \mathcal{L}(\mathcal{V})$ define a linear operator $A^{(n)}$ on $\mathcal{V}^{(n)}$ by

$$A^{(n)}(\psi_1, \ldots, \psi_n) := (A\psi_1, \ldots, A\psi_n).$$

Set $\mathcal{B}^{(n)} = \{B^{(n)} : B \in \mathcal{B}\}$. We claim that $A \in (\mathcal{B}^{\mathrm{C}})^{\mathrm{C}}$ implies that $A^{(n)} \in ((\mathcal{B}^{(n)})^{\mathrm{C}})^{\mathrm{C}}$. Indeed, with respect to an obvious basis for $\mathcal{V}^{(n)}$, each $B^{(n)} \in \mathcal{B}^{(n)}$ has the block-diagonal form (for example for $n = 3$):

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

Now any $C \in \mathcal{L}(\mathcal{V}^{(n)})$ 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 $\mathcal{V}$ into $\mathcal{V}$. We see that

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

and therefore C commutes with each $B^{(n)}$ in $\mathcal{B}^{(n)}$ if and only if $C_{ij} \in \mathcal{B}^{\mathrm{C}}$ for each i, j. In particular, if $C \in (\mathcal{B}^{(n)})^{\mathrm{C}}$ and $A \in (\mathcal{B}^C)^C$, then A commutes with each C_{ij} and therefore $A^{(n)}$ commutes with C, i.e., $A^{(n)} \in ((\mathcal{B}^{(n)})^{\mathrm{C}})^{\mathrm{C}}$.

Now fix $A \in (\mathcal{B}^C)^C$, let $\phi_1, \ldots, \phi_n$ be a basis of $\mathcal{V}$ and put $\phi := (\phi_1, \ldots, \phi_n) \in \mathcal{V}^n$. Put $\mathcal{B}^{(n)}\phi := \{B^{(n)}\phi : B \in \mathcal{B}\}$. We need to show that there exists a $B \in \mathcal{B}$ such that $A\phi_i = B\phi_i$ for each $i = 1, \ldots, n$, or equivalently, that $A^{(n)}\phi \in \mathcal{B}^{(n)}\phi$.

Let us write $\mathcal{V}^{(n)} := \mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_n$ where the $\mathcal{V}_i$ are copies of $\mathcal{V}$. Choose inductively $\mathcal{V}_{i_1}, \mathcal{V}_{i_2}, \ldots$ such that $(\mathcal{B}^{(n)}\phi \oplus \mathcal{V}_{i_1} \oplus \cdots \oplus \mathcal{V}_{i_k}) \cap \mathcal{V}_{i_{k+1}} = \{0\}$, until for some k this is no longer possible. We claim that then $\mathcal{B}^{(n)}\phi \oplus \mathcal{V}_{i_1} \oplus \cdots \oplus \mathcal{V}_{i_k} = \mathcal{V}^{(n)}$. It suffices to show that any $\mathcal{V}_i$ that we have not had yet is contained in $\mathcal{B}^{(n)}\phi \oplus \mathcal{V}_{i_1} \oplus \cdots \oplus \mathcal{V}_{i_k}$. Indeed, $(\mathcal{B}^{(n)}\phi \oplus \mathcal{V}_{i_1} \oplus \cdots \oplus \mathcal{V}_{i_k}) \cap \mathcal{V}_i$ is a $\mathcal{B}$-invariant subspace of $\mathcal{V}_i = \mathcal{V}$ and therefore, since $\mathcal{B}$ has no proper invariant subspaces, either $\mathcal{V}_i$ or $\{0\}$; the latter contradicts our assumption that the process of adding $\mathcal{V}_{i_1}, \mathcal{V}_{i_2}, \ldots$ was complete.

Set $\mathcal{U} := \mathcal{V}_{i_1} \oplus \cdots \oplus \mathcal{V}_{i_k}$. Then any $\psi \in \mathcal{V}^{(n)}$ can be written in a unique way as $\psi = \psi' + \psi''$ with $\psi' \in \mathcal{B}^{(n)}\phi$ and $\psi'' \in \mathcal{U}$. Let $P\psi := \psi'$ denote the projection on $\mathcal{B}^{(n)}\phi$ with respect to this decomposition. We claim that $P \in (\mathcal{B}^{(n)})^{\mathrm{C}}$. Indeed, if $B^{(n)} \in \mathcal{B}^{(n)}$ and $\psi = \psi' + \psi''$ as before then $B^{(n)}\psi = B^{(n)}\psi' + B^{(n)}\psi''$ where $B^{(n)}\psi' \in \mathcal{B}^{(n)}\phi$ and $B^{(n)}\psi'' \in \mathcal{U}$. It follows that $B^{(n)}P\psi = B^{(n)}\psi' = PB^{(n)}\psi$, i.e., P commutes with $B^{(n)}$. Since $A^{(n)} \in ((\mathcal{B}^{(n)})^{\mathrm{C}})^{\mathrm{C}}$, P commutes with $A^{(n)}$, and we conclude that $A^{(n)}\phi = A^{(n)}P\phi = PA^{(n)}\phi \in \mathcal{B}^{(n)}\phi$. ∎

Lemma 3.3.5 (Algebras with no invariant subspaces) *Let $\mathcal{V}$ be a linear space over $\mathbb{C}$ and let $\mathcal{A} \subset \mathcal{L}(\mathcal{V})$ be an algebra with no invariant subspaces. Then $\mathcal{A} = \mathcal{L}(\mathcal{V})$.*

Proof Immediate from Lemma 3.3.2 and Theorem 3.3.4. ∎

Theorem 3.3.6 (Characterisation of factor algebras) *Let $\mathcal{A}$ be an algebra over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Consider the statements:*

(i) *$\mathcal{A}$ is a factor algebra,*
(ii) *$\mathcal{A}$ has a faithful irreducible representation,*
(iii) *$\mathcal{A} \cong \mathcal{L}(\mathcal{V})$ for some linear space $\mathcal{V}$ over $\mathbb{K}$.*

If $\mathbb{K} = \mathbb{R}$, then (i) $\Leftrightarrow$ (ii) $\Leftarrow$ (iii). *If $\mathbb{K} = \mathbb{C}$, then* (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii).

Proof (iii) $\Rightarrow$ (i): We must show that $\mathcal{L}(\mathcal{V})$ has no proper ideals. Assume that $\mathcal{I} \subset \mathcal{L}(\mathcal{V})$ is an ideal and $\mathcal{I} \neq \{0\}$. Choose $A \in \mathcal{I}$, $A \neq 0$. Choose a basis $\{\phi_1, \ldots, \phi_n\}$ for $\mathcal{V}$ and let $P(i,j) \in \mathcal{L}(\mathcal{V})$ denote the linear operator with matrix elements $P(i,j)_{ij} = 1$ and $P(i,j)_{kl} = 0$ if $(i,j) \neq (k,l)$. Since $A \neq 0$, $A_{ij} \neq 0$ for some i,j. Since $\mathcal{I}$ is an ideal, $P(k,i)AP(j,l) \in \mathcal{I}$ for all k,l. The matrix of $P(k,i)AP(j,l)$ is

$$(P(k,i)AP(j,l))_{m,p} = \sum_{q,r} P(k,i)_{m,q}A_{q,r}P(j,l)_{r,p} = \begin{cases} A_{ij} & \text{if } (m,p) = (k,l), \\ 0 & \text{otherwise.} \end{cases}$$

We see that $P(k,i)AP(j,l) = A_{ij}P(k,l) \in \mathcal{I}$ for all k,l. Since $A_{ij} \neq 0$ and since the operators $P(k,l)$ span $\mathcal{L}(\mathcal{V})$, we see that $\mathcal{I} = \mathcal{L}(\mathcal{V})$. Thus, $\mathcal{L}(\mathcal{V})$ has no proper ideals.

(i) ⇒ (ii): We have already seen that every algebra has an irreducible representation. By Proposition 3.2.10, every irreducible representation of a factor algebra is faithful.

(ii) ⇒ (iii) in case $\mathbb{K} = \mathbb{C}$: Let $\mathcal{V}$ be a faithful irreducible representation. Then $\mathrm{Ran}(l_{\mathcal{V}})$ has no proper invariant subspaces, and $\mathcal{A}$ is isomorphic to $\mathrm{Ran}(l_{\mathcal{V}})$. By Lemma 3.3.5, $\mathcal{A} \cong \mathrm{Ran}(l_{\mathcal{V}}) = \mathcal{L}(\mathcal{V})$.

(ii) ⇒ (i): If $\mathbb{K} = \mathbb{C}$ this follows from the fact that (ii) ⇒ (iii). Since we will not really need the case $\mathbb{K} = \mathbb{R}$ in what follows, we only sketch the proof. By assumption $\mathcal{B} := \mathrm{Ran}(l_{\mathcal{V}})$ has no proper invariant subspaces, and $\mathcal{A}$ is isomorphic to $\mathcal{B}$. By Lemma 3.3.1, every nonzero element in $\mathcal{B}^{\mathrm{C}}$ has an inverse. This means that $\mathcal{B}^{\mathrm{C}}$ is a division ring (see Section 2.1). As a result, we can view $\mathcal{V}$ as a finite dimensional linear space over $\mathcal{B}^{\mathrm{C}}$. In particular, we can find a '$\mathcal{B}^{\mathrm{C}}$-basis of $\mathcal{V}$', i.e., vectors $\{\phi_1, \ldots, \phi_k\}$ such that every $\psi \in \mathcal{V}$ can in a unique way be written as

$$\psi = A_1\phi_1 + \cdots + A_k\phi_k,$$

with $A_1, \ldots, A_k \in \mathcal{B}^{\mathrm{C}}$. (To see this, read the proof on page 21 of the fact that every finite dimensional linear space over a division ring has a basis.)

By Theorem 3.3.4, $\mathcal{B} = (\mathcal{B}^{\mathrm{C}})^{\mathrm{C}}$. Here $(\mathcal{B}^{\mathrm{C}})^{\mathrm{C}}$ consists of all maps $B : \mathcal{V} \to \mathcal{V}$ that satisfy $B(A_1\phi_1 + A_2\phi_2) = A_1B\phi_1 + A_2B\phi_2$, i.e., all '$\mathcal{B}^{\mathrm{C}}$-linear' maps. Choosing a $\mathcal{B}^{\mathrm{C}}$-basis $\{\phi_1, \ldots, \phi_k\}$ for $\mathcal{V}$ one can show that there is a one-to-one correspondence between $\mathcal{B}^{\mathrm{C}}$-linear maps and $k \times k$ matrices with coefficients from $\mathcal{B}^{\mathrm{C}}$. Then, much in the same way as we proved the step (iii) ⇒ (i), one can prove that the space of such matrices has no proper ideals, and hence $\mathcal{B} = (\mathcal{B}^{\mathrm{C}})^{\mathrm{C}}$ is a factor algebra. ∎

Proposition 3.3.7 (Isomorphisms of factor algebras) *Let $\mathcal{V}_1, \mathcal{V}_2$ be nonzero linear spaces and let $l : \mathcal{L}(\mathcal{V}_1) \to \mathcal{L}(\mathcal{V}_2)$ be an algebra isomorphism. Then there exists a linear bijection $U : \mathcal{V}_1 \to \mathcal{V}_2$ such that*

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

Proof $\mathcal{V}_1$, together with the map $(A, \phi) \mapsto A\phi$, is a representation of $\mathcal{L}(\mathcal{V}_1)$. It is easy to see that $\mathcal{L}(\mathcal{V}_1)$ has no proper invariant subspaces and therefore this representation is irreducible. Now $\mathcal{V}_2$ with the map $(A, \phi) \mapsto l(A)\phi$ is another representation of $\mathcal{L}(\mathcal{V}_1)$. Since l is surjective, also this representation is irreducible. By Theorem 3.3.6, $\mathcal{L}(\mathcal{V}_1)$ is a factor algebra. By Proposition 3.2.10, all irreducible representations of a factor algebra are equivalent. Applied to the two representations of $\mathcal{L}(\mathcal{V}_1)$ mentioned above, this means that there is a linear bijection $U : \mathcal{V}_1 \to \mathcal{V}_2$ such that

$$UA\phi = l(A)U\phi \qquad (A \in \mathcal{A},\ \phi \in \mathcal{V}_1),$$

as required. ∎

3.4 Semisimple algebras

In this section, we set out to find *all* nonequivalent irreducible representations of an algebra.

By definition, a (left) ideal $\mathcal{I}$ of an algebra $\mathcal{A}$ is called *maximal* if $\mathcal{I} \neq \mathcal{A}$ and every (left) ideal that contains $\mathcal{I}$ is either $\mathcal{I}$ or $\mathcal{A}$.

Lemma 3.4.1 (Maximal left ideals) *Let $\mathcal{A}$ be an algebra and $\mathcal{I} \subset \mathcal{A}$. Then $\mathcal{I}$ is a maximal left ideal if and only if the following equivalent conditions hold.*

(i) *$\mathcal{A}/\mathcal{I}$ is an irreducible representation of $\mathcal{A}$.*

(ii) *There exists an irreducible representation $\mathcal{V}$ of $\mathcal{A}$ and $0 \neq \phi \in \mathcal{V}$ such that $\mathcal{I} = \{A \in \mathcal{A} : A\phi = 0\}$.*

Here we equip $\mathcal{A}/\mathcal{I}$ with the structure of a representation of $\mathcal{A}$ as on page 36.

Proof (i): Assume that $\mathcal{I}'$ is a left ideal such that $\mathcal{I} \subset \mathcal{I}' \subset \mathcal{A}$. Then $\mathcal{I}'/\mathcal{I} = \{B + \mathcal{I} : B \in \mathcal{I}'\}$ is an invariant subspace of $\mathcal{A}/\mathcal{I} = \{A + \mathcal{I} : A \in \mathcal{A}\}$. Thus, $\mathcal{A}/\mathcal{I}$ is irreducible $\Leftrightarrow \mathcal{I}'/\mathcal{I} = \mathcal{I}$ or $\mathcal{I}'/\mathcal{I} = \{0\} \Leftrightarrow \mathcal{I}' = \mathcal{I}$ or $\mathcal{I}' = \mathcal{A} \Leftrightarrow \mathcal{I}$ maximal.

(ii): If $\mathcal{I}$ is a maximal left ideal, then put $\mathcal{V} := \mathcal{A}/\mathcal{I}$ and $\phi := 1 + \mathcal{I}$. Then by (i), $\mathcal{V}$ is an irreducible representation of $\mathcal{A}$, and $\mathcal{I} = \{A \in \mathcal{A} : A\phi = 0\}$. Conversely, assume that $\mathcal{V}$ is an irreducible representation of $\mathcal{A}$ and $\mathcal{I} = \{A \in \mathcal{A} : A\phi = 0\}$ for some $0 \neq \phi \in \mathcal{V}$. Consider the linear map $l_\phi : \mathcal{A} \to \mathcal{V}$ given by $l_\phi(A) := A\phi$. Then $\mathcal{I} = \mathrm{Ker}(l_\phi)$ and since $\mathcal{V}$ is irreducible, $\mathrm{Ran}(l_\phi) = \mathcal{V}$. It follows that $\mathcal{A}/\mathcal{I} = \mathcal{A}/\mathrm{Ker}(l_\phi) \cong \mathrm{Ran}(l_\phi) = \mathcal{V}$, so $\mathcal{A}/\mathcal{I}$ is irreducible. By (i), this implies that $\mathcal{I}$ is maximal. ∎

Lemma 3.4.2 (Maximal ideals) *Let $\mathcal{A}$ be an algebra and $\mathcal{I} \subset \mathcal{A}$. Then $\mathcal{I}$ is a maximal ideal if and only if the following equivalent conditions hold.*

(i) *$\mathcal{A}/\mathcal{I}$ is a factor algebra.*

(ii) *There exists a factor algebra $\mathcal{B}$ and a surjective algebra isomorphism $l : \mathcal{A} \to \mathcal{B}$ such that $\mathcal{I} = \mathrm{Ker}(l)$.*

(iii) *There exists an irreducible representation $\mathcal{V}$ of $\mathcal{A}$ such that $\mathcal{I} = \mathrm{Ker}(l_\mathcal{V})$ $= \{A \in \mathcal{A} : A\phi = 0 \ \forall \phi \in \mathcal{V}\}$.*

Proof (i): Imagine that $\mathcal{I}'$ is an ideal such that $\mathcal{I} \subset \mathcal{I}' \subset \mathcal{A}$. Then, by Exercise 3.1.6, $\mathcal{I}'/\mathcal{I}$ is an ideal of $\mathcal{A}/\mathcal{I}$. Thus, $\mathcal{A}/\mathcal{I}$ is a factor $\Leftrightarrow$ the only ideals of $\mathcal{A}/\mathcal{I}$ are $\mathcal{A}/\mathcal{I}$ and $\{0\}$, $\Leftrightarrow \mathcal{I}'/\mathcal{I} = \mathcal{A}/\mathcal{I}$ or $\mathcal{I}'/\mathcal{I} = \{0\} \Leftrightarrow \mathcal{I}' = \mathcal{A}$ or $\mathcal{I}' = \mathcal{I}$.

(ii): In Exercise 3.1.5 we have seen that $\mathcal{I}$ is an ideal if and only if there exists an algebra $\mathcal{B}$ and a surjective algebra homomorphism $l : \mathcal{A} \to \mathcal{B}$ such that $\mathrm{Ker}(l) = \mathcal{I}$. Now $\mathcal{B} = \mathrm{Ran}(l) \cong \mathcal{A}/\mathrm{Ker}(l) = \mathcal{A}/\mathcal{I}$ so this is just a translation of (i).

(iii): If $\mathcal{I}$ is maximal then by (ii) there exists a factor algebra $\mathcal{B}$ and a surjective algebra homomorphism $l : \mathcal{A} \to \mathcal{B}$ such that $\mathrm{Ker}(l) = \mathcal{I}$. By Theorem 3.3.6 $\mathcal{B}$ has a faithful irreducible representation $\mathcal{V}$. Equip $\mathcal{V}$ with the structure of an $\mathcal{A}$-representation by putting $A\phi := l(A)\phi$ $(A \in \mathcal{A},\ \phi \in \mathcal{V})$. Then $\mathcal{V}$ is an irreducible representation of $\mathcal{A}$ and $\mathcal{I} = \{A \in \mathcal{A} : A\phi = 0\ \forall\phi \in \mathcal{V}\}$. Conversely, if $\mathcal{V}$ is an irreducible representation of $\mathcal{A}$ and $\mathcal{I} = \{A \in \mathcal{A} : A\phi = 0\ \forall\phi \in \mathcal{V}\}$, then $l_{\mathcal{V}} : \mathcal{A} \to \mathrm{Ran}(l_{\mathcal{V}})$ is a surjective algebra homomorphism. Obviously, $\mathrm{Ran}(l_{\mathcal{V}})$ has a faithful irreducible representation, so by Theorem 3.3.6, $\mathrm{Ran}(l_{\mathcal{V}})$ is a factor algebra. By (ii), this implies that $\mathcal{I}$ is a maximal ideal. ∎

Lemma 3.4.3 (Nonequivalent irreducible representations) *Let $\mathcal{A}$ be an algebra and let $\mathcal{V},\mathcal{U}$ be irreducible representations of $\mathcal{A}$. Then $\mathcal{V}$ and $\mathcal{U}$ are equivalent if and only if* $\mathrm{Ker}(l_{\mathcal{V}}) = \mathrm{Ker}(l_{\mathcal{U}})$.

Proof Assume that $\mathrm{Ker}(l_{\mathcal{V}}) = \mathrm{Ker}(l_{\mathcal{U}}) =: \mathcal{I}$. We can view $\mathcal{V}$ and $\mathcal{U}$ as representations of the algebra $\mathcal{A}/\mathcal{I}$, by putting $(A + \mathcal{I})\phi := A\phi$ $(A \in \mathcal{A})$. Since $\mathcal{V},\mathcal{U}$ are obviously faithful irreducible representations of $\mathcal{A}/\mathcal{I}$, by Theorem 3.3.6 the latter is a factor algebra. Now it follows from Proposition 3.2.10 that $\mathcal{V}$ and $\mathcal{U}$ are equivalent.

Conversely, if $\mathcal{V}$ and $\mathcal{U}$ are equivalent, then by definition of equivalence there is a linear bijection $U : \mathcal{V} \to \mathcal{U}$ such that $UA\phi = AU\phi$ $(A \in \mathcal{A},\ \phi \in \mathcal{V})$. Now $A \in \mathrm{Ker}(l_{\mathcal{V}}) \ \Rightarrow\ A\phi = 0\ \forall\phi \in \mathcal{V} \ \Rightarrow\ AU\phi = UA\phi = U0 = 0\ \forall\phi \in \mathcal{V} \ \Rightarrow\ A\psi = 0\ \forall\psi \in \mathcal{U} \ \Rightarrow\ A \in \mathrm{Ker}(l_{\mathcal{U}})$. This shows that $\mathrm{Ker}(l_{\mathcal{V}}) \subset \mathrm{Ker}(l_{\mathcal{U}})$. By symmetry also $\mathrm{Ker}(l_{\mathcal{V}}) \supset \mathrm{Ker}(l_{\mathcal{U}})$. ∎

By definition, the *radical*[1] of an algebra $\mathcal{A}$ is the ideal $\mathcal{J} = \mathcal{J}(\mathcal{A})$ given by

$$\begin{aligned}\mathcal{J} :=& \bigcap\{\mathcal{I} : \mathcal{I} \text{ maximal ideal of } \mathcal{A}\}\\ =& \bigcap\{\mathcal{I} : \mathcal{I} \text{ maximal left ideal of } \mathcal{A}\}.\end{aligned}$$

[1]There exist many radicals, the historically oldest one being Wedderburn's radical and the best-known one being Jacobson's radical. Since many of the definitions coincide for finite dimensional algebras it is not obvious what is the best name for the radical we consider.

Exercise 3.4.4 Show that the two definitions of $\mathcal{J}$ coincide.

Exercise 3.4.5 Let $\mathcal{E}$ be the algebra from Exercise 3.1.1. Determine its radical $\mathcal{J}(\mathcal{E})$.

Theorem 3.4.6 (The radical) *Every algebra $\mathcal{A}$ has finitely many maximal ideals $\mathcal{I}_1, \ldots, \mathcal{I}_n$. Therefore, the radical of $\mathcal{A}$ is given by $\mathcal{J}(\mathcal{A}) = \mathcal{I}_1 \cap \cdots \cap \mathcal{I}_n$. For any algebra $\mathcal{A}$,*

$$\mathcal{A}/\mathcal{J}(\mathcal{A}) \cong \mathcal{A}/\mathcal{I}_1 \oplus \cdots \oplus \mathcal{A}/\mathcal{I}_n. \tag{3.2}$$

One has $\mathcal{J}(\mathcal{A}) = \{0\}$ if and only if

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

where $\mathcal{A}_1, \ldots, \mathcal{A}_n$ are factor algebras.

Proof Assume that $\mathcal{I}_1, \ldots, \mathcal{I}_k$ are maximal ideals of $\mathcal{A}$, all different. For each $i \neq j$, $\mathcal{I}_i + \mathcal{I}_j$ is an ideal that contains $\mathcal{I}_i$ but that is not equal to $\mathcal{I}_i$; therefore, by maximality, $\mathcal{I}_i + \mathcal{I}_j = \mathcal{A}$ for all $i \neq j$. This means that the Chinese remainder theorem is applicable, i.e.,

$$\mathcal{A}/(\mathcal{I}_1 \cap \cdots \cap \mathcal{I}_k) \cong \mathcal{A}/\mathcal{I}_1 \oplus \cdots \oplus \mathcal{A}/\mathcal{I}_k.$$

By Lemma 3.4.2, each $\mathcal{A}/\mathcal{I}_i$ is a factor algebra. In particular, $\dim(\mathcal{A}/\mathcal{I}_i) \geq 1$. Now $\dim(\mathcal{A}/(\mathcal{I}_1 \cap \cdots \cap \mathcal{I}_n)) \leq \dim(\mathcal{A})$, which is finite, and therefore $\mathcal{A}$ can have at most finitely many distinct maximal ideals $\mathcal{I}_1, \ldots, \mathcal{I}_n$ and (3.2) holds.

If $\mathcal{J}(\mathcal{A}) = \mathcal{I}_1 \cap \cdots \cap \mathcal{I}_n = \{0\}$ then

$$\mathcal{A} \cong \mathcal{A}/\{0\} \cong \mathcal{A}/\mathcal{I}_1 \oplus \cdots \oplus \mathcal{A}/\mathcal{I}_n,$$

which is a sum of factor algebras. Conversely, if $\mathcal{A} \cong \mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$ is a direct sum of factor algebras, then define algebra homomorphisms $l_i : (\mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n) \to \mathcal{A}_i$ by the projections

$$l_i(A_1, \ldots, A_n) := A_i \qquad (A_1 \in \mathcal{A}_1, \ldots, A_n \in \mathcal{A}_n).$$

By Lemma 3.4.2, $\mathrm{Ker}(l_1), \ldots, \mathrm{Ker}(l_n)$ are maximal ideals of $\mathcal{A}$. Now $\mathrm{Ker}(l_i) = \{(A_1, \ldots, A_n) : A_i = 0\}$ and therefore $\mathcal{J}(\mathcal{A}) \subset \bigcap_{i=1}^n \mathrm{Ker}(l_i) = \{0\}$. ∎

By definition, an algebra is called *semisimple* if $\mathcal{J}(\mathcal{A}) = \{0\}$ or equivalently, if $\mathcal{A} \cong \mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$, where $\mathcal{A}_1, \ldots, \mathcal{A}_n$ are factor algebras. Note that each $\mathcal{A}_i$ is an ideal of $\mathcal{A}$. Up to the order of the factors, the decomposition of a semisimple algebra into factors is unique:

Lemma 3.4.7 (Decomposition of simple algebras) *Let $\mathcal{A}$ be an algebra and let $\mathcal{A}_1, \ldots, \mathcal{A}_n$ be subalgebras of $\mathcal{A}$ such that each $\mathcal{A}_i$ is a factor algebra and $\mathcal{A} = \mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$. Then each $\mathcal{A}_i$ is an ideal. Apart from $\mathcal{A}_1, \ldots, \mathcal{A}_n$, there are no other ideals that are factor algebras.*

Proof Assume that $\mathcal{I} \subset \mathcal{A}$ is an ideal of $\mathcal{A}$ and a factor algebra. Then, for each i, $\mathcal{I}\mathcal{A}_i$ is an ideal. Now, using Exercise 3.1.8, $\{0\} \neq \mathcal{I} = \mathcal{I}\mathcal{A} = \mathcal{I}(\mathcal{A}_1 + \cdots + \mathcal{A}_n) = \mathcal{I}\mathcal{A}_1 + \cdots + \mathcal{I}\mathcal{A}_n$ so $\mathcal{I}\mathcal{A}_i \neq \{0\}$ for some i. Since $\mathcal{I}\mathcal{A}_i$ is an ideal of $\mathcal{A}_i$ and $\mathcal{A}_i$ is a factor it follows that $\mathcal{I}\mathcal{A}_i = \mathcal{A}_i$. Likewise $\mathcal{I}\mathcal{A}_i$ is an ideal of $\mathcal{I}$ and $\mathcal{I}$ is a factor so $\mathcal{I}\mathcal{A}_i = \mathcal{I}$. Thus, $\mathcal{I} = \mathcal{A}_i$ for some i, as claimed. ∎

We conclude this section by proving the following important property of the radical.

Proposition 3.4.8 (Nilpotent radical) *The radical $\mathcal{J} = \mathcal{J}(\mathcal{A})$ of an algebra $\mathcal{A}$ is nilpotent, i.e., $\mathcal{J}^n = \{0\}$ for some $n \geq 1$.*

We start with a preparatory lemma.

Lemma 3.4.9 (Maximal left ideals and invertible elements) *Let $\mathcal{A}$ be an algebra and $A \in \mathcal{A}$. Then A is contained in some maximal left ideal if and only if A is not invertible.*

Proof Note that $\mathcal{A}A := \{BA : B \in \mathcal{A}\}$ is the smallest left ideal containing A. If $\mathcal{A}A = \mathcal{A}$ then obviously A is not contained in some maximal left ideal. On the other hand, if $\mathcal{A}A \neq \mathcal{A}$ then $\mathcal{A}A$ may be contained some larger left ideal $\mathcal{I}$ that is not yet $\mathcal{A}$ and $\mathcal{I}$ may be contained in some even larger $\mathcal{I}'$ but by finite dimensionality this process stops somewhere so A is contained in some maximal left ideal.

Thus, we must show that $\mathcal{A}A = \mathcal{A}$ if and only if A is invertible. If $\mathcal{A}A = \mathcal{A}$ then in particular $1 \in \mathcal{A}A$ so $BA = 1$ for some $B \in \mathcal{A}$. By Lemma 3.2.4 this implies that A is invertible. Conversely, if A is invertible then $\mathcal{A}A \supset \mathcal{A}AA^{-1} = \mathcal{A}$. ∎

Proof of Proposition 3.4.8 $\mathcal{J} \supset \mathcal{J}^2 \supset \mathcal{J}^3$ so by finite dimensionality, $\mathcal{J}^k = \mathcal{J}^{k+1}$ for some k. Let us say that a finite set $\{B_1, \ldots, B_n\}$ is a left generator of $\mathcal{J}^k$ if every $B \in \mathcal{J}^k$ can be written as

$$B = A_1 B_1 + \cdots + A_n B_n \qquad \text{with } A_1, \ldots, A_n \in \mathcal{A}.$$

By finite dimensionality, such a left generator exists. (For example, we can take for $\{B_1, \ldots, B_n\}$ a basis for $\mathcal{J}^k$ and all A_i of the form $\lambda_i 1$.) Now $B_1 \in \mathcal{J}^k = \mathcal{J}\mathcal{J}^k$

so B_1 can be written as $B_1 = \sum_i \lambda_i C_i D_i$ with $C_i \in \mathcal{J}$ and $D_i \in \mathcal{J}^k$. Each D_i can be written as $D_i = A_{i,1}B_1 + \cdots + A_{i,n}B_n$ so

$$B_1 = \tilde{A}_1 B_1 + \cdots + \tilde{A}_n B_n$$

where $\tilde{A}_k := \sum_i \lambda_i C_i A_{k,i} \in \mathcal{J}$ since $C_i \in \mathcal{J}$ and $\mathcal{J}$ is a radical. It follows that

$$(1 - \tilde{A}_1)B_1 = \tilde{A}_2 B_2 + \cdots + \tilde{A}_n B_n.$$

We claim that $(1-\tilde{A}_1)$ is invertible in $\mathcal{A}$. For assume that $(1-\tilde{A}_1)$ is not invertible. Then by Lemma 3.4.9, $(1 - \tilde{A}_1)$ is contained in some maximal left ideal $\mathcal{I}$. But $\mathcal{J}$ is the intersection of all maximal left ideals so $\tilde{A}_1 \in \mathcal{J} \subset \mathcal{I}$. It follows that $1 \in \mathcal{I}$. By Lemma 3.4.9, this implies that 1 is not invertible, which is absurd.

So $(1 - \tilde{A}_1)$ has some inverse $(1 - \tilde{A}_1)^{-1} \in \mathcal{A}$, and we can write

$$B_1 = (1 - \tilde{A}_1)^{-1}\tilde{A}_2 B_2 + \cdots + (1 - \tilde{A}_1)^{-1}\tilde{A}_n B_n.$$

It is not hard to see that this implies that $\{B_2, \ldots, B_n\}$ is a left generator of $\mathcal{J}^k$. We can continue this process till we see that $\mathcal{J}^k$ is generated by an empty set, and therefore $\mathcal{J}^k = \{0\}$. ∎

3.5 Fully reducible representations

Lemma 3.5.1 (Fully reducible representations) *Let $\mathcal{V}$ be a representation of an algebra $\mathcal{A}$. Then the following conditions are equivalent.*

(i) *There exist irreducible invariant subspaces $\mathcal{U}_1, \ldots, \mathcal{U}_n$ such that $\mathcal{V} = \mathcal{U}_1 \oplus \cdots \oplus \mathcal{U}_n$*

(ii) $\mathcal{V} = \mathrm{span}\{\mathcal{U} \subset \mathcal{V} : \mathcal{U}$ *irreducible invariant subspace*$\}$.

(iii) *For every invariant subspace $\mathcal{U}$ there exists an invariant subspace $\mathcal{W}$ such that $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$.*

(iv) *For every invariant subspace $\mathcal{V}'$ of $\mathcal{V}$ and for every invariant subspace $\mathcal{U}'$ of $\mathcal{V}'$ there exists an invariant subspace $\mathcal{W}'$ of $\mathcal{V}'$ such that $\mathcal{V}' = \mathcal{U}' \oplus \mathcal{W}'$.*

A representation $\mathcal{V}$ satisfying (i)–(iv) is called *fully reducible.*

Proof (i)⇒(ii): Obvious.

(ii)⇒(iii): Let $\mathcal{U}$ be any invariant subspace of $\mathcal{V}$. Then either $\mathcal{U} = \mathcal{V}$ or there exists an irreducible invariant subspace $\mathcal{U}_1$ of $\mathcal{V}$ that is not contained in $\mathcal{U}$. Since $\mathcal{U} \cap \mathcal{U}_1$ is an invariant subspace of $\mathcal{U}_1$, by irreducibility, we have $\mathcal{U} \cap \mathcal{U}_1 = \{0\}$. Now either $\mathcal{U} \oplus \mathcal{U}_1 = \mathcal{V}$ or there exists an irreducible invariant subspace $\mathcal{U}_1$ of $\mathcal{V}$ that is not contained in $\mathcal{U} \oplus \mathcal{U}_1$. By finite dimensionality, this process stops somewhere, so we can write $\mathcal{V} = \mathcal{U} \oplus \mathcal{U}_1 \oplus \cdots \oplus \mathcal{U}_n$ where the $\mathcal{U}_1, \ldots, \mathcal{U}_n$ are irreducible invariant subspaces. In particular, setting $\mathcal{W} := \mathcal{U}_1 \oplus \cdots \oplus \mathcal{U}_n$ yields (iii).

(iii)⇒(iv): By our assumption we can write $\mathcal{V} = \mathcal{U}' \oplus \mathcal{W}$ for some invariant subspace $\mathcal{W}$ of $\mathcal{V}$. Put $\mathcal{W}' := \mathcal{W} \cap \mathcal{V}'$. Then $\mathcal{W}'$ is an invariant subspace of $\mathcal{V}'$. We claim that $\mathcal{V}' = \mathcal{U}' \oplus \mathcal{W}'$. It is clear that $\mathcal{U}' \cap \mathcal{W}' = \{0\}$ so we must show that every $\phi \in \mathcal{V}'$ can be written as $\phi = \phi_{\rm u} + \phi_{\rm w}$ with $\phi_{\rm u} \in \mathcal{U}'$ and $\phi_{\rm w} \in \mathcal{W}'$. Since $\mathcal{V} = \mathcal{U}' \oplus \mathcal{W}$, we know that every $\phi \in \mathcal{V}'$ can be written as $\phi = \phi_{\rm u} + \phi_{\rm w}$ with $\phi_{\rm u} \in \mathcal{U}'$ and $\phi_{\rm w} \in \mathcal{W}$. Now $\phi_{\rm w} = \phi - \phi_{\rm u} \in \mathcal{V}'$ so $\phi_{\rm w} \in \mathcal{W} \cap \mathcal{V}' = \mathcal{W}'$.

(iv)⇒(i): If $\mathcal{V}$ is irreducible or $\{0\}$ the statement is obvious. If $\mathcal{V}$ is not irreducible then $\mathcal{V}$ contains a proper invariant subspace $\mathcal{U}$. By our assumption, $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$ for some invariant subspace $\mathcal{W}$. Continuing in this fashion we can split up $\mathcal{U}$ and $\mathcal{W}$ into further invariant subspaces. By finite dimensionality, this process stops somewhere, and we arrive at a decomposition of $\mathcal{V}$ into irreducible representations.∎

Lemma 3.5.2 (Preservation of full reducibility) *If $\mathcal{V}$ is a fully reducible representation of an algebra $\mathcal{A}$ and $\mathcal{U}$ is an invariant subspace of $\mathcal{V}$, then $\mathcal{U}$ and $\mathcal{V}/\mathcal{U}$ are fully reducible.*

Proof It is obvious from the equivalence of conditions (iii) and (iv) of Lemma 3.5.1 that an invariant subspace $\mathcal{U}$ of a fully reducible representation is itself fully reducible. Furthermore, by Lemma 3.5.1 (iii), we can write $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$ for some invariant subspace $\mathcal{W}$. Since $\mathcal{V}/\mathcal{U} \cong \mathcal{W}$ this also shows that $\mathcal{V}/\mathcal{U}$ is fully reducible.∎

We will prove:

Theorem 3.5.3 (Representations of semisimple algebras) *Every representation of a semisimple algebra is fully reducible.*

We start with a lemma.

Lemma 3.5.4 (Canonical representation of semisimple algebras) *Let $\mathcal{A}$ be a semisimple algebra. Then $\mathcal{A}$, considered as a representation of itself, is fully reducible.*

Proof We first treat the case that $\mathcal{A}$ is an algebra over $\mathbb{C}$. As we have seen in the proof of Proposition 3.2.10, the irreducible representations of any algebra $\mathcal{A}$, considered as a representation of itself, are the minimal left ideals of $\mathcal{A}$. If $\mathcal{A}$ is a factor algebra, then by Theorem 3.3.6 $\mathcal{A} \cong \mathcal{L}(\mathcal{V})$ for some linear space $\mathcal{V}$. For example, if $\mathcal{V}$ is three-dimensional, then it is not hard to see that the spaces of all matrices of the form

$$\begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & a_{12} & 0 \\ 0 & a_{22} & 0 \\ 0 & a_{32} & 0 \end{pmatrix}, \quad \text{or} \quad \begin{pmatrix} 0 & 0 & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}$$

are minimal left ideals. It follows that all minimal left ideals (that is, all irreducible invariant subspaces) span $\mathcal{A}$, so by Lemma 3.5.1 (ii), $\mathcal{A}$ is fully reducible. If $\mathcal{A} \cong \mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$ is semisimple then the same argument applied to each factor $\mathcal{A}_i$ shows that the minimal left ideals of $\mathcal{A}$ span $\mathcal{A}$, so again, $\mathcal{A}$, considered as a representation of itself, is fully reducible.

If $\mathcal{A}$ is an algebra over $\mathbb{R}$ then the argument is the same, except that we must use matrices with coefficients in some division ring, as in the proof of the implication (ii)$\Rightarrow$(i) in Theorem 3.3.6. ∎

Proof of Theorem 3.5.3 Let $\mathcal{A}$ be a semisimple algebra, let $\mathcal{V}$ be a representation of $\mathcal{A}$, and let $\{\phi_1, \ldots, \phi_n\}$ be a basis of $\mathcal{V}$. Consider $\mathcal{A}$ as a representation of itself and consider the direct sum (of representations)

$$\underbrace{\mathcal{A} \oplus \cdots \oplus \mathcal{A}}_{n \text{ times}}.$$

Define $U : \mathcal{A} \oplus \cdots \oplus \mathcal{A} \to \mathcal{V}$ by $U(A_1, \ldots, A_n) := A_1\phi_1 + \cdots + A_n\phi_n$. We claim that U is a representation homomorphism. Indeed: $UA(A_1, \ldots, A_n) = U(AA_1, \ldots, AA_n) = AA_1\phi_1 + \cdots + AA_n\phi_n = AU(A_1, \ldots, A_n)$. Moreover, since $U(\lambda_1 1, \ldots, \lambda_n 1) = \lambda_1\phi_1 + \cdots + \lambda_n\phi_n$, we see that $\mathrm{Ran}(U) = \mathcal{V}$. It follows that $\mathcal{V} = \mathrm{Ran}(U) \cong (\mathcal{A} \oplus \cdots \oplus \mathcal{A})/\mathrm{Ker}(U)$, which is fully reducible by Lemmas 3.5.2 and 3.5.4. ∎

The next theorem summarizes our knowledge about the structure of semisimple algebras.

Theorem 3.5.5 (Structure theorem for semisimple algebras) *Let $\mathcal{A}$ be a semisimple algebra. Then $\mathcal{A}$ has finitely many nonequivalent irreducible representations $\mathcal{V}_1, \ldots, \mathcal{V}_n$, and the map*

$$A \mapsto (l_{\mathcal{V}_1}(A), \ldots, l_{\mathcal{V}_n}(A))$$

defines an algebra isomorphism

$$\mathcal{A} \cong \mathrm{Ran}(l_{\mathcal{V}_1}) \oplus \cdots \oplus \mathrm{Ran}(l_{\mathcal{V}_n}).$$

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

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

with $m_i \geq 0$ $(i = 1, \ldots, n)$. $\mathcal{V}$ *is faithful if and only if* $m_i \geq 1$ *for all* $i = 1, \ldots, n$. *If* $\mathcal{A}$ *is an algebra over the complex numbers then moreover* $\mathrm{Ran}(l_{\mathcal{V}_i}) = \mathcal{L}(\mathcal{V}_i)$ $(i = 1, \ldots, n)$.

Proof By Theorem 3.4.6, $\mathcal{A}$ has finitely many maximal ideals $\mathcal{I}_1, \ldots, \mathcal{I}_n$ and $\mathcal{A} \cong (\mathcal{A}/\mathcal{I}_1) \oplus \cdots \oplus (\mathcal{A}/\mathcal{I}_n)$, where $\mathcal{A}/\mathcal{I}_1, \ldots, \mathcal{A}/\mathcal{I}_n$ are factor algebras by Lemma 3.4.2. Also by lemma 3.4.2, for each $i = 1, \ldots, n$ there exists an irreducible representation $\mathcal{V}_i$ of $\mathcal{A}$ such that $\mathrm{Ker}(l_{\mathcal{V}_i}) = \mathcal{I}_i$, and by Lemma 3.4.3 these are all nonequivalent irreducible representations of $\mathcal{A}$. Consider the algebra homomorphism $l : \mathcal{A} \to \mathcal{L}(\mathcal{V}_1) \oplus \cdots \mathcal{L}(\mathcal{V}_n)$ given by

$$l(A) := (l_{\mathcal{V}_1}(A), \ldots, l_{\mathcal{V}_n}(A)) \qquad (A \in \mathcal{A}).$$

Since $\mathcal{A}/\mathcal{I}_i = \mathcal{A}/\mathrm{Ker}(l_{\mathcal{V}_i}) \cong \mathrm{Ran}(l_{\mathcal{V}_i})$, Theorem 3.4.6 says that

$$\mathrm{Ran}(l) = \mathrm{Ran}(l_{\mathcal{V}_1}) \oplus \cdots \oplus \mathrm{Ran}(l_{\mathcal{V}_n}).$$

Moreover,

$$\mathrm{Ker}(l) = \mathrm{Ker}(l_{\mathcal{V}_1}) \cap \cdots \cap \mathrm{Ker}(l_{\mathcal{V}_n}) = \mathcal{J}(\mathcal{A}) = \{0\}$$

since $\mathcal{A}$ is semisimple. This shows that l is an algebra isomorphism. By Theorem 3.5.3, every representation $\mathcal{V}$ of $\mathcal{A}$ can be written as a direct sum of the irreducible representations $\mathcal{V}_1, \ldots, \mathcal{V}_n$, with $\mathcal{V}_i$ occuring m_i times $(i = 1, \ldots, n)$. Then $\mathrm{Ker}(l_{\mathcal{V}}) = \bigcap\{\mathrm{Ker}(l_{\mathcal{V}_i}) : m_i \geq 1\}$ so $\mathcal{V}$ is faithful if $m_i \geq 1$ for each $i = 1, \ldots, n$. Conversely, since $A \mapsto (l_{\mathcal{V}_1}(A), \ldots, l_{\mathcal{V}_n}(A))$ is an isomorphism, for each i we can find some $A \in \mathcal{A}$ with $l_{\mathcal{V}_i}(A) \neq 0$ and $l_{\mathcal{V}_j}(A) = 0$ for all $j \neq i$. This shows that $\mathcal{V}$ is not faithful if $m_i = 0$ for some i. Finally, $\mathrm{Ran}(l_{\mathcal{V}_i})$ is a subalgebra of $\mathcal{L}(\mathcal{V}_i)$ that has no invariant subspaces. Therefore, by Lemma 3.3.5, in the special case that $\mathcal{A}$ is an algebra over the complex numbers $\mathrm{Ran}(l_{\mathcal{V}_i}) = \mathcal{L}(\mathcal{V}_i)$. ∎

The numbers m_i appearing in Theorem 3.5.5 are called the *multiplicity* of the irreducible representation $\mathcal{V}_i$. Although the way of decomposing $\mathcal{V}$ into irreducible representations is not unique, the multiplicities do not depend on this decomposition. This follows from the next exercise.

Exercise 3.5.6 Write

$$\mathcal{V} \cong \mathcal{W}_1 \oplus \cdots \oplus \mathcal{W}_n \quad \text{with} \quad \mathcal{W}_i = \underbrace{\mathcal{V}_i \oplus \cdots \oplus \mathcal{V}_i}_{m_i \text{ times}}.$$

Let $0 \neq \phi \in \mathcal{V}$ and consider the invariant subspace $\mathcal{A}\phi := \{A\phi : A \in \mathcal{A}\}$. Show that $\mathcal{A}\phi$ is irreducible if and only if $\phi \in \mathcal{W}_i$ for some i. Show that $\mathcal{W}_i = \{\phi \in \mathcal{V} : \mathcal{A}\phi \text{ equivalent to } \mathcal{V}_i\} \cup \{0\}$. Show that $m_i = \dim(\mathcal{W}_i)/\dim(\mathcal{V}_i)$.

Exercise 3.5.7 Let $\mathcal{E}$ be the algebra from Exercise 3.1.1 and let $\tilde{\mathcal{E}} := \{(a, 0, c) : a \in \mathbb{C},\ c \in \mathbb{R}\}$. Show that $\tilde{\mathcal{E}}$ is simple. Let $\mathcal{V}$ be the space $\mathcal{V} := \{\binom{x}{y} : x, y \in \mathbb{R}\}$ equipped with the map

$$(a, 0, c)\binom{x}{y} := \binom{cx}{cy}.$$

Then $\mathcal{V}$ is a representation of $\tilde{\mathcal{E}}$. Determine all nonequivalent representations of $\tilde{\mathcal{E}}$ and determine their multiplicities in $\mathcal{V}$.

3.6 Positive *-algebras

We have finally collected enough knowledge about the structure of algebras to prove the theorems about positive *-algebras from Section 1. Our proofs loosely follow [GHJ89, Appendix II.a]. The definition of a positive *-algebra was given in Section 1.3. In this and the next section, all linear spaces are finite dimensional linear spaces over the complex numbers.

Lemma 3.6.1 (Positive *-algebras are semisimple) *Every positive *-algebra is semisimple.*

Proof Let $\mathcal{A}$ be any *-algebra and let $\mathcal{J} = \mathcal{J}(\mathcal{A})$ be its radical. We claim that $\mathcal{J}^* = \mathcal{J}$. If $\mathcal{I}$ is a left ideal of $\mathcal{A}$, then $\mathcal{I}^*$ is a right ideal. To see this, just note that $A \in \mathcal{A}$, $B \in \mathcal{I}^* \Rightarrow BA = (A^*B^*)^* = C^*$ for some $C \in \mathcal{I}$. In the same way we see that if $\mathcal{I}$ is a right ideal then $\mathcal{I}^*$ is a left ideal, and therefore, if $\mathcal{I}$ is a (two-sided) ideal, then $\mathcal{I}^*$ is also an ideal. It is easy to see that if $\mathcal{I}$ is a maximal ideal, then $\mathcal{I}^*$ is also a maximal ideal. (But not necessarily $\mathcal{I} = \mathcal{I}^*$!) Now

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

as claimed.

Now let $\mathcal{A}$ be a positive *-algebra. By Theorem 3.4.6, we must show that $\mathcal{J} = \{0\}$. Imagine that $0 \neq A \in \mathcal{J}$. Then by what we have just proved $A^* \in \mathcal{J}$ and therefore

$A^*A \in \mathcal{J}$. By positivity, $A^*A \neq 0$, $(A^*A)^*(A^*A) = (A^*A)^2 \neq 0$, and by induction, $(A^*A)^{2^n} \neq 0$ for all $n \geq 1$. On the other hand, by Proposition 3.4.8, $(A^*A)^m = 0$ for some $m \geq 1$. We arrive at a contradiction. ∎

If $\mathcal{A}_1, \ldots, \mathcal{A}_n$ are *-algebras and $\mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$ is their direct sum, in the sense of algebras, then we make $\mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$ into a *-algebra by equipping it with the adjoint operation

$$(A_1, \ldots, A_n)^* := (A_1^*, \ldots, A_n^*) \qquad (A_1 \in \mathcal{A}_1, \ldots, A_n \in \mathcal{A}_n).$$

If the A_i are positive *-algebras then it is easy to see that $\mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$ is a positive *-algebra.

Proposition 3.6.2 (Decomposition of positive *-algebras) *Every positive *-algebra $\mathcal{A}$ is isomorphic to a direct sum*

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

where $\mathcal{A}_1, \ldots, \mathcal{A}_n$ are positive *-algebras that are factor algebras.

Proof Let $\mathcal{A}$ be a positive *-algebra, with adjoint operation denoted by $A \mapsto A^*$. By Lemma 3.6.1, $\mathcal{A}$ is isomorphic *as an algebra* to a direct sum of the form $\mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$, where each $\mathcal{A}_i$ is a factor algebra. We must show that each $\mathcal{A}_i$ can be equipped with a positive adjoint operation $A_i \mapsto A_i^*$ such that this is a *-algebra isomorphism.

By Theorem 3.3.6, each $\mathcal{A}_i$ is of the form $\mathcal{A}_i \cong \mathcal{L}(\mathcal{V}_i)$ for some complex linear space $\mathcal{V}_i$. Choose some arbitrary inner product on $\mathcal{V}_i$ and let $A_i^\dagger$ denote the usual adjoint of an operator $A_i \in \mathcal{L}(\mathcal{V}_i)$ with respect to this inner product. Using Exercise 2.4.8 it is easy to see that this is a positive adjoint operation on $\mathcal{L}(\mathcal{V}_i)$. Thus, on each $\mathcal{A}_i$ we can choose some positive adjoint operation $A_i \mapsto A_i^\dagger$. Now

$$(A_1, \ldots, A_n)^\dagger := (A_1^\dagger, \ldots, A_n^\dagger) \qquad (A_i \in \mathcal{A}_i,\ i = 1, \ldots, n)$$

defines a positive adjoint operation $A \mapsto A^\dagger$ on $\mathcal{A}$. Consider the map $l : \mathcal{A} \to \mathcal{A}$ given by

$$l(A) := (A^\dagger)^*.$$

We check that l is an algebra isomorfism:

$$l(\lambda_1 A_1 + \lambda_1 A_2) = (\overline{\lambda}_1 A_1^\dagger + \overline{\lambda}_1 A_2^\dagger)^* = \lambda_1 (A_1^\dagger)^* + \lambda_2 (A_2^\dagger)^*,$$

$$l(AB) = ((AB)^\dagger)^* = (B^\dagger A^\dagger)^* = (A^\dagger)^*(B^\dagger)^*,$$

and

$$\mathrm{Ker}(l) = \{A : (A^\dagger)^* = 0\} = \{A : A^\dagger = 0\} = \{0\}.$$

Lemma 3.4.7 shows that the decomposition of a semisimple algebra into factors is unique up to the order of the factors. As a result, any algebra isomorfism $l : \mathcal{A} \to \mathcal{A}$ must be of the form:

$$l(A_1,\ldots,A_n) = (l_{\pi(1),1}(A_{\pi(1)}),\ldots,l_{\pi(n),n}(A_{\pi(n)})) \qquad (A_i \in \mathcal{L}(\mathcal{H}_i),\ i = 1,\ldots,n),$$

where $\pi : \{1,\ldots,n\} \to \{1,\ldots,n\}$ is a permutation and

$$l_{\pi(i),i} : \mathcal{H}_{\pi(i)} \to \mathcal{H}_i \qquad (i = 1,\ldots,n)$$

are algebra isomorphisms. We claim that for the algebra isomorphism $A \mapsto l(A) := (A^\dagger)^*$, the permutation π is actually the identity. For imagine that $\pi(i) \neq i$ for some i. Let A be any nonzero element of $\mathcal{A}_{\pi(i)}$. Then $A^*A = l(A^\dagger)A = l_{\pi(i),i}(A^\dagger)A = 0$ since $l_{\pi(i),i}(A^\dagger) \in \mathcal{A}_i$ and $A \in \mathcal{A}_{\pi(i)}$, so we arrive at a contradiction with the fact that $A \mapsto A^*$ is positive.

Thus, l is of the form

$$l(A_1,\ldots,A_n) = (l_1(A_1),\ldots,l_n(A_1)) \qquad (A_i \in \mathcal{A}_i,\ i = 1,\ldots,n),$$

where $l_i : \mathcal{A}_i \to \mathcal{A}_i$ are algebra isomorhisms. Define positive adjoint operations $A_i \mapsto A_i^*$ on $\mathcal{A}_i$ $(i = 1,\ldots,n)$ by $A_i^* := l_i(A_i^\dagger)$. Since $A^* = l(A^\dagger)$ we then have

$$(A_1,\ldots,A_n)^* = (l_1(A_1^\dagger),\ldots,l_n(A_n^\dagger)) = (A_1^*,\ldots,A_n^*),$$

i.e., $\mathcal{A} \cong \mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_n$ in the sense of a direct sum of positive *-algebras. ∎

3.7 Representations of positive *-algebras

Let $\mathcal{A}$ be a positive *-algebra. By definition, a **-representation* of $\mathcal{A}$ is a complex inner product space $\mathcal{H}$ together with a map $(A,\phi) \mapsto A\phi$ from $\mathcal{A} \times \mathcal{H}$ into $\mathcal{H}$ such that

$$\begin{array}{lll}
\text{(a)} & A(\lambda_1\phi_1 + \lambda_2\phi_2) = \lambda_1 A\phi_1 + \lambda_2 A\phi_2 & (\lambda_1,\lambda_2 \in \mathbb{C},\ A \in \mathcal{A},\ \phi_1,\phi_2 \in \mathcal{H}),\\
\text{(b)} & (\lambda_1 A_1 + \lambda_2 A_2)\phi = \lambda_1 A_1\phi + \lambda_2 A_2\phi & (\lambda_1,\lambda_2 \in \mathbb{C},\ A_1,A_2 \in \mathcal{A},\ \phi \in \mathcal{H}),\\
\text{(c)} & (AB)\phi = A(B\phi) & (A,B \in \mathcal{A},\ \phi \in \mathcal{H}),\\
\text{(d)} & 1\phi = \phi & (\phi \in \mathcal{H}),\\
\text{(e)} & \langle\phi|A\psi\rangle = \langle A^*\phi|\psi\rangle & (A \in \mathcal{A},\ \phi,\psi \in \mathcal{H}).
\end{array}$$

I.e., a *-representation is a representation such that additionaly (e) holds. Equivalently, a *-representation is a representation such that additionally the algebra homomorphism $l_{\mathcal{H}} : \mathcal{A} \to \mathcal{L}(\mathcal{H})$ is a *-algebra homomorphism. We call $\mathcal{H}$ *faithful* if $l_{\mathcal{H}} : \mathcal{A} \to \mathrm{Ran}(l_{\mathcal{H}})$ is a *-algebra isomorphism. Invariant subspaces and irreducible *-representations are defined in the same way as for representations. We say that two *-representations $\mathcal{H}_1, \mathcal{H}_2$ of the same positive *-algebra $\mathcal{A}$ are *equivalent* if there exists a unitary map $U : \mathcal{H}_1 \to \mathcal{H}_2$ such that

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

Exercise 3.7.1 Show that two *-representations $\mathcal{H}_1, \mathcal{H}_2$ of a positive *-algebra $\mathcal{A}$ are equivalent if and only if there exists a unitary map $U : \mathcal{H}_1 \to \mathcal{H}_2$ such that

$$l_{\mathcal{H}_2}(A) = U l_{\mathcal{H}_1}(A) U^{-1} \qquad (A \in \mathcal{A}).$$

Lemma 3.7.2 (Positive adjoint operations on factor algebras) *Let $\mathcal{V}$ be a complex linear space and let $A \mapsto A^*$ be a positive adjoint operation on $\mathcal{L}(\mathcal{V})$. Then there exists an inner product $\langle\cdot|\cdot\rangle$ on $\mathcal{V}$ such that $A \mapsto A^*$ is the usual adjoint with respect to this inner product.*

Proof Choose any inner product $\langle\cdot|\cdot\rangle'$ on $\mathcal{V}$ and let $A \mapsto A^\dagger$ denote the adjoint with respect to this inner product. Define $l : \mathcal{L}(\mathcal{V}) \to \mathcal{L}(\mathcal{V})$ by $l(A) := (A^*)^\dagger$. As in the proof of Proposition 3.6.2 we see that l is an algebra isomorphism, and therefore, by Proposition 3.3.7, there exists a linear bijection $U : \mathcal{V} \to \mathcal{V}$ such that $l(A) = UAU^{-1}$. Define a new inner product $\langle\cdot|\cdot\rangle$ on $\mathcal{V}$ by

$$\langle\phi|\psi\rangle := \langle U\phi|U\psi\rangle' \qquad (\phi, \psi \in \mathcal{V}).$$

Then

$$\begin{aligned}\langle\phi|A\psi\rangle &= \langle U\phi|UA\psi\rangle' = \langle U\phi|UAU^{-1}U\psi\rangle' \\ &= \langle U\phi|(A^*)^\dagger U\psi\rangle' = \langle A^*U\phi|U\psi\rangle' = \langle A^*\phi|\psi\rangle,\end{aligned}$$

which proves that A^* is the adjoint of A with respect to the inner product $\langle\phi|\psi\rangle$. ∎

Lemma 3.7.3 (Isomorphisms of factor positive *-algebras) *Let $\mathcal{H}_1, \mathcal{H}_2$ be inner product spaces and let $l : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2)$ be a *-algebra isomorphism. Then there exists a unitary map $U : \mathcal{H}_1 \to \mathcal{H}_2$ such that*

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

Moreover, if two irreducible *-representations of a positive *- algebra are equivalent as representations, then they are equivalent as *-representations.

Proof By Proposition 3.6.2 there exists a linear bijection $U : \mathcal{H}_1 \to \mathcal{H}_2$ such that $l(A) = UAU^{-1}$. We must show that U is unitary. Since $l(A^*) = l(A)^*$ we have $UA^*U^{-1} = (UAU^{-1})^* = (U^{-1})^*A^*U^*$ for each $A \in \mathcal{L}(\mathcal{H}_1)$. In particular, if $A = |\psi\rangle\langle\phi|$ with $\phi, \psi \in \mathcal{H}_1$, then

$$|U^*\phi\rangle\langle(U^{-1})^*\psi| = U|\phi\rangle\langle\psi|U^{-1} = (U^{-1})^*|\phi\rangle\langle\psi|U^* = |U^{-1}\phi\rangle\langle U\psi|.$$

Since this holds for all $\phi, \psi \in \mathcal{H}_1$, we must have $U^{-1} = U^*$, i.e., U is unitary.

Now if two irreducible *-representations $\mathcal{H}_1, \mathcal{H}_2$ of a positive *-algebra are equivalent as representations, then there exists a linear bijection $U : \mathcal{H}_1 \to \mathcal{H}_2$ such that $l_{\mathcal{H}_2}(A) = Ul_{\mathcal{H}_1}(A)U^{-1}$. Since $\mathcal{H}_1$ and $\mathcal{H}_2$ are *-representations, $l_{\mathcal{H}_2}(A)^* = l_{\mathcal{H}_2}(A^*) = Ul_{\mathcal{H}_1}(A^*)U^{-1} = Ul_{\mathcal{H}_1}(A)^*U^{-1}$. By Lemma 3.3.5, $\mathrm{Ran}(l_{\mathcal{H}_1}) = \mathcal{L}(\mathcal{H}_1)$ and $\mathrm{Ran}(l_{\mathcal{H}_2}) = \mathcal{L}(\mathcal{H}_2)$, so $A \mapsto UAU^{-1}$ is a *-algebra isomorphism from $\mathcal{L}(\mathcal{H}_1)$ to $\mathcal{L}(\mathcal{H}_2)$. By what we have just proved, U is unitary, so by Exercise 3.7.1 $\mathcal{H}_1$ and $\mathcal{H}_2$ are equivalent as *-representations. ∎

Let $\mathcal{A}$ be a positive *-algebra and let $\mathcal{H}_1, \mathcal{H}_2$ be *-representations of $\mathcal{A}$. Let $\mathcal{H}_1 \oplus \mathcal{H}_2$ be the direct sum of $\mathcal{H}_1$ and $\mathcal{H}_2$ in the sense of inner product spaces. Then we make $\mathcal{H}_1 \oplus \mathcal{H}_2$ into a *-representation by putting

$$A(\phi_1 + \phi_2) := A\phi_1 + A\phi_2 \qquad (\phi_1 \in \mathcal{H}_1,\ \phi_2 \in \mathcal{H}_2).$$

Note that $\mathcal{H}_1$ and $\mathcal{H}_2$ are invariant subspaces of $\mathcal{H}_1 \oplus \mathcal{H}_2$. The next lemma has no analogue for representations!

Lemma 3.7.4 (Decomposition of *-representations) *Let $\mathcal{A}$ be a positive *-algebra and let $\mathcal{H}$ be a *-representation of $\mathcal{A}$. If $\mathcal{F} \subset \mathcal{H}$ is an invariant subspace of $\mathcal{H}$, then $\mathcal{F}^\perp$ is an invariant subspace of $\mathcal{H}$ and $\mathcal{H} \cong \mathcal{F} \oplus \mathcal{F}^\perp$.*

Proof For any $A \in \mathcal{A}$ and $\phi \in \mathcal{F}^\perp$ one has $\langle\phi|\psi\rangle = 0\ \forall\psi \in \mathcal{F} \ \Rightarrow\ \langle A\phi|\psi\rangle = \langle\phi|A^*\psi\rangle = 0\ \forall\psi \in \mathcal{F} \ \Rightarrow\ A\phi \in \mathcal{F}^\perp$, which shows that $\mathcal{F}^\perp$ is an invariant subspace. It is clear that $\mathcal{H} = \mathcal{F} \oplus \mathcal{F}^\perp$ in the sense of inner product spaces and that $A(\phi_1 + \phi_2) := A\phi_1 + A\phi_2$ for all $\phi_1 \in \mathcal{F}$ and $\phi_2 \in \mathcal{F}^\perp$. ∎

We are ready to prove Theorems 1.3.7 and 1.4.1 from the introduction. Theorem 1.3.7 follows from the following much stronger theorem.

Theorem 3.7.5 (Structure theorem for positive *-algebras) *Let $\mathcal{A}$ be a positive *-algebra. Then $\mathcal{A}$ has finitely many nonequivalent irreducible *-representations $\mathcal{H}_1, \ldots, \mathcal{H}_n$, and the map*

$$A \mapsto (l_{\mathcal{H}_1}(A), \ldots, l_{\mathcal{H}_n}(A))$$

defines an algebra isomorphism

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

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

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

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

If the multiplicities m_i from Theorem 3.7.5 are all one, then $\mathcal{H}$ is a faithful *-representation of $\mathcal{A}$ with the smallest possible dimension. In this case we say that $\mathcal{H}$ is a *minimal* *-representation of $\mathcal{A}$.

Proof of Theorem 3.7.5 By Lemma 3.6.1, $\mathcal{A}$ is semisimple, so by Theorem 3.5.5 $\mathcal{A}$ has finitely many nonequivalent irreducible representations $\mathcal{V}_1, \ldots, \mathcal{V}_n$, and the map

$$A \mapsto (l_{\mathcal{V}_1}(A), \ldots, l_{\mathcal{V}_n}(A))$$

defines an algebra isomorphism

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

Since the decomposition of a semisimple algebra into factor algebras is unique (Lemma 3.4.7), we conclude from Proposition 3.6.2 that on each $\mathcal{L}(\mathcal{V}_i)$ there exists a positive adjoint operation such that the algebra isomorphism above is a *-algebra isomorphism, i.e.,

$$A^* \mapsto (l_{\mathcal{V}_1}(A)^*, \ldots, l_{\mathcal{V}_n}(A)^*).$$

By Lemma 3.7.2, we can choose on each $\mathcal{V}_i$ an inner product such that $l_{\mathcal{V}_1}(A)^*$ is the usual adjoint with respect to this inner product. Let $\mathcal{H}_i$ denote the space $\mathcal{V}_i$ equipped with this inner product. Then $\mathcal{H}_1, \ldots, \mathcal{H}_n$ are irreducible *-representations of $\mathcal{A}$. By Lemma 3.7.3, $\mathcal{H}_1, \ldots, \mathcal{H}_n$ are up to equivalence all irreducible *-representations of $\mathcal{A}$. Finally, by Lemma 3.7.4, every *-representation $\mathcal{H}$ of $\mathcal{A}$ can be decomposed into a direct sum of *-representations. By Theorem 3.5.5, $\mathcal{H}$ is faithful if and only if all multiplicities m_i are nonzero. ∎

Our proof of Theorem 1.3.7 has been quite long and complicated, but it has yielded much more information than the simple fact that each positive *-algebra has a faithful representation. In fact, if we had known before that every positive *-algebra $\mathcal{A}$ has a faithful *-representation $\mathcal{H}$, we could have given a shorter proof of Theorem 3.7.5. In particular, we could have used Lemma 3.7.4 to decompose

$\mathcal{H}$ into irreducible *-representations, and have concluded that $\mathcal{A}$ is semisimple. I don't know if a proof of Theorem 3.7.5 along these lines is possible.[2] The difficulty seems to be that the condition that the adjoint operation is positive is rather subtle.

Proof of Theorem 1.4.1 It is clear from Theorem 3.7.5 that the algebra $\mathcal{A}$ is abelian if and only if the spaces $\mathcal{H}_i$ are all one-dimensional. In that case, $\mathcal{L}(\mathcal{H}) \cong \mathbb{C}$ and $\mathcal{A} \cong \mathbb{C}^n$. ∎

It is often useful to formulate Theorem 1.4.1 a bit differently.

Theorem 3.7.6 (Abelian positive *-algebras) *Let $\mathcal{A}$ be an abelian positive *-algebra. Then there exists a partition of the identity $\{P_1, \ldots, P_n\}$ such that*

$$\mathcal{A} = \big\{ \sum_{i=1}^{n} \lambda_i P_i : \lambda_i \in \mathbb{C} \big\}.$$

Exercise 3.7.7 Prove Theorem 3.7.6.

3.8 Representation independent quantities

From now on we will only be interested in positive *-algebras. To simplify terminology, when we speak about a *representation* $\mathcal{H}$ of a positive *-algebra $\mathcal{A}$, we will from now on always mean a faithful *-representation, unless we stress otherwise. If $\mathcal{H}$ satisfies only properties (a)–(d) from page 54, we will say that $\mathcal{H}$ is a representation of $\mathcal{A}$ *as an algebra.*

In practice, when calculating in a positive *-algebra $\mathcal{A}$, one always chooses a representation $\mathcal{H}$ to work with. Theorem 3.7.5 tells us the following. Recall the notation $A \oplus B$ from page 25.

Theorem 3.8.1 (Form of a general representation) *Let $\mathcal{A}$ be a positive *-algebra represented on an inner product space $\mathcal{H}$. Then $\mathcal{H}$ can be decomposed into orthogonal subspaces*

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

such that $\mathcal{A}$ consists of all operators of the form

$$(A_1 \oplus \cdots \oplus A_1) \oplus \cdots \oplus (A_n \oplus \cdots \oplus A_n)$$

[2]For example, in [Dav96], first Theorem 1.5.1 is proved and then it is concluded that every C^*-algebra is semisimple.

with $A_1 \in \mathcal{L}(\mathcal{H}_1), \ldots, A_n \in \mathcal{L}(\mathcal{H}_n)$. *The representation is minimal if* $m_i = 1$ $(i = 1, \ldots, n)$.

When working with a given representation, one often tacitly identifies $\mathcal{A}$ with the isomorphic positive *-algebra $\mathrm{Ran}(l_{\mathcal{H}})$. In doing so, it is important to keep track of which quantities and concepts depend on the choice of the representation, and which don't. For example, an operator $P \in \mathcal{L}(\mathcal{H})$ is a projection if and only if $P^* = P$ and $P^2 = P$, so being a projection does not depend on the choice of the representation. We conclude this chapter by showing that the spectrum of an operator is a representation-independent object.

Lemma 3.8.2 (Inverses and representations) *Let* $\mathcal{A}$ *be a positive* **-algebra represented on an inner product space* $\mathcal{H}$. *If* $A \in \mathcal{A}$ *has an inverse* $A^{-1} \in \mathcal{L}(\mathcal{H})$ *then* $A^{-1} \in \mathcal{A}$.

Proof Note that in the formulation of this lemma, we identify $\mathcal{A}$ with the isomorphic algebra $\mathrm{Ran}(l_{\mathcal{H}}) \subset \mathcal{L}(\mathcal{H})$. By Theorem 3.8.1 we can decompose $\mathcal{H}$ into orthogonal subspaces

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

and $\mathcal{A}$ consists of all operators of the form

$$A = (A_1 \oplus \cdots \oplus A_1) \oplus \cdots \oplus (A_n \oplus \cdots \oplus A_n)$$

with $A_1 \in \mathcal{L}(\mathcal{H}_1), \ldots, A_n \in \mathcal{L}(\mathcal{H}_n)$. Now it is clear that A is invertible if and only if each A_i is invertible, and in that case the inverse

$$A^{-1} = (A_1^{-1} \oplus \cdots \oplus A_1^{-1}) \oplus \cdots \oplus (A_n^{-1} \oplus \cdots \oplus A_n^{-1})$$

is also in $\mathcal{A}$. ∎

Proposition 3.8.3 (Spectrum is representation independent) *Let* $\mathcal{A}$ *be a positive* **-algebra represented on an inner product space* $\mathcal{H}$. *Then the spectrum* $\sigma(A)$ *of an operator* $A \in \mathcal{A}$ *does not depend on the choice of the faithful representation.*

Proof This follows from Lemma 3.8.2 and the fact that by Lemma 3.8.2,

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

∎

Warning: although the collection of eigenvalues of an operator $A \in \mathcal{A}$ does not depend on the representation, the dimensions of the corresponding eigenspaces *do* depend on the representation!

Exercise 3.8.4 Is the trace $\mathrm{tr}(A)$ of an operator $A \in \mathcal{A}$ representation independent?

Exercise 3.8.5 Recall that the norm of an element ϕ of an inner product space $\mathcal{H}$ is defined as $\|\phi\| := \sqrt{\langle\phi|\phi\rangle}$. Show that the operator norm

$$\|A\|_\infty := \sup_{\phi\in\mathcal{H}} \frac{\|A\phi\|}{\|\phi\|}.$$

of an element A of a positive *-algebra does not depend on the representation. Show that

$$\|A^*A\|_\infty = \|A\|_\infty^2.$$

(Hint: Theorem 2.4.13.)

Exercise 3.8.6 Let $\mathcal{A}$ be a linear space over the complex numbers satisfying all properties (i)–(viii) from page 8 of a positive *-algebra, except for property (iv). Equip the direct sum (in the sense of linear spaces) $\mathcal{A} \oplus \mathbb{C}$ with the multiplication

$$(A, \alpha)(B, \beta) := (AB + \beta A + \alpha B, \alpha\beta),$$

and the adjoint operation

$$(A, \alpha)^* := (A^*, \overline{\alpha}).$$

Show that $\mathcal{A} \oplus \mathbb{C}$ is a positive *-algebra.

Chapter 4

Quantum mechanics

4.1 Normal operators

Let $\mathcal{A}$ be a positive *-algebra and let $\mathcal{C}$ any subset of $\mathcal{A}$. By definition, the positive *-algebra *generated* by $\mathcal{C}$ is the sub-*-algebra $\alpha(\mathcal{C})$ of $\mathcal{A}$ given by

$$\alpha(\mathcal{C}) := \mathrm{span}\{A_1 \cdots A_n : A_i \in \mathcal{C} \cup \mathcal{C}^* \cup \{1\}\}.$$

It is easy to see that $\alpha(\mathcal{C})$ is the smallest sub-*-algebra of $\mathcal{A}$ containing $\mathcal{C}$ and the identity from A.

As announced at the end of Chapter 2, we now give the proof that any normal operator is of the form (2.1). This follows from the next, stronger theorem.

Theorem 4.1.1 (Commuting normal operators) *Let $\mathcal{N} \subset \mathcal{L}(\mathcal{H})$ be a collection of commuting normal operators on $\mathcal{H}$, i.e., each $A \in \mathcal{N}$ is normal and $AB = BA$ for all $A, B \in \mathcal{N}$. Then there exists an orthonormal basis $\{\phi_1, \ldots, \phi_n\}$ of $\mathcal{H}$ such that each A_α is diagonal with respect to $\{\phi_1, \ldots, \phi_n\}$.*

Remark Note that the *same* orthonormal basis $\{\phi_1, \ldots, \phi_n\}$ diagonalizes *all* normal operators $A \in \mathcal{N}$.

Proof of Theorem 4.1.1 It is easy to see that $\alpha(\mathcal{N})$, the the positive *-algebra generated by $\mathcal{N}$, is abelian.[1] By Theorem 3.7.6, there exists a partition of the identity $\{P_1, \ldots, P_n\}$ such that

$$\alpha(\mathcal{N}) = \big\{ \sum_{i=1}^{n} \lambda_i P_i : \lambda_i \in \mathbb{C} \big\}.$$

[1] This is actually not easy, although it is true.

Let $\mathcal{H}_1, \ldots, \mathcal{H}_n$ be the orthogonal subspaces of $\mathcal{H}$ such that P_i projects on $\mathcal{H}_i$. Let $\{\phi_{i,1}, \ldots, \phi_{i,m_i}\}$ be an orthonormal basis of $\mathcal{H}_i$. Then $\{\phi_{i,j} : i = 1, \ldots, n,\ j = 1, \ldots, m_i\}$ is an orthonormal basis of $\mathcal{H}$. Now

$$P_i = \sum_{j=1}^{n} |\phi_{i,j}\rangle\langle\phi_{i,j}|,$$

and each $A \in \mathcal{A}$ (in particular, each $A \in \mathcal{N}$) is of the form

$$A = \sum_{i=1}^{n} \sum_{j=1}^{m_i} \lambda_i |\phi_{i,j}\rangle\langle\phi_{i,j}|,$$

i.e., A is diagonal. ∎

We call an element A of a positive *-algebra normal if $AA^* = A^*A$. The proof of Theorem 4.1.1 carries over from $\mathcal{L}(\mathcal{H})$ to a general positive *-algebra $\mathcal{A}$, yielding the next theorem. Recall that the spectrum $\sigma(A)$ of an operator A is representation independent (Proposition 3.8.3).

Theorem 4.1.2 (Commuting normal operators) *Let $\mathcal{A}$ be a positive *-algebra and let $\mathcal{N} \subset \mathcal{A}$ be a collection of commuting normal operators in $\mathcal{A}$. Then there exists a partition of the identity $\{P_1, \ldots, P_n\}$ such that each $A \in \mathcal{N}$ is of the form*

$$A = \sum_i \lambda_i P_i \qquad (\lambda_1, \ldots, \lambda_n \in \sigma(A)). \tag{4.1}$$

Exercise 4.1.3 Let $\mathcal{A}$ be a positive *-algebra and let $A \in \mathcal{A}$ be normal. Show that the partition of the identity $\{P_1, \ldots, P_n\}$ can be chosen such that the λ_i in (4.1) are all different. Show that if the λ_i are all different, then the decomposition in (4.1) is unique up to the order of the projections P_i.

Two important examples of normal operators are hermitian operators and unitary operators. By definition, an operator $A \in \mathcal{A}$ is *hermitian*, or *self-adjoint* if $A = A^*$. An operator $U \in \mathcal{A}$ is called *unitary* if U is invertible and $U^{-1} = U^*$.

Exercise 4.1.4 Show that an operator is A is hermitian if and only if A is normal and all its eigenvalues are real.

Exercise 4.1.5 Show that an operator U is unitary if and only if U is normal and all its eigenvalues λ_i satisfy $|\lambda_i| = 1$.

Recall that the spectrum $\sigma(A)$ of an operator $A \in \mathcal{A}$ is well-defined and representation independent. A hermitian operator $A \in \mathcal{A}$ is called *positive* if all its eigenvalues are nonnegative. If $\mathcal{H}$ is a representation of $\mathcal{A}$ then this is equivalent to

$$\langle \phi | A \phi \rangle \geq 0 \qquad (\phi \in \mathcal{H}).$$

Let $\mathcal{A}_r := \{A \in \mathcal{A} : A^* = A\}$ be the space of hermitian elements of $\mathcal{A}$ and let $\mathcal{A}_+ := \{A \in \mathcal{A}_r : A \text{ is positive}\}$. We define a partial order on $\mathcal{A}_r$ by

$$A \leq B \ \Leftrightarrow \ B - A \in \mathcal{A}_+.$$

For commuting normal operators one can define a kind of 'functional calculus'. If $(A_\alpha)_{\alpha=1,\ldots,k}$ are commuting normal operators then by Theorem 4.1.2 there exists a partition of the identity $\{P_1, \ldots, P_n\}$ such that

$$A_\alpha = \sum_i \lambda_{\alpha,i} P_i \qquad (\alpha = 1, \ldots, k).$$

Now let $f : \mathbb{C}^k \to \mathbb{C}$ be any function. Then we define a normal operator $f(A_1, \ldots, A_k)$ by

$$f(A_1, \ldots, A_k) := \sum_i f(\lambda_{1,i}, \ldots, \lambda_{k,i}) P_i. \tag{4.2}$$

Exercise 4.1.6 Check that the definition in (4.2) does not depend on the choice of the partition of the identity $\{P_1, \ldots, P_n\}$.

Exercise 4.1.7
(a) If A_1, A_2 are commuting normal operators, $\alpha_1, \alpha_2 \in \mathbb{C}$, and $f : \mathbb{C}^2 \to \mathbb{C}$ is defined as $f(\lambda_1, \lambda_2) := \alpha_1 \lambda_1 + \alpha_2 \lambda_2$, then show that $f(A_1, A_2) := \alpha_1 A_1 + \alpha_2 A_2$.

(b) If A is a normal operator and $f : \mathbb{C} \to \mathbb{C}$ is defined as $f(\lambda) := \lambda^n$ $(n \geq 0)$, then show that $f(A) = A^n$.

(c) If A is an invertible normal operator and $f : \mathbb{C} \backslash \{0\} \to \mathbb{C}$ is defined as $f(\lambda) := \lambda^{-1}$, then show that $f(A) = A^{-1}$.

(d) If A is a normal operator and $f : \mathbb{C} \to \mathbb{C}$ is defined as $f(\lambda) := \overline{\lambda}$, then show that $f(A) = A^*$. If $f(\lambda) := \mathrm{Re}(\lambda)$ or $f(\lambda) := \mathrm{Im}(\lambda)$, then $f(A) = \mathrm{Re}(A)$ or $f(A) = \mathrm{Im}(A)$, respectively.

Exercise 4.1.8 Let $\mathcal{A}$ be a positive *-algebra and let $A \in \mathcal{A}$ be a positive hermitian operator. Show that there exists a unique positive hermitian operator $\sqrt{A} \in \mathcal{A}$ such that $(\sqrt{A})^2 = A$.

Exercise 4.1.9 Show that a probability ρ on a positive *-algebra $\mathcal{A}$ is uniquely determined by all probabilities $\rho(P)$ with $P \in \mathcal{A}$ a projection.

4.2 Observables and states

Let $\mathcal{A}$ be a positive *-algebra and let $A \in \mathcal{A}$ be hermitian. Then A can written as

$$A = \sum_i \lambda_i P_i$$

where $\{P_1, \ldots, P_n\}$ is a partition of the identity and $\lambda_1, \ldots, \lambda_n$ are distinct real numbers. By Exercise 4.1.3, this decomposition is unique up to the order of the projections P_i. In quantum mechanics, A is interpreted as an *observable*, i.e., a real-valued physical quantity whose values we can measure. $\{P_1, \ldots, P_n\}$ is interpreted as an ideal measurement of the observable A and P_i represents the observation that the observable A has the value λ_i. Note that if ρ is a probability on $\mathcal{A}$ then

$$\rho(A) = \sum_i \lambda_i \rho(P_i)$$

where $\rho(P_i)$ is the probability that an ideal measurement of A will yield the value λ_i. Thus, we can interpret $\rho(A)$ as the *expected value* or *expectation* of the observable A.

If $(A_1, \ldots, A_d)$ are commuting hermitian operators, then there exists a partition of the identity $\{P_1, \ldots, P_n\}$ and real numbers $\lambda_{i,\alpha}$ such that

$$A_\alpha = \sum_i \lambda_{i,\alpha} P_i \qquad (\alpha = 1, \ldots, d).$$

If we moreover insist that

$$(\lambda_{i,1}, \ldots, \lambda_{i,d}) \neq (\lambda_{j,1}, \ldots, \lambda_{j,d}) \qquad (i \neq j),$$

then this way of writing the A_α is unique up to the order of the projections P_i. We interpret $\{P_1, \ldots, P_n\}$ as an ideal simultaneous measurement of the observables $(A_1, \ldots, A_d)$. The projection P_i corresponds to the observation that the values of $(A_1, \ldots, A_d)$ are $(\lambda_{i,1}, \ldots, \lambda_{i,d})$. We interpret $(A_1, \ldots, A_d)$ as an observable with values in $\mathbb{R}^d$.

If A, B are noncommuting hermitian operators, then there exist no simultaneous ideal measurements of A and B and we cannot interpret (A, B) as an observable with values in $\mathbb{R}^2$.

Let $\mathcal{A}$ be a positive *-algebra and let $\mathcal{A}'$ be the space of linear forms on $\mathcal{A}$, i.e., linear maps $l : \mathcal{A} \to \mathbb{C}$. A linear form $l \in \mathcal{A}'$ is called *real* if

$$l(A^*) = \overline{l(A)} \qquad (A \in \mathcal{A}).$$

A *positive* linear form is a linear form such that

$$l(A^*A) \geq 0 \qquad (A \in \mathcal{A}).$$

We say that l is *strictly* positive if moreover

$$l(A^*A) = 0 \ \Rightarrow A = 0.$$

Recall the definition of a probability ρ on $\mathcal{A}$ from Section 1.3. Note that a probability is a positive real form ρ with $\rho(1) = 1$. A *pseudotrace* on $\mathcal{A}$ is a strictly positive linear form $\tau : \mathcal{A} \to \mathbb{C}$ such that

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

Proposition 4.2.1 (Pseudotraces) *On each positive *-algebra $\mathcal{A}$ there exists a pseudotrace τ. If $\mathcal{A}$ is a factor algebra, then τ is unique up to a multiplicative strictly positive constant.*

Proof By Theorem 1.3.7, $\mathcal{A}$ is isomorphic to a sub-*-algebra of some $\mathcal{L}(\mathcal{H})$. Now the trace on $\mathcal{L}(\mathcal{H})$ has all the properties of a pseudotrace (see Exercise 2.4.11). This shows the existence of a pseudotrace (and explains its name). Now let $\mathcal{A} \cong \mathcal{L}(\mathcal{H})$ be a factor algebra, and let τ be a pseudotrace on $\mathcal{L}(\mathcal{H})$. Fix some $\phi \in \mathcal{H}$ with $\|\phi\| = 1$. Then, for arbitrary $\psi_1, \psi_2 \in \mathcal{H}$,

$$\tau(|\psi_1\rangle\langle\psi_2|) = \tau(|\psi_1\rangle\langle\phi|\phi\rangle\langle\psi_2|) = \tau(|\phi\rangle\langle\psi_2|\psi_1\rangle\langle\phi|) = \tau(|\phi\rangle\langle\phi|)\langle\psi_2|\psi_1\rangle.$$

By linearity, it follows that τ is unique up to the strictly positive constant $\tau(|\phi\rangle\langle\phi|)$. ∎

Note that in the choice of a pseudotrace on a general positive *-algebra $\mathcal{A}$ there is a free multiplicative constant for each of the factors that $\mathcal{A}$ consists of; if $\mathcal{A}$ is abelian then every strictly positive linear form is a pseudotrace.

Proposition 4.2.2 (Density operator) *Let $\mathcal{A}$ be a positive *-algebra and let τ be a 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.

We call R the *density* of ρ with respect to τ.

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, note that it is straightforward to check that the formula

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

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 4.2.3 Show that every real linear form l can be written as $l = p\rho_+ - q\rho_-$ where $p, q \geq 0$ and ρ_+, ρ_- are probabilities. Show that every linear form l can be written as $l = \mathrm{Re}(l) + i\mathrm{Im}(l)$ where $\mathrm{Re}(l)$ and $\mathrm{Im}(l)$ are real linear forms.

Let $\mathcal{A}$ be a positive *-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*.

Proposition 4.2.4 (Pure states and state vectors) *Let $\mathcal{A}$ be a positive *-algebra, minimally represented on an inner product space $\mathcal{H}$. Let $\mathcal{H}_1, \ldots, \mathcal{H}_n$ be subspaces of $\mathcal{H}$ such that $\mathcal{H} = \mathcal{H}_1 \oplus \ldots \oplus \mathcal{H}_n$ and*

$$\mathcal{A} = \{A_1 \oplus \cdots \oplus A_n : A_1 \in \mathcal{L}(\mathcal{H}_1), \ldots, A_n \in \mathcal{L}(\mathcal{H}_n)\}.$$

Let ρ be a probability on $\mathcal{A}$. Then ρ is a pure state if and only if there exists a $\mathcal{H}_i$ and a $\psi \in \mathcal{H}_i$ with $\|\psi\| = 1$ such that

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

Every probability ρ on $\mathcal{A}$ can be written as

$$\rho(A) = \sum_j p_j \langle\psi_j|A\psi_j\rangle$$

where $A \mapsto \langle\psi_j|A\psi_j\rangle$ are pure states. It is possible to choose the ψ_j orthonormal.

Proof If $\psi \in \mathcal{H}$ satisfies $\|\psi\| = 1$, then it is easy to see that

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

defines a probability on $\mathcal{A}$. Now let ρ be any probability on $\mathcal{A}$ and let R be the density of ρ with respect to the trace tr. Then R is of the form

$$R = R_1 \oplus \cdots \oplus R_n,$$

with $R_i \in \mathcal{L}(\mathcal{H}_i)$ self-adjoint and positive. Therefore (recall Theorem 2.4.13) there exist orthonormal bases $\{\phi_{i,1}, \ldots, \phi_{i,d_i}\}$ for $\mathcal{H}_i$ $(i = 1, \ldots, n)$ and nonnegative numbers $p_{i,j}$ such that

$$R_i = \sum_{j=1}^{d_i} p_{i,j} |\phi_{i,j}\rangle\langle\phi_{i,j}|.$$

It follows that

$$\rho(A) = \mathrm{tr}(RA) = \sum_{i=1}^{n} \sum_{j=1}^{d_i} p_{i,j} \langle\phi_{i,j}|A|\phi_{i,j}\rangle,$$

i.e.,

$$\rho = \sum_{i=1}^{n} \sum_{j=1}^{d_i} p_{i,j} \rho_{\phi_{i,j}}.$$

From the fact that $\rho(1) = 1$ it is easy to see that the $p_{i,j}$ sum up to one. Thus, every probability can be written as a convex combination of probabilities of the form ρ_{ψ_j} such that each ψ_j is in some $\mathcal{H}_i$, and the ψ_j can be chosen orthonormal. We must show that such ρ_{ψ_j} are pure states, i.e., they cannot be decomposed further. Assume that $\psi \in \mathcal{H}_i$ for some i, $\|\psi\| = 1$, and that $\rho_\psi = p\rho_1 + (1-p)\rho_2$

for some probabilities ρ_1, ρ_2 on $\mathcal{A}$ and $0 < p < 1$. By what we have proved we can write

$$\rho_1 = \sum_{i=1}^{n} \sum_{j=1}^{d_i} p^1_{i,j} \rho_{\phi^1_{i,j}} \quad \text{and} \quad \rho_2 = \sum_{i=1}^{n} \sum_{j=1}^{d_i} p^2_{i,j} \rho_{\phi^2_{i,j}}$$

where $\{\phi^1_{i,1}, \ldots, \phi^1_{i,d_i}\}$ and $\{\phi^2_{i,1}, \ldots, \phi^2_{i,d_i}\}$ are orthonormal bases for $\mathcal{H}_i$. For any probability ρ write

$$N_\rho := \{\phi \in \mathcal{H} : \rho(|\phi\rangle\langle\phi|) = 0\}.$$

Then

$$\begin{aligned} N_{\rho_1} &= \mathrm{span}\{\phi^1_{i,j} : p^1_{i,j} = 0\}, \\ N_{\rho_2} &= \mathrm{span}\{\phi^2_{i,j} : p^2_{i,j} = 0\}, \end{aligned}$$

and

$$N_{\rho_\psi} = \{\phi \in \mathcal{H} : \langle\phi|\psi\rangle = 0\}.$$

In particular, $\dim(N_{\rho_\psi}) = \dim(\mathcal{H}) - 1$. Since

$$N_{p\rho_1 + (1-p)\rho_2} = N_{\rho_1} \cap N_{\rho_2},$$

we see that $\rho_\psi = p\rho_1 + (1-p)\rho_2$ is only possible if $\dim(N_{\rho_1}) = \dim(N_{\rho_2}) = \dim(\mathcal{H}) - 1$ and

$$N_{\rho_\psi} = N_{\rho_1} = N_{\rho_2}.$$

But this is only possible if $\rho_\psi = \rho_1 = \rho_2$. Thus ρ_ψ cannot be written as a nontrivial convex combination of other states. ∎

Let $\mathcal{A}$ be a positive *-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. The following reformulation of Proposition 4.2.4 in a representation-free language is sometimes useful.

Lemma 4.2.5 (Pure states and minimal projections) *If P is a minimal projection in a positive *-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 probability 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.

Proof Let $\mathcal{H} = \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$ be a minimal representation of $\mathcal{A}$ as in Proposition 4.2.4 and let P be a minimal projection in $\mathcal{A}$. Let $\mathcal{F}$ be the space that P projects on, i.e., $P = P_{\mathcal{F}}$. Since each $P_{\mathcal{H}_i}$ commutes with every operator in $\mathcal{A}$ we see that $P_{\mathcal{H}_i}P_{\mathcal{F}} = P_{\mathcal{H}_i\cap\mathcal{F}}$ is a projection that ist smaller than $P_{\mathcal{F}}$. By mininality, either $P_{\mathcal{H}_i\cap\mathcal{F}} = 0$ or $P_{\mathcal{H}_i\cap\mathcal{F}} = P_{\mathcal{F}}$. It follows that $\mathcal{F}$ is a subspace of some $\mathcal{H}_i$. Using minimality once more, it is easy to see that $\mathcal{F}$ must be one-dimensional, so the minimal projections in $\mathcal{A}$ are exactly the operators P of the form

$$P = |\psi\rangle\langle\psi| \quad \text{with } \psi \in \mathcal{H}_i \text{ for some } i = 1, \ldots, n, \ \|\psi\| = 1.$$

Then

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

so the statements follow from Proposition 4.2.4. ∎

Note that Proposition 4.2.4 or Lemma 4.2.5 show in particular that every probability can be written as a convex combination of pure states. This decomposition is not unique! In the special case that $\mathcal{A}$ is a factor algebra and $\mathcal{H}$ is a minimal representation, Proposition 4.2.4 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.

For simple quantum mechanical systems, it is in many cases easy to find a minimal representation $\mathcal{H}$ for the positive *-algebra $\mathcal{A}$ describing the system. Moreover, such a minimal representation $\mathcal{H}$ appears to have a 'natural' construction. Most of the time, $\mathcal{A}$ is actually a factor algebra so that $\mathcal{A} \cong \mathcal{L}(\mathcal{H})$. If $\mathcal{A}$ is not a factor, then physicists say that there is a *super selection rule* forbidding certain hermitian operators in $\mathcal{L}(\mathcal{H})$ from being observables.[2]

As the quantum mechanical systems that one considers get more complicated, one soon discovers that minimal representations are not so easy to find and do not appear 'natural' at all. In such cases, the most efficient way to describe the system is to describe the relevant positive *-algebra in a representation free language.

In Section 1.4, we called a probability ρ on a positive *-algebra a precise state if $\rho(P) \in \{0, 1\}$ for all projections $P \in \mathcal{A}$. As announced there, we can prove:

[2]For example, if one considers a number of elementary particles, such as electrons, in a finite region of space, then any observable that can be measured in this finite region must commute with the hermitian operator corresponding to the number of particles.

Lemma 4.2.6 (No decomposition in precise states) *If $\mathcal{A}$ is a positive *-algebra and $\mathcal{A}$ is not abelian, then not every probability can be written as a convex combination of precise states.*

Proof Since pure states cannot be written as a nontrivial convex combination of other states, it suffices to prove that there exist pure states that are not precise. Let $\mathcal{H}$ be a minimal representation of $\mathcal{A}$ and let $\mathcal{H}_1, \ldots, \mathcal{H}_n$ be subspaces as in Proposition 4.2.4. By nonabelianness, at least one $\mathcal{H}_i$ is not one-dimensional and we can find $\psi, \phi \in \mathcal{H}_i$ with $\|\psi\| = \|\phi\| = 1$ and $\langle\psi|\phi\rangle \neq 0, 1$. Then

$$\rho_\psi(|\phi\rangle\langle\phi|) = \langle\psi|\phi\rangle\langle\phi|\psi\rangle = |\langle\psi|\phi\rangle|^2 \notin \{0, 1\},$$

hence ρ_ψ is not precise. ∎

4.3 The Schrödinger equation

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

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

Scalar means that we assume that the particle has no internal degrees of freedom (such as spin). *Massive* means that the particle has a mass $m > 0$, measured in kilogram. A *point particle* means a particle that is so small that we can neglect its spatial extensions. A *potential*, finally, is a function $V : \mathbb{R}^d \to \mathbb{R}$ which tells you that if the particle is at the point $x \in \mathbb{R}^d$, then its potential energy[3] is $V(x)$, measured in Joule, where one Jou is one kg met^2 sec^{-2}.

With these assumptions, the *Schrödinger equation*[4] describing this particle is

$$i\hbar\tfrac{\partial}{\partial t}\psi_t(x) = V(x)\psi_t(x) - \frac{\hbar^2}{2m}\sum_{k=1}^{d} \tfrac{\partial^2}{\partial x_k{}^2}\psi_t(x) \qquad (t \in \mathbb{R},\ x \in \mathbb{R}^d). \tag{4.3}$$

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

[4]Suggested by Schrödinger to describe the behavior of an electron in the Coulomb potential from the kern of an atom.

Here $\hbar$ is Planck's constant,[5]

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

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

To relate (4.3) to the formalism of quantum mechanics, consider the linear space

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

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

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

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

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

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

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

corresponds to the energy of the particle. The operator H is also called the *Hamiltonian*. Note that with this notation, equation (4.3) takes the form

$$i\hbar\tfrac{\partial}{\partial t}\psi_t = H\psi_t \qquad (t \in \mathbb{R}). \tag{4.4}$$

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

One can show that for each $\psi_0 \in \mathcal{H}$ equation (4.3) (suitably interpreted) has a unique solution $(\psi_t)_{t\in\mathbb{R}}$ in $\mathcal{H}$ satisfying

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

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

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

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

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

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

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

4.4 Deterministic time evolution

In this section we return to finite dimensions. We moreover choose units such that $\hbar = 1$. Assume that $\mathcal{A}$ is a positive *-algebra, minimally represented on some inner product space $\mathcal{H}$. Let $H \in \mathcal{A}$ be any hermitian operator. Motivated by the Schrödinger equation in its form (4.4) we look at the differential equation

$$i\tfrac{\partial}{\partial t}\psi_t = H\psi_t. \tag{4.5}$$

This is just a linear differential equation in finite dimensions, so it is easy to see that for each initial condition $\psi_0 \in \mathcal{H}$ there is a unique continuously differentiable solution $t \mapsto \psi_t$. In fact, this solution is given by

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

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

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

We interpret the vector ψ_t, normalized such that $\|\psi_t\| = 1$, as the pure state of the physical system under consideration. More precisely, we interpret (4.5) as saying that if the system at time $t = 0$ is described by the pure state

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

then at time t the state is

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

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

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

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

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

Exercise 4.4.3 Show that deterministic time evolution of the type just described preserves pure states, i.e., if ρ_0 is a pure state then ρ_t is a pure state. (If $\mathcal{A}$ is not a factor algebra this is not completely trivial!)

We will later see that a deeper converse of Exercise 4.4.3 holds: any time evolution, satisfying certain natural conditions, that maps pure states into pure states is of the type just described.

The fact that our time evolution maps pure states, which are probabilities that give 'maximal knowledge' about our system, into pure states, is why we have called the time evolution described by a Hamiltonian deterministic. A remarkable property of this sort of time evolution is that it is *reversable*, in the sense that from the state at time t we can deduce the state at all earlier times. This is true regardless of the fact that if the Schrödinger equation (4.3) with $V \equiv 0$ is started in a pure state ψ_0 that is localized in a small region of space, then as time tends to infinity the solution of (4.3) becomes more and more spread out in space.

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

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

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

Chapter 5

Subalgebras and independence

5.1 The tensor product

Let $\mathcal{V}_1, \mathcal{V}_2$, and $\mathcal{W}$ be linear spaces (as usual finite dimensional). By definition, a map $b : \mathcal{V}_1 \times \mathcal{V}_2 \to \mathcal{W}$ is *bilinear* if

$$\begin{aligned} \phi_1 &\mapsto b(\phi_1, \phi_2) \quad \text{is linear for each fixed } \phi_2 \in \mathcal{V}_2, \\ \phi_2 &\mapsto b(\phi_1, \phi_2) \quad \text{is linear for each fixed } \phi_1 \in \mathcal{V}_1. \end{aligned}$$

Proposition 5.1.1 (Definition of the tensor product) *For any two linear spaces $\mathcal{V}_1, \mathcal{V}_2$ there exists a linear space $\mathcal{V}_1 \otimes \mathcal{V}_2$, called the tensor product of $\mathcal{V}_1$ and $\mathcal{V}_2$, and a bilinear map $(\phi_1, \phi_2) \mapsto \phi_1 \otimes \phi_2$ from $\mathcal{V}_1 \times \mathcal{V}_2$ into $\mathcal{V}_1 \otimes \mathcal{V}_2$, satisfying the following equivalent properties:*

(i) *There exist bases $\{\phi^1_1, \ldots, \phi^1_{d_1}\}$ and $\{\phi^2_1, \ldots, \phi^2_{d_2}\}$ for $\mathcal{V}_1$ and $\mathcal{V}_2$, such that $\{\phi^1_i \otimes \phi^2_j : i = 1, \ldots, d_1,\ j = 1, \ldots, d_2\}$ is a basis for $\mathcal{V}_1 \otimes \mathcal{V}_2$.*

(ii) *For any bases $\{\phi^1_1, \ldots, \phi^1_{d_1}\}$ and $\{\phi^2_1, \ldots, \phi^2_{d_2}\}$ for $\mathcal{V}_1$ and $\mathcal{V}_2$, $\{\phi^1_i \otimes \phi^2_j : i = 1, \ldots, d_1,\ j = 1, \ldots, d_2\}$ is a basis for $\mathcal{V}_1 \otimes \mathcal{V}_2$.*

(iii) *For any $l_1 \in \mathcal{V}_1'$ and $l_2 \in \mathcal{V}_2'$ there exists a unique $l \in (\mathcal{V}_1 \otimes \mathcal{V}_2)'$ such that $l(\phi_1 \otimes \phi_2) = l_1(\phi_1) l_2(\phi_2)$ for all $\phi_1 \in \mathcal{V}_1$, $\phi_2 \in \mathcal{V}_2$.*

(iv) *For any linear space $\mathcal{W}$ and for any bilinear map $b : \mathcal{V}_1 \times \mathcal{V}_2 \to \mathcal{W}$, there exists a unique linear map $\overline{b} : \mathcal{V}_1 \otimes \mathcal{V}_2 \to \mathcal{W}$ such that $\overline{b}(\phi_1 \otimes \phi_2) = b(\phi_1, \phi_2)$ for all $\phi_1 \in \mathcal{V}_1, \phi_2 \in \mathcal{V}_2$.*

The tensor product of $\mathcal{V}_1$ and $\mathcal{V}_2$ is unique in the following sense: if a linear space $\mathcal{V}_1 \tilde{\otimes} \mathcal{V}_2$ together with a bilinear map $(\phi_1, \phi_2) \mapsto \phi_1 \tilde{\otimes} \phi_2$ satisfies the properties (i)–(iv)*, then there exists a unique linear bijection $l : \mathcal{V}_1 \tilde{\otimes} \mathcal{V}_2 \to \mathcal{V}_1 \otimes \mathcal{V}_2$ such that $l(\phi_1 \tilde{\otimes} \phi_2) = \phi_1 \otimes \phi_2$ for all $\phi_1 \in \mathcal{V}_1$ and $\phi_2 \in \mathcal{V}_2$.*

Proof We start by showing that there exists a linear space $\mathcal{V}_1 \otimes \mathcal{V}_2$ and a bilinear map $(\phi_1, \phi_2) \mapsto \phi_1 \otimes \phi_2$ from $\mathcal{V}_1 \times \mathcal{V}_2$ into $\mathcal{V}_1 \otimes \mathcal{V}_2$, satisfying property (i). Fix bases $\{\phi_1^1, \ldots, \phi_{d_1}^1\}$ and $\{\phi_1^2, \ldots, \phi_{d_2}^2\}$ for $\mathcal{V}_1$ and $\mathcal{V}_2$, respectively. Choose for $\mathcal{V}_1 \otimes \mathcal{V}_2$ any linear space with dimension $d_1 d_2$, choose a basis for $\mathcal{V}_1 \otimes \mathcal{V}_2$, give the $d_1 d_2$ basis vectors the names

$$\phi_i^1 \otimes \phi_j^2 \qquad \text{with } i = 1, \ldots, d_1,\ j = 1, \ldots, d_2,$$

and define a bilinear map $(\phi_1, \phi_2) \mapsto \phi_1 \otimes \phi_2$ from $\mathcal{V}_1 \times \mathcal{V}_2$ into $\mathcal{V}_1 \otimes \mathcal{V}_2$ by

$$\Big(\sum_{i=1}^{d_1} \lambda_i \phi_i^1\Big) \otimes \Big(\sum_{j=1}^{d_2} \mu_j \phi_j^2\Big) := \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} \lambda_i \mu_j \, \phi_i^1 \otimes \phi_j^2.$$

Then obviously (i) holds. We now show that (i)⇒(iv)⇒(iii)⇒(ii)⇒(i).

(i)⇒(iv): Define

$$\overline{b}(\phi_i^1 \otimes \phi_j^2) := b(\phi_i^1, \phi_j^2) \qquad (i = 1, \ldots, d_1,\ j = 1, \ldots, d_2).$$

Since the $\phi_i^1 \otimes \phi_j^2$ are a basis of $\mathcal{V}_1 \otimes \mathcal{V}_2$, this definition extends to a unique linear map $\overline{b} : \mathcal{V}_1 \otimes \mathcal{V}_2 \to \mathcal{W}$. Since b is bilinear, it is easy to see that

$$\overline{b}(\phi_1 \otimes \phi_2) = b(\phi_1, \phi_2) \qquad \forall\, \phi_1 \in \mathcal{V}_1,\ \phi_2 \in \mathcal{V}_2.$$

This proves (iv).

(iv)⇒(iii): Obvious, since $(\phi_1, \phi_2) \mapsto l_1(\phi_1) l_2(\phi_2)$ is bilinear.

(iii)⇒(ii): Let $\{\phi_1^1, \ldots, \phi_{d_1}^1\}$ and $\{\phi_1^2, \ldots, \phi_{d_2}^2\}$ be bases for $\mathcal{V}_1$ and $\mathcal{V}_2$, respectively. We claim that $\{\phi_i^1 \otimes \phi_j^2 : i = 1, \ldots, d_1,\ j = 1, \ldots, d_2\}$ is a basis for $\mathcal{V}_1 \otimes \mathcal{V}_2$. Assume that

$$\sum_{ij} \lambda_{ij} \, \phi_i^1 \otimes \phi_j^2 = 0.$$

For any $l_1 \in \mathcal{V}_i'$ and $l_2 \in \mathcal{V}_i'$, there exists a linear form l on $\mathcal{V}_1 \otimes \mathcal{V}_2$ such that $l(\phi_1 \otimes \phi_2) = l_1(\phi_1) l_2(\phi_2)$ for all $\phi_1 \in \mathcal{V}_1$, $\phi_2 \in \mathcal{V}_2$, and therefore,

$$\sum_{ij} \lambda_{ij} \, l_1(\phi_i^1) l_2(\phi_j^2) = l\Big(\sum_{ij} \lambda_{ij} \, \phi_i^1 \otimes \phi_j^2\Big) = l(0) = 0,$$

In particular, we may choose

$$l_1(\phi_i^1) = \delta_{ik} \quad \text{and} \quad l_2(\phi_i^2) = \delta_{jl}.$$

This shows that $\lambda_{kl} = 0$ for all k, l, i.e., the vectors $\phi^1_i \otimes \phi^2_j$ are linearly independent. It is easy to see that if the vectors $\phi^1_i \otimes \phi^2_j$ would not span $\mathcal{V}_1 \otimes \mathcal{V}_2$, then the linear form l would not be unique. Thus $\{\phi^1_i \otimes \phi^2_j : i = 1, \ldots, d_1,\ j = 1, \ldots, d_2\}$ is a basis for $\mathcal{V}_1 \otimes \mathcal{V}_2$.

(ii)$\Rightarrow$(i): trivial.

Now imagine that $\mathcal{V}_1 \tilde{\otimes} \mathcal{V}_2$ together with the bilinear map $(\phi_1, \phi_2) \mapsto \phi_1 \tilde{\otimes} \phi_2$ also satisfies the properties (i)–(iv). Then, since $(\phi_1, \phi_2) \mapsto \phi_1 \otimes \phi_2$ is bilinear, by property (iv) there exists a unique linear map $l : \mathcal{V}_1 \tilde{\otimes} \mathcal{V}_2 \to \mathcal{V}_1 \otimes \mathcal{V}_2$ such that $l(\phi_1 \tilde{\otimes} \phi_2) = \phi_1 \otimes \phi_2$ for all $\phi_1 \in \mathcal{V}_1,\ \phi_2 \in \mathcal{V}_2$. By property (i), l maps a basis of $\mathcal{V}_1 \tilde{\otimes} \mathcal{V}_2$ to a basis of $\mathcal{V}_1 \otimes \mathcal{V}_2$ so l is a linear bijection. ∎

If $l_1 \in \mathcal{V}_1'$ and $l_2 \in \mathcal{V}_2'$ are linear forms on $\mathcal{V}_1$ and $\mathcal{V}_2$, respectively, then we denote the linear form l on $\mathcal{V}_1 \otimes \mathcal{V}_2$ from property (iii) of Proposition 5.1.1 by $l_1 \otimes l_2$. I.e.,

$$l_1 \otimes l_2(\phi_1 \otimes \phi_2) := l_1(\phi_1) l_2(\phi_2) \qquad (\phi_1 \in \mathcal{V}_1,\ \phi_2 \in \mathcal{V}_2).$$

The next lemma says that this is good notation.

Lemma 5.1.2 (Tensor product and dual spaces) *The linear space $(\mathcal{V}_1 \otimes \mathcal{V}_2)'$ together with the bilinear map $(l_1, l_2) \mapsto l_1 \otimes l_2$ is a version of the tensor product $\mathcal{V}_1' \otimes \mathcal{V}_2'$.*

Proof Choose bases $\{\phi^1_1, \ldots, \phi^1_{d_1}\}$ and $\{\phi^2_1, \ldots, \phi^2_{d_2}\}$ for $\mathcal{V}_1$ and $\mathcal{V}_2$, respectively. Let $\{l^1_1, \ldots, l^1_{d_1}\}$ and $\{l^2_1, \ldots, l^2_{d_2}\}$ be the corresponding dual bases for $\mathcal{V}_1'$ and $\mathcal{V}_2'$, i.e.,

$$l^1_k(\phi^1_i) := \delta_{ik} \quad \text{and} \quad l^2_l(\phi^2_j) := \delta_{jl}.$$

Then

$$l^1_k \otimes l^2_l\,(\phi^1_i \otimes \phi^2_j) = \delta_{ik}\delta_{jl} = \delta_{(i,j),(k,l)},$$

which shows that $\{l^1_k \otimes l^2_l : k = 1, \ldots, d_1,\ l = 1, \ldots, d_2\}$ is the dual basis to $\{\phi^1_i \otimes \phi^2_j : i = 1, \ldots, d_1,\ j = 1, \ldots, d_2\}$. Therefore, by property (i) from Proposition 5.1.1, $(\mathcal{V}_1 \otimes \mathcal{V}_2)'$ together with the bilinear map $(l_1, l_2) \mapsto l_1 \otimes l_2$ is a version of the tensor product $\mathcal{V}_1' \otimes \mathcal{V}_2'$. ∎

Exercise 5.1.3 Let $\mathcal{V}_1, \mathcal{V}_2$ and $\mathcal{W}$ be linear spaces over the complex numbers and let $c : \mathcal{V}_1 \times \mathcal{V}_2 \to \mathcal{W}$ be a map such that $\phi_1 \mapsto c(\phi_1, \phi_2)$ is colinear for each fixed $\phi_2 \in \mathcal{V}_2$ and $\phi_2 \mapsto c(\phi_1, \phi_2)$ is colinear for each fixed $\phi_1 \in \mathcal{V}_1$. Show that there exists a unique colinear map $\overline{c} : \mathcal{V}_1 \otimes \mathcal{V}_2 \to \mathcal{W}$ such that $\overline{c}(\phi_1 \otimes \phi_2) = c(\phi_1, \phi_2)$ $\forall\ \phi_1 \in \mathcal{V}_1, \phi_2 \in \mathcal{V}_2$.

Exercise 5.1.4 Let Ω_1, Ω_2 be finite sets. Show that $\mathbb{C}^{\Omega_1} \otimes \mathbb{C}^{\Omega_2} \cong \mathbb{C}^{\Omega_1 \times \Omega_2}$. Here $\cong$ means that there is a natural linear bijection between the two spaces.

Exercise 5.1.5 Express $\dim(\mathcal{V}_1 \oplus \mathcal{V}_2)$ and $\dim(\mathcal{V}_1 \otimes \mathcal{V}_2)$ in $\dim(\mathcal{V}_1)$ and $\dim(\mathcal{V}_2)$.

Exercise 5.1.6 Assume that $\dim(\mathcal{V}_1) \geq 2$ and $\dim(\mathcal{V}_2) \geq 2$. Show that there exists a $\psi \in \mathcal{V}_1 \otimes \mathcal{V}_2$ that is not of the form $\psi = \phi_1 \otimes \phi_2$ with $\phi_1 \in \mathcal{V}_1$ and $\phi_2 \in \mathcal{V}_2$.

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

Exercise 5.1.7 Show that the tensor product is associative: $(\mathcal{V}_1 \otimes \mathcal{V}_2) \otimes \mathcal{V}_3 \cong \mathcal{V}_1 \otimes (\mathcal{V}_2 \otimes \mathcal{V}_3)$. Here $\cong$ means that there is a natural linear bijection between the two spaces.

We now give a short introduction to tensor calculus. If $\mathcal{V}_1, \ldots, \mathcal{V}_n$ are linear spaces and

$$\{e^1_1, \ldots, e^1_{d_1}\}, \ldots, \{e^n_1, \ldots, e^n_{d_1}\}$$

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

$$\big\{e^1_{i_1} \otimes \cdots e^n_{i_n} : i_1 = 1, \ldots, d_1, \ldots, i_n = 1, \ldots, d_n\big\}$$

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

$$T = \sum_{i_1=1}^{d_1} \cdots \sum_{i_n=1}^{d_n} T_{i_1 \ldots i_n} e^1_{i_1} \otimes \cdots e^n_{i_n}.$$

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

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

so

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

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

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

Note that this formula is linear in each component, so by (a generalisation of) property (iv) of Proposition 5.1.1, c_{23} extends to a unique linear map

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

If $\{e_1, \ldots, e_n\}$ is a basis for $\mathcal{V}$, $\{l_1, \ldots, l_n\}$ is the corresponding dual basis for $\mathcal{V}'$, and $\{f_1, \ldots, f_m\}$ is a basis for $\mathcal{W}$, then

$$c_{23}\Big(\sum_{ijk} T_{ijk} f_i \otimes l_j \otimes e_k\Big) = \sum_{ijk} T_{ijk} l_j(e_k) f_i = \sum_{ij} T_{ijj} f_i,$$

so with respect to these bases

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

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

Lemma 5.1.8 (Examples of contractions) *For any linear spaces $\mathcal{V}, \mathcal{W}$ there exists a natural linear isomorphism*

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

If $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_m\}$ are bases for $\mathcal{V}$ and $\mathcal{W}$, and $\{l_1, \ldots, l_n\}$ is the dual basis of $\{e_1, \ldots, e_n\}$, then in this isomorphism,

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

where A_{ij} is the matrix of A with respect to the bases $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_m\}$. Moreover,

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

Here, in (i), $c_{12} : \mathcal{V} \otimes \mathcal{V}' \to \mathbb{C}$ is defined as $c_{12}(\phi \otimes l) := l(\phi)$. In (iii), $A' \in \mathcal{L}(\mathcal{W}', \mathcal{V}')$ denotes the adjoint of A (see page 25).

Proof of Lemma 5.1.8 Every $A \in \mathcal{W} \otimes \mathcal{V}'$ defines a linear operator $\tilde{A} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ by

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

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

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

Consistent with our notation for tensors, write a vector $x \in \mathcal{V}$ as $x = \sum_i x_i e_i$. Then

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

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

$$\begin{array}{rll}
\text{(i)} & \operatorname{tr}(A) = \sum_i A_{ii} & (A \in \mathcal{V} \otimes \mathcal{V}'), \\
\text{(ii)} & (Ax)_i = \sum_j A_{ij} x_j & (A \in \mathcal{W} \otimes \mathcal{V}',\ x \in \mathcal{V}), \\
\text{(iii)} & (A'y)_i = \sum_j y_i A_{ij} & (A \in \mathcal{W} \otimes \mathcal{V}',\ y \in \mathcal{W}'), \\
\text{(iv)} & (AB)_{ij} = \sum_k A_{ik} B_{kj} & (A \in \mathcal{W} \otimes \mathcal{V}',\ B \in \mathcal{V} \otimes \mathcal{F}'),
\end{array}$$

which are all well-known facts. ∎

If $\mathcal{A}_1, \mathcal{A}_2$ are algebras and $\mathcal{A}_1 \otimes \mathcal{A}_2$ is their tensor product in the sense of linear spaces, then we equip $\mathcal{A}_1 \otimes \mathcal{A}_2$ with the structure of an algebra by putting

$$(A_1 \otimes A_2)(B_1 \otimes B_2) := A_1 B_1 \otimes A_2 B_2 \qquad (A_1, B_1 \in \mathcal{A}_1,\ A_2, B_2 \in \mathcal{A}_2). \tag{5.1}$$

Exercise 5.1.9 Show that formula (5.1) defines (uniquely!) a multiplication on $\mathcal{A}_1 \otimes \mathcal{A}_2$, and that $\mathcal{A}_1 \otimes \mathcal{A}_2$, with this multiplication, is an algebra.

If $\mathcal{V}_1, \mathcal{V}_2$ are linear spaces and $A_1 \in \mathcal{L}(\mathcal{V}_1), A_2 \in \mathcal{L}(\mathcal{V}_2)$ are linear operators, then we define $A_1 \otimes A_2 \in \mathcal{L}(\mathcal{V}_1 \otimes \mathcal{V}_2)$ by

$$(A_1 \otimes A_2)(\phi_1 \otimes \phi_2) := (A_1\phi_1) \otimes (A_2\phi_2) \qquad (\phi_1 \in \mathcal{V}_1,\ \phi_2 \in \mathcal{V}_2).$$

(Note that by property (iv) of Proposition 5.1.1, this defines a unique linear operator on $\mathcal{V}_1 \otimes \mathcal{V}_2$.) The next lemma shows that this is good notation.

Lemma 5.1.10 (Tensor product of factor algebras) *The algebra $\mathcal{L}(\mathcal{V}_1 \otimes \mathcal{V}_2)$ together with the bilinear map $(A_1, A_2) \mapsto A_1 \otimes A_2$ is a version of the tensor product of algebras $\mathcal{L}(\mathcal{V}_1) \otimes \mathcal{L}(\mathcal{V}_2)$.*

Proof Choose bases $\{\phi^1_1, \ldots, \phi^1_{d_1}\}$ and $\{\phi^2_1, \ldots, \phi^2_{d_1}\}$ for $\mathcal{V}_1$ and $\mathcal{V}_2$, respectively. Then a basis for $\mathcal{L}(\mathcal{V}_1)$ is formed by all linear operators P^1_{ij} $(1 \leq i, j \leq d_1)$ with

$$P^1_{ij}\phi_k := \delta_{jk}\phi^1_i.$$

Likewise, a basis for $\mathcal{L}(\mathcal{V}_2)$ is formed by all operators P^2_{lm} $(1 \leq l, m \leq d_2)$. Now the operators

$$(P^1_{ij} \otimes P^2_{lm})(\phi^1_k \otimes \phi^2_n) = \delta_{jk}\delta_{mn}(\phi^1_i \otimes \phi^2_l)$$

form a basis for $\mathcal{L}(\mathcal{V}_1 \otimes \mathcal{V}_2)$. This shows that $\mathcal{L}(\mathcal{V}_1 \otimes \mathcal{V}_2)$ is a version of $\mathcal{L}(\mathcal{V}_1) \otimes \mathcal{L}(\mathcal{V}_2)$ in the sense of linear spaces. It is trivial to check (5.1) so $\mathcal{L}(\mathcal{V}_1 \otimes \mathcal{V}_2)$ is also a version of $\mathcal{L}(\mathcal{V}_1) \otimes \mathcal{L}(\mathcal{V}_2)$ in the sense of algebras. ∎

If $\mathcal{H}_1, \mathcal{H}_2$ are inner product spaces and $\mathcal{H}_1 \otimes \mathcal{H}_2$ is their tensor product in the sense of linear spaces, then we equip $\mathcal{H}_1 \otimes \mathcal{H}_2$ with an inner product by putting

$$\langle \phi_1 \otimes \phi_2 | \psi_1 \otimes \psi_2 \rangle := \langle \phi_1 | \psi_1 \rangle_1 \langle \phi_2 | \psi_2 \rangle_2, \tag{5.2}$$

where $\langle \cdot | \cdot \rangle_1$ and $\langle \cdot | \cdot \rangle_2$ denote the inner products on $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively.

Exercise 5.1.11 Check that (5.2) defines a unique inner product on $\mathcal{H}_1 \otimes \mathcal{H}_2$.

If $\mathcal{A}_1, \mathcal{A}_2$ are *-algebras and $\mathcal{A}_1 \otimes \mathcal{A}_2$ is their tensor product in the sense of algebras then we equip $\mathcal{A}_1 \otimes \mathcal{A}_2$ with the adjoint operation

$$(A_1 \otimes A_2)^* := A_1^* \otimes A_2^* \qquad (A_1 \in \mathcal{A}_1,\ A_2 \in \mathcal{A}_2). \tag{5.3}$$

Exercise 5.1.12 Check that (5.3) defines a unique adjoint operation on $\mathcal{A}_1 \otimes \mathcal{A}_2$. Show that $\mathcal{A}_1 \otimes \mathcal{A}_2$ is a positive *-algebra if $\mathcal{A}_1$ and $\mathcal{A}_2$ are positive *-algebras.

Lemma 5.1.13 (Tensor product of factor positive *-algebras) *The positive *-algebra $\mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ together with the bilinear map $(A_1, A_2) \mapsto A_1 \otimes A_2$ is a version of the tensor product of *-algebras $\mathcal{L}(\mathcal{H}_1) \otimes \mathcal{L}(\mathcal{H}_2)$.*

Proof By Lemma 5.1.10 it suffices to shows that (5.3) holds. Indeed

$$\begin{aligned}
&\langle \phi_1 \otimes \phi_2 | (A_1 \otimes A_2) \psi_1 \otimes \psi_2 \rangle = \langle \phi_1 \otimes \phi_2 | (A_1 \psi_1) \otimes (A_2 \psi_2) \rangle \\
&\quad = \langle \phi_1 | A_1 \psi_1 \rangle_1 \langle \phi_2 | A_2 \psi_2 \rangle_2 = \langle A_1^* \phi_1 | \psi_1 \rangle_1 \langle A_2^* \phi_2 | \psi_2 \rangle_2 \\
&\quad = \langle (A_1^* \phi_1) \otimes (A_2^* \phi_2) | \psi_1 \otimes \psi_2 \rangle = \langle (A_1^* \otimes A_2^*) \phi_1 \otimes \phi_2 | \psi_1 \otimes \psi_2 \rangle
\end{aligned}$$

for all $\phi_1, \psi_1 \in \mathcal{H}_1$ and $\phi_2, \psi_2 \in \mathcal{H}_2$. By linearity and colinearity it follows that

$$\langle \phi | (A_1 \otimes A_2) \psi \rangle = \langle (A_1^* \otimes A_2^*) \phi | \psi \rangle \qquad (\phi, \psi \in \mathcal{H}_1 \otimes \mathcal{H}_2),$$

which implies (5.3). ∎

If $\mathcal{V}_1, \mathcal{V}_2$ are linear spaces and $\{\phi^2_1, \ldots, \phi^2_{d_2}\}$ is a basis for $\mathcal{V}_2$, then

$$\mathcal{V}_1 \otimes \mathcal{V}_2 \cong (\mathcal{V}_1 \otimes \phi^2_1) \oplus \cdots \oplus (\mathcal{V}_1 \otimes \phi^2_{d_2}),$$

where

$$\mathcal{V}_1 \otimes \phi_1^2 := \{\psi_1 \otimes \phi_1^2 : \psi_1 \in \mathcal{V}_1\} \cong \mathcal{V}_1.$$

Likewise, if $\mathcal{H}_1, \mathcal{H}_2$ are inner product spaces and $\{\phi_1^2, \ldots, \phi_{d_2}^2\}$ is an *orthogonal* basis for $\mathcal{H}_2$, then we have the decomposition of $\mathcal{H}_1 \otimes \mathcal{H}_2$ into *orthogonal* subspaces

$$\mathcal{H}_1 \otimes \mathcal{H}_2 \cong (\mathcal{H}_1 \otimes \phi_1^2) \oplus \cdots \oplus (\mathcal{H}_1 \otimes \phi_{d_2}^2),$$

where

$$\mathcal{H}_1 \otimes \phi_1^2 := \{\psi_1 \otimes \phi_1^2 : \psi_1 \in \mathcal{H}_1\} \cong \mathcal{H}_1.$$

With these observations, we can reformulate Theorem 3.8.1 in the following way:

Theorem 5.1.14 (Form of a general representation) *Let $\mathcal{A}$ be a positive *-algebra represented on an inner product space $\mathcal{H}$. Then there exist inner product spaces $\mathcal{H}_1, \ldots, \mathcal{H}_n$ and $\mathcal{K}_1, \ldots, \mathcal{K}_n$ such that one has the isomorphism of inner product spaces*

$$\mathcal{H} \cong (\mathcal{H}_1 \otimes \mathcal{K}_1) \oplus \cdots \oplus (\mathcal{H}_n \otimes \mathcal{K}_n),$$

and $\mathcal{A}$ consists of all operators of the form

$$(A_1 \otimes 1) \oplus \cdots \oplus (A_n \otimes 1)$$

with $A_1 \in \mathcal{L}(\mathcal{H}_1), \ldots, A_n \in \mathcal{L}(\mathcal{H}_n)$.

Note that $\mathcal{H}_1, \ldots, \mathcal{H}_n$ are the irreducible *-representations of $\mathcal{A}$ and $\dim(\mathcal{K}_i) = m_i$ is the multiplicity of $\mathcal{H}_i$ in $\mathcal{H}$. Every choice of an orthogonal basis for each of the spaces $\mathcal{K}_1, \ldots, \mathcal{K}_n$ leads to a different decomposition of $\mathcal{H}$ as in Theorem 3.8.1.

5.2 Independence

Let $\mathcal{A}$ be a positive *-algebra. By definition, in these lecture notes, a *sub-1-algebra* of $\mathcal{A}$ is a sub-*-algebra $\mathcal{B}$ of $\mathcal{A}$ containing the identity from $\mathcal{A}$. We interpret $\mathcal{B}$ as a subsystem of $\mathcal{A}$. For example, $\mathcal{A}$ may be an algebra describing two particles, and $\mathcal{B}$ may be the subalgebra containing all observables that say something about the first particle only. Or $\mathcal{B}$ may be the algebra containing all observables that say something about certain aspects of a particle, such as its position, but neglect other aspects, such as its spin. In a classical setting, our system might consist of a deck of cards, and $\mathcal{B}$ might consist of all functions that depend only on the color of the cards, but not on their values. Note that in all these cases it is natural to require that $\mathcal{B}$ contains the trivial observation 1 from $\mathcal{A}$. If ρ is a probability on $\mathcal{A}$, describing our knowledge about the system $\mathcal{A}$, then

$$\rho\big|_{\mathcal{B}},$$

the restriction of ρ to $\mathcal{B}$, is a probability on $\mathcal{B}$ describing our knowledge about the subsystem $\mathcal{B}$.

Proposition 5.2.1 (Commuting subalgebras) *Let $\mathcal{A}$ be a positive *-algebra and let $\mathcal{B}_1, \mathcal{B}_2$ be sub-1-algebras of $\mathcal{A}$. Then the following are equivalent:*

$$\begin{array}{rll} \text{(i)} & \sum_{i=1}^n \rho(P_{2,i}P_1P_{2,i}) = \rho(P_1) & \forall P_1 \in \mathcal{B}_1 \text{ projection},\ \{P_{2,1},\ldots,P_{2,n}\} \subset \mathcal{B}_2 \\ & & \quad \text{partition of the identity},\ \rho \text{ probability} \\ & & \quad \text{on } \mathcal{A}, \\ \text{(ii)} & P_1P_2 = P_2P_1 & \forall P_1 \in \mathcal{B}_1,\ P_2 \in \mathcal{B}_2,\ P_1, P_2 \text{ projections}, \\ \text{(iii)} & B_1B_2 = B_2B_1 & \forall B_1 \in \mathcal{B}_1,\ B_2 \in \mathcal{B}_2. \end{array}$$

Note that (i) says that if we know that someone performs an ideal measurement on the subsystem $\mathcal{B}_2$ (but we don't know the outcome), then this does not change our knowledge about the subsystem $\mathcal{B}_1$. Condition (iii) says that every operator from $\mathcal{B}_1$ commutes with every operator from $\mathcal{B}_2$. In this case we say that $\mathcal{B}_1$ and $\mathcal{B}_2$ *commute*.

Proof of Proposition 5.2.1 (i)$\Rightarrow$(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 $\mathcal{A}$

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

This holds for all ρ if and only if[1]

$$2P_2P_1P_2 = P_1P_2 + P_2P_1.$$

Represent $\mathcal{A}$ on some inner product space $\mathcal{H}$. Then there exists subspaces $\mathcal{H}_1, \mathcal{H}_2$ such that P_1, P_2 are the orthogonal projections on $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. For any $\phi_2 \in \mathcal{H}_2$, one has

$$P_1\phi_2 = P_1P_2\phi_2 = P_2P_1\phi_2 - 2P_2P_1P_2\phi_2 \in \mathcal{H}_2.$$

[1] To see this, note that by Exercise 4.2.3, every linear form can be written as a linear combination of (at most four) probabilities. Therefore, by linearity, $2l(P_2P_1P_2) = l(P_1P_2) + l(P_2P_1)$ for every $l \in \mathcal{A}'$, which implies $2P_2P_1P_2 = P_1P_2 + P_2P_1$.

This implies that $P_2P_1P_2 = P_1P_2$ and therefore

$$2P_1P_2 = P_1P_2 + P_2P_1,$$

which implies $P_1P_2 = P_2P_1$.

(ii)⇒(iii): If B_1 and B_2 are self-adjoint, then we can write $B_1 = \sum_{i=1}^{n_1} \lambda_i P_{1,i}$ and $B_2 = \sum_{j=1}^{n_2} \mu_j P_{2,j}$ where λ_i, μ_j are real numbers and $P_{1,i}, P_{2,j}$ are projections from $\mathcal{B}_1$ and $\mathcal{B}_2$, respectively, so the statement follows by linearity. To prove the general case we use that every operator $B_1 \in \mathcal{B}_1$ can be written as $B_1 = \mathrm{Re}(B_1) + i\mathrm{Im}(B_1)$ with $\mathrm{Re}(B_1), \mathrm{Im}(B_1) \in \mathcal{B}_1$ self-adjoint, and likewise for any $B_2 \in \mathcal{B}_2$ (see Exercise 1.3.4), and we again use linearity.

(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 $\{P_{2,1}, \ldots, P_{2,n}\} \subset \mathcal{B}_2$. ∎

Recall the definition of $\alpha(\mathcal{C})$, the sub-1-algebra generated by a set $\mathcal{C} \subset \mathcal{A}$, from page 61. If $\mathcal{B}_1, \mathcal{B}_2$ are commuting sub-1-algebras of some larger positive *-algebra, then

$$\alpha(\mathcal{B}_1 \cup \mathcal{B}_2) = \mathrm{span}\{B_1B_2 : B_1 \in \mathcal{B}_1,\ B_2 \in \mathcal{B}_2\}.$$

is the smallest *-algebra containing $\mathcal{B}_1$ and $\mathcal{B}_2$. This formula does not hold in general if $\mathcal{B}_1$ and $\mathcal{B}_2$ do not commute!

Proposition 5.2.2 (Logically independent algebras) *Let $\mathcal{B}_1, \mathcal{B}_2$ be sub-1-algebras of some larger positive *-algebra, and assume that $\mathcal{B}_1$ and $\mathcal{B}_2$ commute. Then the following 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 probabilities ρ_1 on $\mathcal{B}_1$ and ρ_2 on $\mathcal{B}_2$ there exists a unique probability 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 $B_1 \in \mathcal{B}_1,\ B_2 \in \mathcal{B}_2$.*

(iii) *$\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$ together with the bilinear map $(B_1, B_2) \mapsto B_1B_2$ is a version of the tensor product $\mathcal{B}_1 \otimes \mathcal{B}_2$.*

By definition, in these lecture notes, we will say that two sub-1-algebras $\mathcal{B}_1, \mathcal{B}_2$ of some larger positive *-algebra $\mathcal{A}$ are *logically independent* if $\mathcal{B}_1$ and $\mathcal{B}_2$ commute and they satisfy the equivalent properties (i)–(iii) from Propsition 5.2.2. In classical probability, property (i) is sometimes called 'qualitative independence' [Ren70]. Property (i) 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 $\mathcal{B}_2$.

Proof of Proposition 5.2.2 (i)$\Rightarrow$(ii): We first prove the statement when ρ_1 and ρ_2 are pure states. In that case, by Lemma 4.2.5, there exists a projection $P_1 \in \mathcal{B}_1$ that is minimal in $\mathcal{B}_1$ (i.e., the only projections Q with $Q \leq P_1$ *and* $Q \in \mathcal{B}_1$ are $Q = 0$ and $Q = P_1$) and a projection $P_2 \in \mathcal{B}_2$ that is minimal in $\mathcal{B}_2$ such that $\rho_1 = \rho_{P_1}$ and $\rho_2 = \rho_{P_2}$. 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

$$(P_1P_2)(B_1B_2)(P_1P_2) = P_1B_1P_1P_2B_2P_2 = \rho_1(B_1)\rho_2(B_2)P_1P_2 \quad (B_1 \in \mathcal{B}_1,\ B_2 \in \mathcal{B}_2).$$

Therefore, by linearity, there exists a function $\rho : \alpha(\mathcal{B}_1 \cup \mathcal{B}_2) \to \mathbb{C}$ such that

$$(P_1P_2)A(P_1P_2) = P_1B_1P_1P_2B_2P_2 = \rho(A)P_1P_2 \qquad (A \in \alpha(\mathcal{B}_1 \cup \mathcal{B}_2)).$$

We claim that P_1P_2 is minimal in $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$. By assumption, $P_1P_2 \neq 0$. Now assume that $Q \in \alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$ is a projection such that $Q \leq P_1P_2$. Then $Q = (P_1P_2)Q(P_1P_2) = \rho(Q)P_1P_2$, which shows that P_1P_2 is minimal. It follows that the funtion ρ we have just introduced is the pure state $\rho_{P_1P_2}$ and

$$\rho(B_1B_2) = \rho_1(B_1)\rho_2(B_2) \qquad (B_1 \in \mathcal{B}_1,\ B_2 \in \mathcal{B}_2).$$

Since operators of the form B_1B_2 span $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$, by linearity, ρ is the only state with this property.

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 $B_1 \in \mathcal{B}_1,\ B_2 \in \mathcal{B}_2$. We put

$$\rho := \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} p_iq_j\rho_{ij}.$$

Then

$$\begin{aligned}\rho(B_1B_2) &= \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} p_iq_j\rho_{1,i}(B_1)\rho_{2,j}(B_2)\\ &= \Big(\sum_{i=1}^{n_1} p_i\rho_{1,i}(B_1)\Big)\Big(\sum_{j=1}^{n_2} q_j\rho_{2,j}(B_2)\Big) = \rho_1(B_1)\rho_2(B_2),\end{aligned}$$

as required. Uniqueness of ρ follows in the same way as before.

(ii)⇒(iii): Using Exercise 4.2.3 it is easy to see that for any linear forms l_1 on $\mathcal{B}_1$ and l_2 on $\mathcal{B}_2$ 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 $B_1 \in \mathcal{B}_1$ and $B_2 \in \mathcal{B}_2$. Therefore, by Proposition 5.1.1, property (iii), $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$ together with the bilinear map $(B_1, B_2) \mapsto B_1 B_2$ is a version of the tensor product $\mathcal{B}_1 \otimes \mathcal{B}_2$ in the sense of linear spaces. Since $\mathcal{B}_1$ and $\mathcal{B}_2$ commute,

$$(B_1 B_2)(C_1 C_2) = (B_1 C_1)(B_2 C_2) \qquad (B_1, C_1 \in \mathcal{B}_1,\ B_2, C_2 \in \mathcal{B}_2)$$

and

$$(B_1 B_2)^* = (B_2 B_1)^* = B_1^* B_2^*,$$

so $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)$ is a version of the tensor product $\mathcal{B}_1 \otimes \mathcal{B}_2$ in the sense of *-algebras.

(iii)⇒(i): Obvious. ∎

Lemma 5.2.3 (Independent algebras) *Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be logically independent sub-1-algebras of some larger positive *-algebra $\mathcal{A}$ and let ρ be a probability on $\mathcal{A}$. Then the following are equivalent:*

(i) $\rho(P_2 B_1 P_2)/\rho(P_2) = \rho(B_1)$ *for every $B_1 \in \mathcal{B}_1$ and projection $P_2 \in \mathcal{B}_2$ with $\rho(P_2) \neq 0$.*

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

Property (i) says that doing an observation on the subsystem $\mathcal{B}_2$ does not change our knowledge about the subsystem $\mathcal{B}_1$. By definition, we say that two sub-1-algebras $\mathcal{B}_1$ and $\mathcal{B}_2$ of some larger positive *-algebra are *independent* under a probability ρ if $\mathcal{B}_1$ and $\mathcal{B}_2$ are logically independent and they satisfy the equivalent properties (i) and (ii) from Lemma 5.2.3. Note that in this case

$$\rho\big|_{\alpha(\mathcal{B}_1 \cup \mathcal{B}_2)} = \rho\big|_{\mathcal{B}_1} \otimes \rho\big|_{\mathcal{B}_2},$$

i.e., the probability on $\alpha(\mathcal{B}_1 \cup \mathcal{B}_2) \cong \mathcal{B}_1 \otimes \mathcal{B}_2$ is the *product probability* of its *marginals* $\rho\big|_{\mathcal{B}_1}$ and $\rho\big|_{\mathcal{B}_2}$.

Proof of Lemma 5.2.3 Since $\mathcal{B}_1$ and $\mathcal{B}_2$ commute, $\rho(P_2 B_1 P_2) = \rho(B_1 P_2 P_2) = \rho(B_1 P_2)$, so (i) is equivalent to

$$\rho(B_1 P_2) = \rho(B_1)\rho(P_2) \quad \forall B_1 \in \mathcal{B}_1,\ P_2 \in \mathcal{B}_2 \text{ nonzero projection.}$$

Since $\mathcal{B}_2$ is spanned by its projections, by linearity, this is equivalent to (ii). ∎

(Logical) independence of three or more algebras is defined in the obvious way, with tensor products of three or more algebras and products of three or more probabilities. We remind the reader of the fact that already in classical probability, pairwise independence does not imply independence, i.e., it is possible to find a probability ρ on a positive *-algebra of the form $\mathcal{B}_1 \otimes \mathcal{B}_2 \otimes \mathcal{B}_3$ such that $\mathcal{B}_1$ and $\mathcal{B}_2$ are independent, $\mathcal{B}_2$ and $\mathcal{B}_3$ are independent, and $\mathcal{B}_3$ and $\mathcal{B}_1$ are independent, but ρ is not of the form $\rho_1 \otimes \rho_2 \otimes \rho_3$, i.e., $\mathcal{B}_1, \mathcal{B}_2$, and $\mathcal{B}_3$ are not independent.

Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be positive *-algebras and let $\mathcal{A}_1 \otimes \mathcal{A}_2$ be their tensor product. We ask ourselves the following question: is every probability ρ on $\mathcal{A}_1 \otimes \mathcal{A}_2$ a convex combination of product probabilities, i.e., can ρ be written as

$$\rho = \sum_{i=1}^{n} p_i \rho_i^1 \otimes \rho_i^2$$

with ρ_i^1, ρ_i^2 probabilities on $\mathcal{A}_1$ and $\mathcal{A}_2$ and p_i nonnegative numbers summing up to one? It turns out that the answer is negative if and only if $\mathcal{A}_1$ and $\mathcal{A}_2$ are both nonabelian. Here is the simplest possible counterexample: let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two-dimensional inner product spaces with orthonormal bases $\{\phi_1^1, \phi_2^1\}$ and $\{\phi_1^2, \phi_2^2\}$, respectively. Put $\mathcal{A}_1 := \mathcal{L}(\mathcal{H}_1)$, $\mathcal{A}_2 := \mathcal{L}(\mathcal{H}_2)$, and $\mathcal{A}_1 \otimes \mathcal{A}_2 = \mathcal{L}(\mathcal{H}_1) \otimes \mathcal{L}(\mathcal{H}_2) \cong \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ (recall Lemma 5.1.13). Define a vector ψ in $\mathcal{H}_1 \otimes \mathcal{H}_2$ by

$$\psi := \tfrac{1}{\sqrt{2}}\, \phi_1^1 \otimes \phi_1^2 + \tfrac{1}{\sqrt{2}}\, \phi_2^1 \otimes \phi_2^2.$$

It is easy to check that $\|\psi\| = 1$ so

$$\rho(A) := \rho_\psi(A) = \langle \psi | A | \psi \rangle \qquad (A \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2))$$

defines a pure state on $\mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2) \cong \mathcal{A}_1 \otimes \mathcal{A}_2$. We claim that ρ cannot be written as a convex combination of product probabilities. Indeed, since ρ is a pure state, ρ cannot be written as a nontrivial convex combination of other probabilities. Therefore, it suffices to show that ρ is itself not a product probability. Consider the marginals of ρ, which are the probabilities ρ_1 and ρ_2 on $\mathcal{A}_1$ and $\mathcal{A}_2$, respectively, defined by

$$\begin{aligned} \rho_1(A_1) &:= \rho(A_1 \otimes 1) \qquad (A_1 \in \mathcal{A}_1), \\ \rho_2(A_2) &:= \rho(1 \otimes A_2) \qquad (A_2 \in \mathcal{A}_2). \end{aligned}$$

We see that

$$\begin{aligned} \rho_1(A_1) &= \langle \psi | A_1 \otimes 1 | \psi \rangle \\ &= \tfrac{1}{2}\langle \phi_1^1 \otimes \phi_1^2 | A_1 \otimes 1 | \phi_1^1 \otimes \phi_1^2 \rangle + \tfrac{1}{2}\langle \phi_1^1 \otimes \phi_1^2 | A_1 \otimes 1 | \phi_2^1 \otimes \phi_2^2 \rangle \\ &\quad + \tfrac{1}{2}\langle \phi_2^1 \otimes \phi_2^2 | A_1 \otimes 1 | \phi_1^1 \otimes \phi_1^2 \rangle + \tfrac{1}{2}\langle \phi_2^1 \otimes \phi_2^2 | A_1 \otimes 1 | \phi_2^1 \otimes \phi_2^2 \rangle \end{aligned}$$

$$
\begin{aligned}
&= \tfrac{1}{2}\langle\phi_1^1|A_1|\phi_1^1\rangle\langle\phi_1^2|\phi_1^2\rangle + \tfrac{1}{2}\langle\phi_1^1|A_1|\phi_2^1\rangle\langle\phi_1^2|\phi_2^2\rangle \\
&\quad +\tfrac{1}{2}\langle\phi_2^1|A_1|\phi_1^1\rangle\langle\phi_2^2|\phi_1^2\rangle + \tfrac{1}{2}\langle\phi_2^1|A_1|\phi_2^1\rangle\langle\phi_2^2|\phi_2^2\rangle \\
&= \tfrac{1}{2}\langle\phi_1^1|A_1|\phi_1^1\rangle + \tfrac{1}{2}\langle\phi_2^1|A_1|\phi_2^1\rangle = \tfrac{1}{2}\mathrm{tr}(A_1),
\end{aligned}
$$

and likewise $\rho_2(A_2) = \frac{1}{2}\mathrm{tr}(A_2)$ for all $A_2 \in \mathcal{A}_2$. In particular, ρ_1 and ρ_2 are not pure states! It is easy to see that

$$\rho_1 \otimes \rho_2(A) = \tfrac{1}{4}\mathrm{tr}(A) \qquad (A \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2),$$

and $\frac{1}{4}\mathrm{tr} \neq \rho$, which proves that ρ is not a product probability.

By definition, a probability on $\mathcal{A}_1 \otimes \mathcal{A}_2$ that is not a convex combination of product probabilities is called an *entangled state*.[2] Entanglement is a typical phenomenon from the quantum world, and has been the subject of a lot of research. As we will see in Section 6.2, entanglement can be measured, demonstrating quite convincingly that quantum probability is real.

Exercise 5.2.4 Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be positive *-algebras. Show that there exist entangled states on $\mathcal{A}_1 \otimes \mathcal{A}_2$ if and only if $\mathcal{A}_1$ and $\mathcal{A}_2$ are *both* nonabelian.

5.3 Completely positive maps

For any positive *-algebra $\mathcal{A}$, we write

$$
\begin{aligned}
\mathcal{A}'_r &:= \{l : l \text{ real linear form on } \mathcal{A}\}, \\
\mathcal{A}'_+ &:= \{l : l \text{ positive real linear form on } \mathcal{A}\}, \\
\mathcal{A}'_{\mathrm{prob}} &:= \{\rho : \rho \text{ probability on } \mathcal{A}\}.
\end{aligned}
$$

We will be interested in operations on $\mathcal{A}$, i.e., the effects of our physical interference with the system $\mathcal{A}$, which changes the probability ρ on $\mathcal{A}$ into another probability. Thus, we will be interested in maps $f : \mathcal{A}'_{\mathrm{prob}} \to \mathcal{A}'_{\mathrm{prob}}$ that take a probability ρ on $\mathcal{A}$ and map it into a new probability $f(\rho)$ on $\mathcal{A}$. A natural requirement on f is that it be affine, i.e.,

$$f(p\rho_1 + (1-p)\rho_2) = pf(\rho_1) + (1-p)\rho_2 \qquad (\rho_1, \rho_2 \in \mathcal{A}'_{\mathrm{prob}},\ 0 \leq p \leq 1).$$

This says that if before we performed our operation f, our knowledge about the system is with probability p described by ρ_1 and with probability $1-p$ by ρ_2,

[2] The word entanglement is a translation of the German ‘Verschränktheit’, coined in a paper by Schrödinger from 1935 [Sch35].

then after we have performed the operation f, the system is with probability p described by $f(\rho_1)$ and with probability $1-p$ by $f(\rho_2)$.

At first, it might seem that nothing more can be said about f and that any affine map $f : \mathcal{A}'_{\rm prob} \to \mathcal{A}'_{\rm prob}$ describes a legal operation on the system $\mathcal{A}$. However, it turns out that this is not the case. Assume that $\mathcal{B}$ is some other system, logically independent of $\mathcal{A}$. (Since our algebra $\mathcal{A}$ will typically not describe the whole universe, there will typically be lots of such systems!) Then we must be able to say what happens with a probability ρ on $\mathcal{A}\otimes\mathcal{B}$ when we perform our operation on $\mathcal{A}$ and do nothing with $\mathcal{B}$. For product probabilities, it is natural to require that the effect of our operation on $\mathcal{A}$, doing nothing with $\mathcal{B}$, is to map $\rho_1\otimes\rho_2$ to

$$F(\rho_1\otimes\rho_2) := f(\rho_1)\otimes\rho_2.$$

Now we need to require from f that F can be extended to an affine map $F : (\mathcal{A}\otimes\mathcal{B})'_{\rm prob} \to (\mathcal{A}\otimes\mathcal{B})'_{\rm prob}$. We will see that this is a nonempty requirement on f! As one might guess, problems occur for entangled probabilities.

Slightly generalizing our set-up, let us consider affine maps $f : \mathcal{A}'_{\rm prob} \to \mathcal{B}'_{\rm prob}$ from the probabilities on one positive *-algebra $\mathcal{A}$ into the probabilities on another positive *-algebra $\mathcal{B}$.

Lemma 5.3.1 (Affine maps on probabilities) *Let $\mathcal{A}$ and $\mathcal{B}$ be positive *-algebras and let $T : \mathcal{B}\to\mathcal{A}$ be a linear map satisfying*

$$\begin{array}{rl}\text{(i)} & T(B^*) = T(B)^*,\\ \text{(ii)} & B\geq 0 \;\Rightarrow\; T(B)\geq 0,\\ \text{(iii)} & T(1) = 1.\end{array} \tag{5.4}$$

Let $T^* : \mathcal{A}'\to\mathcal{B}'$ be the adjoint of T, i.e.,

$$T^*(l)(A) = l(T(B))) \qquad (l\in\mathcal{A}',\ B\in\mathcal{B}).$$

Then the restriction of T^* to $\mathcal{A}'_{\rm prob}$ is an affine map from $\mathcal{A}'_{\rm prob}$ into $\mathcal{B}'_{\rm prob}$, and conversely, every affine map from $\mathcal{A}'_{\rm prob}$ into $\mathcal{B}'_{\rm prob}$ arises in this way.

Proof Assume that $f : \mathcal{A}'_{\rm prob}\to\mathcal{B}'_{\rm prob}$ is affine. Since $\mathcal{A}'_{\rm prob}$ is a convex subset of $\mathcal{A}'$ not containing the origin and $\mathcal{A}'_{\rm prob}$ spans $\mathcal{A}'$ (see Exercise 4.2.3), f can be extended to a unique linear map $\hat f : \mathcal{A}'\to\mathcal{B}'$. It is not hard to see that $\hat f$ satisfies

$$\begin{array}{rl}\text{(i)} & l \text{ real} \Rightarrow \hat f(l) \text{ real},\\ \text{(ii)} & l \text{ positive} \Rightarrow \hat f(l) \text{ positive},\\ \text{(iii)} & \hat f(l)(1) = l(1),\end{array} \tag{5.5}$$

and conversely, if a linear map $\hat{f} : \mathcal{A}' \to \mathcal{B}'$ satisfies (5.5) then $\hat{f}$ maps $\mathcal{A}'_{\rm prob}$ into $\mathcal{B}'_{\rm prob}$ and is affine on $\mathcal{A}'_{\rm prob}$. Let T denote the adjoint of $\hat{f}$ in the sense of page 25, i.e., $T : \mathcal{B} \to \mathcal{A}$ is the unique linear map such that

$$\hat{f}(l)(B) = l(T(B)) \qquad (l \in \mathcal{A}',\ B \in \mathcal{B}).$$

Then it is not hard to see that the conditions (5.4) (i)–(iii) are equivalent to the conditions (5.5) (i)–(iii). ∎

Exercise 5.3.2 Check that (5.4) is equivalent to (5.5) if $\hat{f} = T^*$, the adjoint of T.

A linear map $T : \mathcal{B} \to \mathcal{A}$ satisfying (5.4) (i) and (ii) is called *positive*.

Now let $\mathcal{A}, \mathcal{B}$ be positive *-algebras and let $f : \mathcal{A}'_{\rm prob} \to \mathcal{B}'_{\rm prob}$ be affine. We want to know when it is true that for any positive *-algebra $\mathcal{C}$, there exists an affine map $F : (\mathcal{A} \otimes \mathcal{C})'_{\rm prob} \to (\mathcal{B} \otimes \mathcal{C})'_{\rm prob}$ such that

$$F(\rho_1 \otimes \rho_2) := f(\rho_1) \otimes \rho_2 \qquad (\rho_1 \in \mathcal{A}'_{\rm prob},\ \rho_2 \in \mathcal{C}'_{\rm prob}). \tag{5.6}$$

Since any positive *-algebra is isomorphic to a sub-*-algebra of some $\mathcal{L}(\mathcal{K})$, it suffices to consider the case that $\mathcal{C} = \mathcal{L}(\mathcal{K})$ for some inner product space $\mathcal{K}$. By Lemma 5.3.1, there is a positive linear map $T : \mathcal{B} \to \mathcal{A}$ satisfying $T(1) = 1$ such that $f = T^*$. Likewise, if there exists a map F satisfying (5.6) then F must be of the form $F = S^*$ for some positive linear map $S : \mathcal{B} \otimes \mathcal{L}(\mathcal{K}) \to \mathcal{A} \otimes \mathcal{L}(\mathcal{K})$ satisfying $S(1) = 1$. Now (5.6) implies that

$$S^* = T^* \otimes 1 = (T \otimes 1)^*,$$

i.e., we must have $S = (T \otimes 1)$. This leads to the following definition: We say that a linear map $T : \mathcal{B} \to \mathcal{A}$ is *completely positive* if the map

$$T \otimes 1 : \mathcal{B} \otimes \mathcal{L}(\mathcal{K}) \to \mathcal{A} \otimes \mathcal{L}(\mathcal{K})$$

is positive for every inner product space $\mathcal{K}$. Surprisingly, we will see below that positivity does not imply complete positivity.

By definition, an *operation* from a positive-*-algebra $\mathcal{A}$ to another positive-*-algebra $\mathcal{B}$ is a map of the form $T^* : \mathcal{A}' \to \mathcal{B}'$, where $T : \mathcal{B} \to \mathcal{A}$ is a completely positive linear map satisfying $T(1) = 1$. Operations correspond to things that one can do with a physical system, changing its state, and possibly even the algebra needed to describe the system.

The next lemma is almost a direct consequence of the way complete positivity has been defined, so we skip the proof.

Lemma 5.3.3 (Operations on logically independent algebras) *Let $\mathcal{A},\mathcal{B}$ and $\mathcal{C}$ be positive *-algebras and let $T^*:\mathcal{A}'\to\mathcal{B}'$ be an operation. Then there exists a unique operation $S^*:(\mathcal{A}\otimes\mathcal{C})'\to(\mathcal{B}\otimes\mathcal{C})'$ such that*

$$S^*(\rho_1\otimes\rho_2)=(T^*\rho_1)\otimes\rho_2 \qquad (\rho_1\in\mathcal{A}'_{\mathrm{prob}},\ \rho_2\in\mathcal{C}'_{\mathrm{prob}}).$$

This operation is given by $S^*=(T\otimes 1)^*$. If ρ is any probability on $\mathcal{A}\otimes\mathcal{C}$, then

$$\big((T\otimes 1)^*\rho\big)(1\otimes C)=\rho(1\otimes C) \qquad (C\in\mathcal{C}).$$

Note that the second formula says that if $\mathcal{A}$ and $\mathcal{C}$ are logically independent, then performing an operation T on $\mathcal{A}$ does not change our knowledge about $\mathcal{C}$.

We set out to show that positivity does not imply complete positivity and to characterise all completely positive maps. Before we show this, we first give a slightly different formulation of complete positivity, that is often useful. Let $\mathcal{A}$ be a positive *-algebra and let $\mathcal{K}$ be an inner product space. Let $\{\psi_1,\ldots,\psi_m\}$ be an orthonormal basis for $\mathcal{K}$. Then the linear operators

$$|\psi_i\rangle\langle\psi_j| \qquad (i,j=1,\ldots,m)$$

form a basis for $\mathcal{L}(\mathcal{K})$, and we can decompose $\mathcal{A}\otimes\mathcal{L}(\mathcal{K})$ as a linear space (see page 81) as

$$\mathcal{A}\otimes\mathcal{L}(\mathcal{K})\cong\bigoplus_{i,j=1,\ldots,m}\mathcal{A}\otimes|\psi_i\rangle\langle\psi_j|.$$

Thus, every $A\in\mathcal{A}\otimes\mathcal{L}(\mathcal{K})$ has a unique decomposition

$$A=\sum_{i,j=1}^{m}A_{ij}\otimes|\psi_i\rangle\langle\psi_j|,$$

with $A_{ij}\in\mathcal{A}$. Now

$$\Big(\sum_{ij}A_{ij}\otimes|\psi_i\rangle\langle\psi_j|\Big)\Big(\sum_{kl}B_{kl}\otimes|\psi_k\rangle\langle\psi_l|\Big)=\sum_{il}\Big(\sum_j A_{ij}B_{jl}\Big)|\psi_i\rangle\langle\psi_l,|$$

which shows that

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

Moreover,

$$\Big(\sum_{ij}A_{ij}\otimes|\psi_i\rangle\langle\psi_j|\Big)^*=\sum_{ij}A^*_{ij}\otimes|\psi_j\rangle\langle\psi_i|,$$

which shows that

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

Thus, we see that with respect to an orthonormal basis for $\mathcal{K}$ there is a natural isomorphism between the positive *-algebra $\mathcal{A} \otimes \mathcal{L}(\mathcal{K})$ and the space of $m \times m$ matrices with entries from $\mathcal{A}$. Now if $T : \mathcal{B} \to \mathcal{A}$ is a linear map, then the linear map $(T \otimes 1) : \mathcal{B} \otimes \mathcal{L}(\mathcal{K}) \to \mathcal{A} \otimes \mathcal{L}(\mathcal{K})$ satisfies

$$(T \otimes 1)\Big(\sum_{ij} B_{ij} \otimes |\psi_i\rangle\langle\psi_j|\Big) = \Big(\sum_{ij} T(B_{ij}) \otimes |\psi_i\rangle\langle\psi_j|\Big),$$

which shows that

$$\big((T \otimes 1)(B)\big)_{ij} = T(B_{ij}). \tag{5.7}$$

If $\mathcal{A}$ is a positive *-algebra then we let $M_m(\mathcal{A})$ denote the space of $m \times m$ matrices with entries from $\mathcal{A}$, equipped with the structure of a positive *-algebra by putting $(AB)_{ij} := \sum_k A_{ik}B_{kj}$ and $(A^*)_{ij} := (A_{ji})^*$.

Lemma 5.3.4 (Different formulation of complete positivity) *Let $\mathcal{A}, \mathcal{B}$ be positive *-algebra and let $T : \mathcal{B} \to \mathcal{A}$ be a linear map. Then T is completely positive if and only if the map from $M_m(\mathcal{B})$ to $M_m(\mathcal{A})$ given by*

$$\begin{pmatrix} B_{11} & \cdots & B_{1m} \\ \vdots & & \vdots \\ B_{m1} & \cdots & B_{mm} \end{pmatrix} \mapsto \begin{pmatrix} T(B_{11}) & \cdots & T(B_{1m}) \\ \vdots & & \vdots \\ T(B_{m1}) & \cdots & T(B_{mm}) \end{pmatrix}$$

is positive for each m.

Counterexample 5.3.5 (A positive map that is not completely positive) Let $\mathcal{H}$ be an inner product space of dimension 2, with orthonormal basis $\{\phi_1, \phi_2\}$. Then the linear map $T : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ given with respect to this basis by $T(A)_{ij} := A_{ji}$ is positive and satisfies $T(1) = 1$, but T is not completely positive.

Proof If A is positive then

$$\sum_{ij} \overline{x}_i A_{ij} x_j \geq 0 \qquad \forall x \in \mathcal{H}.$$

This implies (take $y_i := \overline{x}_i$)

$$\sum_{ij} \overline{y}_i T(A)_{ij} y_j = \sum_{ij} y_i A_{ij} \overline{y}_j \geq 0 \qquad \forall y \in \mathcal{H},$$

which shows that $T(A)$ is positive. However, the map

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \mapsto \begin{pmatrix} T(A_{11}) & T(A_{12}) \\ T(A_{21}) & T(A_{22}) \end{pmatrix}$$

is not positive, since under this map

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The matrix on the left is

$$\left| \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right|,$$

which is clearly a positive operator. On the other hand, the matrix on the right has an eigenvalue -1:

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

and is therefore not positive. ∎

Remark An operator $T : \mathcal{B} \to \mathcal{A}$ is called n-positive if the map from $M_n(\mathcal{A})$ to $M_n(\mathcal{B})$ in Lemma 5.3.4 is positive. One can show that $(n+1)$-positive $\Rightarrow$ n-positive but n-positive $\not\Rightarrow$ $(n+1)$-positive.

We set out to describe the general form of completely positive maps $T : \mathcal{B} \to \mathcal{A}$. We first treat the case that $\mathcal{A}$ and $\mathcal{B}$ are factor algebras. The next theorem, which describes completely positive maps between factor algebras, is due to Stinespring [Sti55]; see also [Tak79, Thm IV.3.6].

Theorem 5.3.6 (Stinespring) *Let $\mathcal{H}$ and $\mathcal{F}$ be inner product spaces and let $V(1), \ldots, V(n)$ be linear maps in $\mathcal{L}(\mathcal{H}, \mathcal{F})$. Then*

$$T(A) := \sum_{m=1}^{n} V(m) A V(m)^* \qquad (A \in \mathcal{L}(\mathcal{H}))$$

defines a completely positive linear map $T : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{F})$ and conversely, every completely positive linear map from $\mathcal{L}(\mathcal{H})$ to $\mathcal{L}(\mathcal{F})$ is of this form.

Proof This proof is best done with tensor calculus. We have $\mathcal{L}(\mathcal{H}) \cong \mathcal{H} \otimes \mathcal{H}'$ and $\mathcal{L}(\mathcal{F}) \cong \mathcal{F} \otimes \mathcal{F}'$. A linear map T from $\mathcal{L}(\mathcal{H})$ to $\mathcal{L}(\mathcal{F})$ is therefore an element of

$$\mathcal{L}(\mathcal{H} \otimes \mathcal{H}', \mathcal{F} \otimes \mathcal{F}) \cong \mathcal{F} \otimes \mathcal{F}' \otimes (\mathcal{H} \otimes \mathcal{H}')' \cong \mathcal{F} \otimes \mathcal{F}' \otimes \mathcal{H}' \otimes \mathcal{H}.$$

Thus, there is a tensor

$$\tilde{T} \in \mathcal{F} \otimes \mathcal{F}' \otimes \mathcal{H}' \otimes \mathcal{H}$$

such that with respect to bases for $\mathcal{F}, \mathcal{H}$ and the corresponding dual bases for $\mathcal{F}', \mathcal{H}'$, one has

$$(T(A))_{ij} = \sum_{kl} \tilde{T}_{ijkl} A_{kl}.$$

Note that $A \in \mathcal{L}(\mathcal{H}) \cong \mathcal{H} \otimes \mathcal{H}'$. We have contracted the third coordinate of $\tilde{T}$ with the first coordinate of A, which corresponds to contracting $\mathcal{H}'$ with $\mathcal{H}$, and the fourth coordinate of $\tilde{T}$ with the second coordinate of A, which corresponds to contracting $\mathcal{H}$ with $\mathcal{H}'$. In view of what follows, it will be convenient to order the the spaces in the tensor product $\mathcal{F} \otimes \mathcal{F}' \otimes \mathcal{H}' \otimes \mathcal{H}$ in a different way. Let

$$T \in \mathcal{F} \otimes \mathcal{H}' \otimes \mathcal{H} \otimes \mathcal{F}'$$

be the tensor defined by $T_{ijkl} = \tilde{T}_{iljk}$. Then

$$(T(A))_{ij} = \sum_{kl} T_{iklj} A_{kl}. \tag{5.8}$$

Now T is of the form $T(A) = \sum_m V(m) A V(m)^*$ if and only if

$$\sum_{kl} T_{iklj} A_{kl} = (T(A))_{ij} = \big(\sum_m V(m) A V(m)^*\big)_{ij} = \sum_m \sum_{kl} V_{ik}(m) A_{kl} \overline{V}_{jl}(m),$$

which is equivalent to

$$T_{iklj} = \sum_m V_{ik}(m) \overline{V}_{jl}(m). \tag{5.9}$$

We must now formulate complete positivity in the language of tensor calculus. If $\mathcal{K}$ is another inner product space then $\mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{K}) \cong \mathcal{L}(\mathcal{H} \otimes \mathcal{K})$ and

$$\mathcal{L}(\mathcal{H} \otimes \mathcal{K}) \cong (\mathcal{H} \otimes \mathcal{K}) \otimes (\mathcal{H} \otimes \mathcal{K})' \cong \mathcal{H} \otimes \mathcal{K} \otimes \mathcal{K}' \otimes \mathcal{H}'.$$

If we order the spaces $\mathcal{H}, \mathcal{K}, \mathcal{K}', \mathcal{H}'$ in the tensor product in this way, then the map $(T \otimes 1) : \mathcal{L}(\mathcal{H} \otimes \mathcal{K}) \to \mathcal{L}(\mathcal{F} \otimes \mathcal{K})$ takes the form

$$\big((T \otimes 1)A\big)_{ijkl} = \sum_{mn} T_{imnl} A_{mjkn}, \tag{5.10}$$

i.e., $T \otimes 1$ acts only on the coordinates corresponding to $\mathcal{H} \otimes \mathcal{H}'$ (compare formula (5.7)). By definition, T is completely positive if for each $\mathcal{K}$, $T \otimes 1$ maps positive operators into positive operators. We need the following simple fact, which we leave as an exercise.

Exercise 5.3.7 Show that an operator $A \in \mathcal{L}(\mathcal{H}) \cong \mathcal{H} \otimes \mathcal{H}'$ is positive if and only if there exist $x(n) \in \mathcal{H}$ such that $A = \sum_n |x(n)\rangle\langle x(n)|$. Show that in coordinates this says that $A_{ij} = \sum_n x_i(n)\overline{x}_j(n)$. Show that the $x(n)$ can be chosen orthogonal.

For the space $\mathcal{L}(\mathcal{H} \otimes \mathcal{K}) \cong \mathcal{H} \otimes \mathcal{K} \otimes \mathcal{K}' \otimes \mathcal{H}'$ this means that A is positive if and only if A is of the form

$$A_{ijkl} = \sum_n B_{ij}(n)\overline{B}_{lk}(n) \tag{5.11}$$

for some $B(n) \in \mathcal{H} \otimes \mathcal{K}$.

So far we have only translated all definitions into the language of tensor calculus. Now, we must show that the map $T \otimes 1$ defined in (5.10) maps A of the form (5.11) into operators of the form (5.11) again if and only if T is of the form (5.9).

If T is of the form (5.9) and A is of the form (5.11) then

$$\begin{aligned}\big((T \otimes 1)A\big)_{ijkl} &= \sum_n \sum_m \sum_{pq} V_{ip}(m) B_{pj}(n) \overline{V}_{lq}(m) \overline{B}_{qk}(n) \\ &= \sum_{nm} (V(m)B(n))_{ij} \overline{(V(m)B(n))}_{lk},\end{aligned}$$

which is again something of the form (5.11).

To prove that conversely every completely positive T is of the form (5.9) we use a trick. For the space $\mathcal{K}$ we may in particular choose $\mathcal{K} = \mathcal{H}'$. Let 1 denote the identity in $\mathcal{L}(\mathcal{H}) \cong \mathcal{H} \otimes \mathcal{H}' = \mathcal{H} \otimes \mathcal{K}$. Then

$$A_{ijkl} := 1_{ij}\overline{1}_{kl}$$

defines an operator of the form (5.11). Here $1_{ij} = \overline{1}_{ij} = \delta_{ij}$. Since T is completely positive, $T \otimes 1$ maps A into an operator of the form (5.11), and therefore

$$T_{ijkl} = \sum_{mn} T_{imnl} 1_{mj} \overline{1}_{kn} = \big((T \otimes 1)A\big)_{ijkl} = \sum_p V_{ij}(p)\overline{V}_{kl}(p)$$

for some $V(p) \in \mathcal{F} \otimes \mathcal{K} = \mathcal{F} \otimes \mathcal{H}'$. This shows that T is of the form (5.9). ∎

Exercise 5.3.8 Show that it is possible to choose the $V(n)$ in Theorem 5.3.6 orthogonal with respect to a suitable inner product on $\mathcal{L}(\mathcal{H}, \mathcal{F})$.

Completely positive maps between algebras that are not factor algebras are described by the next theorem.

Theorem 5.3.9 (Completely positive maps) *Let $\mathcal{A} \cong \mathcal{L}(\mathcal{H}_1) \oplus \cdots \oplus \mathcal{L}(\mathcal{H}_n)$ and $\mathcal{B} \cong \mathcal{L}(\mathcal{F}_1) \oplus \cdots \oplus \mathcal{L}(\mathcal{F}_m)$ be positive-*-algebras, written as the direct sum of factor algebras. Let $V_{ij}(k)$ be linear maps in $\mathcal{L}(\mathcal{H}_i, \mathcal{F}_j)$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, $k = 1, \ldots, l(i,j)$. Then*

$$T(A_1, \ldots, A_n) := \Big(\sum_{i=1}^{n} \sum_{k=1}^{l(i,1)} V_{i1}(k) A_i V_{i1}(k)^*, \ldots, \sum_{i=1}^{n} \sum_{k=1}^{l(i,m)} V_{im}(k) A_i V_{im}(k)^* \Big)$$

defines a completely positive linear map $T : \mathcal{A} \to \mathcal{B}$ and conversely, every completely positive linear map from $\mathcal{A}$ to $\mathcal{B}$ is of this form.

Proof Any linear map $T : \mathcal{A} \to \mathcal{B}$ can be decomposed as

$$T(A_1, \ldots, A_n) := \Big(\big(\sum_{i=1}^{n} T_{i1}(A_i) \big), \ldots, \big(\sum_{i=1}^{n} T_{im}(A_i) \big) \Big),$$

where $T_{ij} : \mathcal{L}(\mathcal{H}_i) \to \mathcal{L}(\mathcal{F}_j)$. If we show that T is completely positive if and only if all the T_{ij} are completely positive, then Theorem 5.3.9 follows from Stinespring's theorem. For any inner product space $\mathcal{K}$ we have

$$\big(\mathcal{L}(\mathcal{H}_1) \oplus \cdots \oplus \mathcal{L}(\mathcal{H}_n) \big) \otimes \mathcal{L}(\mathcal{K}) \cong (\mathcal{L}(\mathcal{H}_1) \otimes \mathcal{L}(\mathcal{K})) \oplus \cdots \oplus (\mathcal{L}(\mathcal{H}_n) \otimes \mathcal{L}(\mathcal{K})),$$

and in this isomorphism,

$$(T \otimes 1)(A_1, \ldots, A_n) := \Big(\big(\sum_{i=1}^{n} (T_{i1} \otimes 1)(A_i) \big), \ldots, \big(\sum_{i=1}^{n} (T_{im} \otimes 1)(A_i) \big) \Big).$$

Now

$$(B_1, \ldots, B_n) \in (\mathcal{L}(\mathcal{F}_1) \otimes \mathcal{L}(\mathcal{K})) \oplus \cdots \oplus (\mathcal{L}(\mathcal{F}_m) \otimes \mathcal{L}(\mathcal{K}))$$

is positive if and only if all the B_j are positive. (Check this!) From this it is easy to see that T is completely positive if all the T_{ij} are completely positive. To see that this is also necessary note that $T \otimes 1$ maps $(0, \ldots, 0, A_i, 0, \ldots, 0)$ into

$$\big((T_{i1} \otimes 1)(A_i), \ldots, (T_{im} \otimes 1)(A_i) \big),$$

which must be positive whenever $A_i \in \mathcal{L}(\mathcal{H}_i) \otimes \mathcal{L}(\mathcal{K})$ is positive. This means that all $(T_{ij} \otimes 1)(B_j)$ must be positive, hence all T_{ij} must be completely positive. ∎

We now look at some special types of operations. We first consider the classical, abelian case. Recall that every abelian positive *-algebra $\mathcal{A}$ is of the form $\mathcal{A} \cong$

$\mathbb{C}^n = \{(a_1, \ldots, a_n) : a_i \in \mathbb{C}\}$ and that general probability ρ on $\mathbb{C}^n$ is of the form (see Exercise 1.4.4)

$$\rho(a_1, \ldots, a_n) = \sum_i p_i a_i, \tag{5.12}$$

where the p_i are nonnegative number summing up to one.

Lemma 5.3.10 (Operations on abelian algebras) *Let $\mathbb{C}^n$ and $\mathbb{C}^m$ be abelian positive *-algebras and let $(\pi_{ij})_{i=1,\ldots,n,\ j=1,\ldots,m}$ be nonnegative numbers such that $\sum_j \pi_{ij} = 1$ for each i. Then*

$$(T^*\rho)(b_1, \ldots, b_m) := \rho\Big(\sum_j \pi_{1j} b_j, \ldots, \sum_j \pi_{n,j} b_j\Big)$$

defines an operation from $(\mathbb{C}^n)'$ to $(\mathbb{C}^m)'$, and conversely, every operation $T^* : (\mathbb{C}^n)' \to (\mathbb{C}^m)'$ is of this form.

Proof Immediate from Theorem 5.3.9, the fact that completely positive maps from $\mathbb{C}$ to $\mathbb{C}$ correspond to multiplication with nonnegative numbers, and the fact that $T(1) = 1$. ∎

Note that if ρ is of the form (5.12) then Lemma 5.3.10 says that T^* is of the form

$$(T^*\rho)(b_1, \ldots, b_m) = \sum_j \big(\sum_i p_i \pi_{ij}\big) b_j.$$

A matrix π_{ij} with $\sum_j \pi_{ij} = 1$ for each i is called a *transition kernel*.

We next take a look at deterministic operations. Let $\mathcal{A}, \mathcal{B}$ be positive *-algebras and let $T^* : \mathcal{B}' \to \mathcal{A}'$ be an operation. We will say T^* is *deterministic* if for each inner product space $\mathcal{K}$, the operation

$$(T^* \otimes 1) : (\mathcal{B} \otimes \mathcal{L}(\mathcal{K}))' \to (\mathcal{A} \otimes \mathcal{L}(\mathcal{K}))'$$

maps pure states into pure states. Since pure states are probabilities describing maximal information about a physical system, deterministic operations are operations without loss of information. Slightly generalizing our definition, if $T : \mathcal{A} \to \mathcal{B}$ is a completely positive map then let us say that T^* is deterministic if for each $\mathcal{K}$, $(T^* \otimes 1)$ maps pure states into scalar multiples of pure states.

Proposition 5.3.11 (Deterministic operations) *Let $\mathcal{A} \cong \mathcal{L}(\mathcal{H}_1) \oplus \cdots \oplus \mathcal{L}(\mathcal{H}_n)$ and $\mathcal{B} \cong \mathcal{L}(\mathcal{F}_1) \oplus \cdots \oplus \mathcal{L}(\mathcal{F}_m)$ be positive-*-algebras, written as the direct sum of factor algebras. Let $f : \{1, \ldots, m\} \to \{1, \ldots, n\}$ be a function and let $U_j \in$*

$\mathcal{L}(\mathcal{H}_{f(j)}, \mathcal{F}_j)$ $(j = 1, \ldots, m)$. *Then the completely positive map* $T : \mathcal{A} \to \mathcal{B}$ *given by*

$$T(A_1, \ldots, A_n) := \big(U_1 A_{f(1)} U_1^*, \ldots, U_m A_{f(m)} U_m^*\big)$$

defines a deterministic $T^* : \mathcal{B}' \to \mathcal{A}'$ *and conversely, every deterministic* $T^* : \mathcal{B}' \to \mathcal{A}'$ *is of this form.* T^* *is an operation if and only if the* U_i *are unitary.*

Proof The spaces $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$ and $\mathcal{F}_1 \oplus \cdots \oplus \mathcal{F}_m$ are in an obvious way minimal representations for the algebras $\mathcal{A}$ and $\mathcal{B}$. There is a one-to-one correspondence between linear forms $l \in \mathcal{A}'$ and their densities $L \in \mathcal{A}$ with respect to the trace on $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n$, given by

$$l(A) = \mathrm{tr}(L^* A) \qquad (A \in \mathcal{A}),$$

so the linear map $T^* : \mathcal{B}' \to \mathcal{A}'$ gives rise to a linear map (also denoted by T^*) from $\mathcal{B}$ to $\mathcal{A}$, such that

$$(T^* l)(A) = \mathrm{tr}((T^* L)^* A) \qquad (A \in \mathcal{A}).$$

The fact that $(T^* \otimes 1) : (\mathcal{B} \otimes \mathcal{L}(\mathcal{K}))' \to (\mathcal{A} \otimes \mathcal{L}(\mathcal{K}))'$ maps pure states into scalar multiples of pure states now means that the corresponding map from $(\mathcal{B} \otimes \mathcal{L}(\mathcal{K}))$ to $(\mathcal{A} \otimes \mathcal{L}(\mathcal{K}))$ maps densities of the form $|\psi\rangle\langle\psi|$ with $\psi \in \mathcal{F}_j \otimes \mathcal{K}$ for some j into densities of the form $|\phi\rangle\langle\phi|$ with $\phi \in \mathcal{H}_i \otimes \mathcal{K}$ for some i. (Recall Proposition 4.2.4.)

In particular, for each j, the map $T^* : \mathcal{L}(\mathcal{F}_j) \to \mathcal{L}(\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n)$ maps operators of the form $|\psi\rangle\langle\psi|$ into operators of this form again. Taking $\mathcal{K} = \mathcal{F}_j'$ and mimicking the proof of Stinespring's theorem, we see that the restriction of T^* to $\mathcal{L}(\mathcal{F}_j)$ must be of the form

$$T^*(L) = V_j L V_j^*$$

for some $V_j \in \mathcal{L}(\mathcal{F}_j, \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n)$. Since V_j must map pure states into scalar multiples of pure states, for every $\phi \in \mathcal{F}_j$ there must be an i such that $V_j \phi \in \mathcal{H}_i$. Since $\mathrm{Ran}(V_j)$ is a linear space, this is only possible if $\mathrm{Ran}(V_j) \subset \mathcal{H}_i$ for some i. Denote this i by $f(j)$, let $U_j \in \mathcal{L}(\mathcal{H}_{f(j)}, \mathcal{F}_j)$ denote the adjoint of V_j, and let P_j and P_i denote the orthogonal projections on $\mathcal{F}_j$ and $\mathcal{H}_i$, respectively. Then

$$T^* L = \sum_j P_{f(j)} U_j^* P_j L P_j U_j P_{f(j)}$$

So the linear form $A \mapsto \mathrm{tr}(L^* A)$ is mapped under T^* to

$$\mathrm{tr}((T^* L)^* A) = \sum_j \mathrm{tr}(P_{f(j)} U_j^* P_j L^* P_j U_j P_{f(j)} A) = \sum_j \mathrm{tr}(L^* P_j U_j P_{f(j)} A P_{f(j)} U_j^* P_j).$$

This shows that

$$T(A) = \sum_j P_j U_j P_{f(j)} A P_{f(j)} U_j^* P_j,$$

as claimed. It is easy to see that conversely if T is of this form then T^* is deterministic. T^* is an operation if and only if $T^*(1) = 1$ which is equivalent to the statement that $U_j U_j^*$ is for each j the identity on $\mathcal{F}_j$, which means that the U_j must be unitary. ∎

Exercise 5.3.12 Let $\mathcal{H}, \mathcal{F}$ be inner product spaces, let $\{\phi_1, \ldots, \phi_n\}$ be a basis for $\mathcal{F}$, and let $\psi \in \mathcal{H}$ satisfy $\|\psi\| = 1$. Define $T : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{F})$ by

$$T(B) := \sum_{i=1}^{n} |\phi_i\rangle\langle\psi|B|\psi\rangle\langle\phi_i| \qquad (B \in \mathcal{L}(\mathcal{H})).$$

Show that $T^* : \mathcal{L}(\mathcal{F})' \to \mathcal{L}(\mathcal{H})'$ is an operation that maps pure states into pure states, but T^* is not deterministic.

We have already met an important example of a deterministic operation when we discussed *deterministic time evolution*. Recall that on page 73 we described the deterministic time evolution of probabilities on a positive *-algebra $\mathcal{A}$ by a hermitian operator $H \in \mathcal{A}$ called Hamiltonian. If ρ_0 is the probability describing our knowledge about the system at time 0, then the probability at time t is given by

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

where

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

Note that the operation T_t^* corresponds to the Schrödinger picture, while the completely positive identity preerving map T_t corresponds to the Heisenberg picture.

The last class of operations that we take a special look at is described by the next lemma.

Lemma 5.3.13 (Inner operations) *Let $\mathcal{A}$ be a positive *-algebra and assume that $V(1), \ldots, V(n) \in \mathcal{A}$ satisfy $\sum_m V(m)V(m)^* = 1$. Let $T : \mathcal{A} \to \mathcal{A}$ be given by*

$$T(A) := \sum_m V(m) A V(m)^* \qquad (A \in \mathcal{A}).$$

Then $T^* : \mathcal{A}' \to \mathcal{A}'$ is an operation. If $\mathcal{A}$ is a factor algebra then every operation $T^* : \mathcal{A}' \to \mathcal{A}'$ is of this form. If $\mathcal{A}$ is abelian, then the only operation of this form is the identity map from $\mathcal{A}'$ to $\mathcal{A}'$.

Proof Write $\mathcal{A}$ as a sum of factors $\mathcal{A} = \mathcal{L}(\mathcal{H}_1) \oplus \cdots \oplus \mathcal{L}(\mathcal{H}_n)$. Let P_i denote the orthogonal projection on $\mathcal{H}_i$ and define $V_i(m) \in \mathcal{L}(\mathcal{H}_i, \mathcal{H}_i)$ by

$$V_i(m)\phi := P_i V(m)\phi \qquad (\phi \in \mathcal{H}_i).$$

Then T is of the form of Theorem 5.3.9, where the maps $V_{ij}(k)$ there are zero for $i \neq j$. The converse is also true: Write a general operation $T^* : \mathcal{A}' \to \mathcal{A}'$ in the form of Theorem 5.3.9. In the special case that the maps $V_{ij}(k)$ there are zero for $i \neq j$, T is of the form

$$T(A_1, \ldots, A_n) = \Big(\sum_{k=1}^{l(1)} V_1(k) A_1 V_1(k)^*, \ldots, \sum_{k=1}^{l(n)} V_n(k) A_1 V_n(k)^* \Big)$$

where $V_i(k) \in \mathcal{L}(\mathcal{H}_i)$. Therefore, if we define $W_i(k) \in \mathcal{A}$ by

$$W_i(k) := (0, \ldots, 0, \underset{\substack{\uparrow \\ i\text{-th place}}}{V_i(k)}, 0, \ldots, 0),$$

then

$$T(A) = \sum_{i=1}^{n} \sum_{k=1}^{l(i)} W_i(k) A W_i(k)^*.$$

is of the form from Lemma 5.3.13. This shows in particular that if $\mathcal{A}$ is a factor algebra then every operation is of the form from Lemma 5.3.13. On the other hand, if $\mathcal{A}$ is abelian, then any operation of this form satisfies $T^*\rho(A) = \rho(T(A)) = \rho\big(\sum_m V(m) A V(m)^*\big) = \rho\big(\sum_m V(m) V(m)^* A\big) = \rho(1A) = \rho(A)$. ∎

In these lecture notes, we will call operations $T^* : \mathcal{A}' \to \mathcal{A}'$ of the form from Lemma 5.3.13 *inner operations*, since they can be constructed with operators $V(m)$ from the algebra $\mathcal{A}$ itself. We see that factor algebras have a double function. They describe all possible *observations* one can do on a physical system (represented by projections $P \in \mathcal{A}$), but also all possible *operations* one can perform on the system (represented by operators $V(1), \ldots, V(n) \in \mathcal{A}$ such that $\sum_m V(m)V(m)^* = 1$). On the other hand, on abelian algebras observations and operations appear like completely unconnected objects.

Note that if $\mathcal{A}$ is a positive *-algebra and $H \in \mathcal{A}$ is hermitian, then time evolution with the Hamiltonian H is an inner operation.

Another important example of an inner operation is an *ideal measurement*. Recall that ideal measurements are described by a partition of the identity $\{P_1, \ldots, P_n\}$.

According to point 4° of our interpretation of quantum probability spaces, if we know that someone performs the ideal measurement $\{P_1, \ldots, P_n\}$ on a physical system (but we don't know the outcome), and before the measurement our knowledge about the system was described by the quantum probability space $(\mathcal{A}, \rho)$, then after the measurement we must describe our knowledge by the quantum probability space $(\mathcal{A}, \rho')$, where

$$\rho'(A) = \sum_{i=1}^{n} \rho(P_i A P_i) = (T^*\rho)(A) \qquad (A \in \mathcal{A}),$$

where $T : \mathcal{A} \to \mathcal{A}$ is given by

$$T(A) = \sum_{i=1}^{n} P_i A P_i \qquad (A \in \mathcal{A}).$$

In Exercise 1.4.7 we proved that $\rho' = \rho$ if $\mathcal{A}$ is abelian. This is a special case of the more general statement that all inner operations on an abelian algebra are trivial.

5.4 Measurements

So far, our interpretation of quantum probability was based on ideal measurements. Using the theory of completely positive maps, we can give a new interpretation of quantum probability, that also allows for measurements that are not ideal. It will be convenient to introduce the folowing notation: when ρ is a positive linear form on a positive *-algebra $\mathcal{A}$, we set

$$|\rho| := \rho(1).$$

We say that a completely positive map T is *contracting* if

$$T(1) \leq 1.$$

If $\mathcal{A}$ is a positive *-algebra and $T : \mathcal{A} \to \mathcal{A}$ is a contracting completely positive map, then we call $T^* : \mathcal{A} \to \mathcal{A}$ a (nonideal) observation.

We formulate our new interpretation of quantum probability spaces as follows (compare Section 1.3).

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

2° We can interact with the physical system by performing measurements. A measurement is described by a finite collection $\{T_i : i \in I\}$ of completely positive maps $T_i : \mathcal{A} \to \mathcal{A}$ such that $\sum_{i \in I} T_i = 1$. We interpret the T_i^* as observations and the labels $i \in I$ as possible outcomes of the measurement.

3° If we perform the measurement $\{T_i : i \in I\}$ on the physical system, then this measurement yields the outcome $i \in I$ with probability $p_i := |T_i^* \rho|$. The probability p_i measures how likely we judge it to be that our measurement yields the outcome i. If $p_i = 1$ then the outcome i is sure. If the probability of an outcome of one measurement is larger than (equal to) the probability of another outcome of another measurement, then the first outcome is more likely than (equally likely as) the other outcome.

4° If we know that the measurement $\{T_i : i \in I\}$ is performed on the system, but we don't know the outcome, then we must describe our changed knowledge about the system with the new probability ρ' on $\mathcal{A}$ defined by $\rho' := \sum_{i \in I} T_i^* \rho$.

5° If the measurement $\{T_i : i \in I\}$ is performed on the system and we learn that it yields the outcome i, then we must describe our changed knowledge about the system with the new probability $\tilde{\rho}$ on $\mathcal{A}$ defined by $\tilde{\rho} := T_i^* \rho / |T_i^* \rho|$. This formula is not defined if $|T_i^* \rho| = 0$ but in that case we were sure that the ideal measurement would yield another outcome than i, so there was something wrong with the way we described our knowledge before the measurement.

By definition, we say that a measurement $\{T_i : i \in I\}$ is *ideal* if there exists a partition of the identity $\{P_i : i \in I\}$ such that $T_i(A) = P_i A P_i$ $(A \in \mathcal{A},\ i \in I)$. We call a measurement *nonperturbing* if

$$\sum_{i \in I} T_i^* = 1,$$

i.e., performing the measurement and not looking at the outcome doesn't change the probability. It is easy to see that ideal measurements are the only nonperturbing if the positive *-algebra is abelian.

We say that two measurements $\{T_i : i \in I\}$ and $\{S_j : j \in J\}$ are *commuting* if

$$T_i S_j \rho = S_j T_i \rho \qquad (i \in I,\ j \in J,\ \rho \in \mathcal{A}').$$

If two measurements commute then it does not matter in which order we perform them. If we first perform the measurement $\{T_i : i \in I\}$ and then $\{S_j : j \in J\}$ and we obtain the outcomes i and j, then this has the same probability as when we first perform the measurement $\{S_j : j \in J\}$ and then $\{T_i : i \in I\}$ and we obtain the outcomes j and i. Moreover, after we have done these observations, we end up with the same new probability, regardless of the order in which we obtained i and j.

It is easy to see that any two *ideal* measurements on an abelian probability space $(\mathcal{A}, \rho)$ commute.

Exercise 5.4.1 Give an example of two (nonideal) measurements on an abelian probability space $(\mathcal{A}, \rho)$ that do *not* commute.

Chapter 6

Peculiarities of quantum probability

6.1 The Kochen-Specker paradox

We have seen in Section 1.4 that a quantum probability space $(\mathcal{A}, \rho)$ is equivalent to a 'classical' probability space $(\Omega, \mathcal{P}(\Omega), \mu)$ if $\mathcal{A}$ is abelian. Our aim is to prove a sort of converse to this statement, i.e., to show that if $\mathcal{A}$ is not abelian, then $(\mathcal{A}, \rho)$ cannot be replaced by a classical probability space. In Lemma 4.2.6, we already saw that if $\mathcal{A}$ is not abelian, then there exist pure states on $\mathcal{A}$ that are not precise. Here pure states were defined as probabilities that cannot be decomposed into other probabilities, so that they correspond, in a way, to maximal knowledge. Precise states were defined as probabilities that assign to every ideal observation either the value 0 or 1.

When thinking about classical probability, we usually assume that all physical quantities have values -we only don't know which ones. We will see that this intuition runs into difficulties in quantum probability. The fact that not every probability can be decomposed into precise states is in itself not very worrying. After all, perhaps this is only an artefact of our theory, and it is possible to extend quantum probability by enlarging our class of probabilities $\mathcal{A}'_{\mathrm{prob}}$ in such a way that in the extended theory every probability *can* be written as a convex combination of precise states. 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.

Our first result in this direction is called the Kochen-Specker paradox [KS67].

Theorem 6.1.1 (Kochen-Specker paradox) *Let $\mathcal{H}$ be an inner product space of dimension at least 3. Then it is not possible to assign to every projection $P \in \mathcal{L}(\mathcal{H})$ a value 'true' or 'false', in such a way that in every ideal measurement $\{P_1, \ldots, P_n\}$ exactly one projection has the value 'true' and all others have the value 'false'. There even exists a finite set $\mathcal{P} \subset \mathcal{L}(\mathcal{H})$ of projections such that it is not possible to assign to every $P \in \mathcal{P}$ a value 'true' or 'false' in the way described above.*

Remark 1 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 2 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 3 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, hidden variable theories do not lead to serious paradoxes. Of course, if one has two logically independent algebras of the form $\mathcal{L}(\mathcal{H})$ with $\dim(\mathcal{H}) = 2$, then they already generate an algebra of the form $\mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{H}) \cong \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$ where $\mathcal{H} \otimes \mathcal{H}$ has dimension four.

Proof of Theorem 6.1.1 As will be obvious from our proof, it suffices to prove the statement for the case $\dim(\mathcal{H}) = 3$. Choose an orthonormal basis $\{\phi_1, \phi_2, \phi_3\}$ of $\mathcal{H}$ and consider projections of the form

$$P := |\psi\rangle\langle\psi| \quad \text{with} \quad \psi = x_1\phi_1 + x_2\phi_2 + x_3\phi_3,$$

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

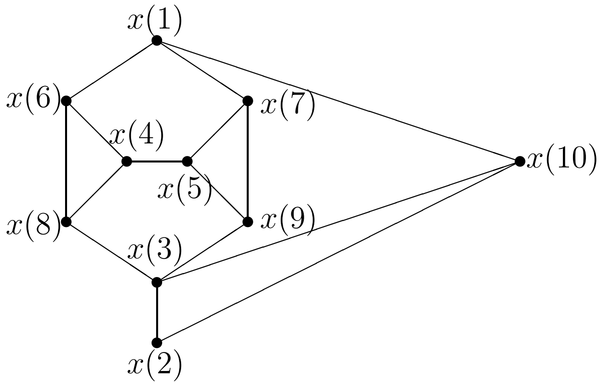

Figure 6.1: Kochen-Specker diagram

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 $\mathcal{P}'$ cannot be colored according to these rules then by adding finitely many points we get a set $\mathcal{P}$ that cannot be colored 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.

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.) ∎

6.2 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: Even in classical probability, there is no reason why it should be possible to associate with every nonideal observation T_i a value 'true' or 'false' that does not depend on the nonideal measurement $\{T_1, \ldots, T_n\}$ that it occurs in. Since it is hard to be sure in practice whether a given measurement is ideal, this is a weakness of the Kochen-Specker paradox. In this section we discuss a test on the nonabelianness of a probability theory that does not assume that certain measurements are ideal.

Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be positive *-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. If T_i^* is an observation on the subsystem $\mathcal{A}_1$, then the corresponding observation on the joint system $\mathcal{A}_1 \otimes \mathcal{A}_2$ is described by $(T_i^* \otimes 1) : (\mathcal{A}_1 \otimes \mathcal{A}_2)' \to (\mathcal{A}_1 \otimes \mathcal{A}_2)'$. If ρ is a probability on $\mathcal{A}_1 \otimes \mathcal{A}_2$, then $\mathcal{A}_1$ and $\mathcal{A}_2$ are independent under ρ if an observation one one subsystem gives no information about the other subsystem. We have seen that this is equivalent to the statement that ρ is a product measure:

$$\rho = \rho_1 \otimes \rho_2$$

where $\rho_j \in (\mathcal{A}_j)'_{\mathrm{prob}}$ $(j = 1, 2)$ are the marginals of ρ. By definition, we called a probability ρ on $\mathcal{A}_1 \otimes \mathcal{A}_2$ an *entangled state* if ρ is *not* a convex combination of product probabilities, i.e., ρ is *not* of the form

$$\rho = \sum_k p_k \, \rho_{1,k} \otimes \rho_{2,k},$$

where the $\rho_{1,k}$ and $\rho_{2,k}$ are probabilities on $\mathcal{A}_1$ and $\mathcal{A}_2$, respectively, and the p_i are nonnegative numbers summing up to one. In Exercise 5.2.4 we proved that there exist entangled states on $\mathcal{A}_1 \otimes \mathcal{A}_2$ if and only if $\mathcal{A}_1$ and $\mathcal{A}_2$ are both nonabelian. Therefore, entanglement can be used to detect nonabelianness. We will see that entanglement can be measured, thereby allowing for an experimental verification of quantum probability.

Let $\{T_{+a}, T_{-a}\}$ and $\{T_{+b}, T_{-b}\}$ be measurements on the subsystem $\mathcal{A}_1$ and let $\{S_{+a}, S_{-a}\}$ and $\{S_{+b}, S_{-b}\}$ be measurements on the subsystem $\mathcal{A}_2$. Let

$$\pi(x, y) := |(T_x \otimes S_y)\rho| \qquad (x, y = \pm a, \pm b)$$

denote the probability of doing the observation x on the subsystem $\mathcal{A}_1$ and the observation y on the subsystem $\mathcal{A}_2$. If the probability ρ on $\mathcal{A}_1 \otimes \mathcal{A}_2$ is not entangled, then these probabilities $\pi(x, y)$ must be of the form

$$\pi(x, y) = \sum_k p_k \, \pi_{1,k}(x)\pi_{2,k}(y) \qquad (x, y = \pm a, \pm b). \tag{6.1}$$

where the $\pi_{i,k}(x)$ are nonnegative numbers such that $\pi_{i,k}(+x) + \pi_{i,k}(-x) = 1$ for all $x = a, b,\ i = 1, 2$, and k, and the p_k are nonnegative numbers summing up to one.

We define *correlation coefficients* between the measurements on different subsystems by

$$\begin{aligned} c(x, y) &:= \pi(+x, +y) + \pi(-x, -y) - \pi(+x, -y) - \pi(-x, +y) \\ &= \sum_{\sigma,\tau=+,-} \pi(\sigma x, \tau y)(\sigma 1)(\tau 1) \qquad (x, y = a, b). \end{aligned}$$

Note that $c(x, y) = 1$ if the measurements $\{T_{+x}, T_{-x}\}$ and $\{S_{+y}, S_{-y}\}$ always both yield the $+$ outcome or both yield the $-$ outcome. On the other hand, $c(x, y) = -1$ if always one measurement yields the $+$ outcome and the other one the $-$ outcome.

Theorem 6.2.1 (Bell inequality) *If the probabilities $\pi(x, y)$ satisfy* (6.1)*, then*

$$\big|c(a, a) + c(a, b) + c(b, a) - c(b, b)\big| \leq 2.$$

Remark Since correlation coefficients can take values between -1 and 1, it seems at first that the expression in Bell's inequality can take values up to 4. The statement of the theorem is that for probabilities satisfying the no entanglement condition (6.1), the maximal value is 2.

Proof of Theorem 6.2.1 If π is of the form

$$\pi(x, y) = \pi_1(x)\pi_2(y) \qquad (x, y = \pm a, \pm b),$$

then, setting

$$\gamma_i(x) := \pi_i(+x) - \pi_i(-x) \qquad (i = 1, 2,\ x = a, b),$$

we have

$$\begin{aligned}
&c(a,a) + c(a,b) + c(b,a) - c(b,b) \\
&\quad = \sum_{\sigma,\tau=+,-} \pi_1(\sigma a)\pi_2(\tau a)(\sigma 1)(\tau 1) + \sum_{\sigma,\tau=+,-} \pi_1(\sigma a)\pi_2(\tau b)(\sigma 1)(\tau 1) \\
&\qquad + \sum_{\sigma,\tau=+,-} \pi_1(\sigma b)\pi_2(\tau a)(\sigma 1)(\tau 1) - \sum_{\sigma,\tau=+,-} \pi_1(\sigma b)\pi_2(\tau b)(\sigma 1)(\tau 1) \\
&\quad = \gamma_1(a)\gamma_2(a) + \gamma_1(a)\gamma_2(b) + \gamma_1(b)\gamma_2(a) - \gamma_1(b)\gamma_2(b) \\
&\quad = \gamma_1(a)\big(\gamma_2(a) + \gamma_2(b)\big) + \gamma_1(b)\big(\gamma_2(a) - \gamma_2(b)\big),
\end{aligned}$$

so

$$\begin{aligned}
&\big|c(a,a) + c(a,b) + c(b,a) - c(b,b)\big| \\
&\qquad \le |\gamma_1(a)|\big|\gamma_2(a) + \gamma_2(b)\big| + |\gamma_1(b)|\big|\gamma_2(a) - \gamma_2(b)\big| \\
&\qquad \le \big|\gamma_2(a) + \gamma_2(b)\big| + \big|\gamma_2(a) - \gamma_2(b)\big|,
\end{aligned}$$

where we have used that $|\gamma_1(x)| \le 1$. If $\gamma_2(a) + \gamma_2(b)$ and $\gamma_2(a) - \gamma_2(b)$ have the same sign then

$$\begin{aligned}
&\big|\gamma_2(a) + \gamma_2(b)\big| + \big|\gamma_2(a) - \gamma_2(b)\big| \\
&\qquad = \big|\gamma_2(a) + \gamma_2(b) + \gamma_2(a) - \gamma_2(b)\big| = 2|\gamma_2(a)|,
\end{aligned}$$

and otherwise this expression is $2|\gamma_2(b)|$. In any case, it is smaller or equal to 2. More generally, if π is of the form (6.1), then

$$c(a,a) + c(a,b) + c(b,a) - c(b,b) = \sum_k p_k\big(c_k(a,a) + c_k(a,b) + c_k(b,a) - c_k(b,b)\big)$$

where c_k is the correlation coefficient under $\pi_k(x,y) := \pi_{1,k}(x)\pi_{2,k}(y)$, so also in this case this quantity is smaller or equal to 2. ∎

Our next aim is to show that entangled states *can* violate the Bell inequality. We demonstrate this with ideal measurements on an entangled state that describes two polarized photons. 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, 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. No other outcomes are possible! It is said that the polarization of a single photon is *quantized*, since it can only take certain integer values, depending on how we measure it.

In the language of quantum mechanics, polarization is described as follows. There is a two-dimensional inner product space $\mathcal{H}$ with orthonormal basis $\{\phi_1, \phi_2\}$. The polarization state of a single photon is described by a probability on $\mathcal{L}(\mathcal{H})$. For every angle α in the plane perpendicular to the direction in which the light travels, the orthogonal projection $P_\alpha = |\eta_\alpha\rangle\langle\eta_\alpha|$ on the vector

$$\eta_\alpha := \cos(\alpha)\phi_1 + \sin(\alpha)\phi_2$$

corresponds to the observation that the photon is polarized in the direction α. Here being polarized in the directions α or $\alpha + \pi$ is the same. An ideal measurement on the polarization state of the photon is described by a partition of the identity of the form $\{P_\alpha, P_{\alpha+\pi/2}\}$, and will determine whether the photon is polarized in the direction α or in the perpendicular direction $\alpha + \frac{1}{2}\pi$.

Exercise 6.2.2 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 $\cos(\beta - \alpha)^2$.

It turns out to be possible to create pairs of photons (usually traveling in opposite directions) that are polarized in the same direction. More precisely, such pairs of photons are described by the entangled state ρ_ψ on $\mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{H}) \cong \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$ given by

$$\psi := \tfrac{1}{\sqrt{2}}\,\phi_1 \otimes \phi_1 + \tfrac{1}{\sqrt{2}}\,\phi_2 \otimes \phi_2.$$

The probability of finding one photon in the polarization direction α and the other photon in the polarization direction β is

$$\begin{aligned}
\rho_\psi(P_\alpha \otimes P_\beta) &= \tfrac{1}{2}\big\langle \phi_1 \otimes \phi_1 + \phi_2 \otimes \phi_2 \big| P_\alpha \otimes P_\beta \big| \phi_1 \otimes \phi_1 + \phi_2 \otimes \phi_2 \big\rangle \\
&= \tfrac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\langle \phi_i \otimes \phi_i | P_\alpha \otimes P_\beta | \phi_j \otimes \phi_j \rangle \\
&= \tfrac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\langle \phi_i | P_\alpha | \phi_j \rangle \langle \phi_i | P_\beta | \phi_j \rangle
\end{aligned}$$

$$\begin{aligned}
&= \tfrac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\langle\phi_i|P_\alpha|\phi_j\rangle\langle\phi_j|P_\beta|\phi_i\rangle = \tfrac{1}{2}\mathrm{tr}(P_\alpha P_\beta) = \tfrac{1}{2}\mathrm{tr}(P_0, P_{\beta-\alpha}) \\
&= \tfrac{1}{2}\Big|\big\langle\phi_1\big|\cos(\beta-\alpha)\phi_1 + \sin(\beta-\alpha)\phi_2\big\rangle\Big|^2 \\
&= \tfrac{1}{2}\cos(\beta-\alpha)^2.
\end{aligned}$$

Here we have used rotation invariance and the facts that the operators P_α are self-adjoint and that their matrices with respect to the basis $\{\phi_1, \phi_2\}$ have real entries. Therefore, if we measure whether the first photon is polarized in the direction α and the second photon is polarized in the direction β, then the probability that we get either two times a positive answer or two times a negative answer is $\cos(\beta-\alpha)^2$. The probability that we get different answers is $\sin(\beta-\alpha)^2$. Let $T_\alpha(A) := P_\alpha A P_\alpha$ be the ideal observation that a photon is polarized in the direction α. If on the first photon we perform the ideal measurement $\{T_\alpha, T_{\alpha+\pi/2}\}$ and on the second photon we perform the ideal measurement $\{T_\beta, T_{\beta+\pi/2}\}$, then the correlation coefficient between these ideal measurements is

$$\cos(\beta-\alpha)^2 - \sin(\beta-\alpha)^2 = 2\cos(\beta-\alpha)^2 - 1 = \cos(2(\beta-\alpha))$$

We now compare this with the set-up of Theorem 6.2.1. For the measurements $\{T_{+a}, T_{-a}\}$ and $\{T_{+b}, T_{-b}\}$ on the subsystem $\mathcal{A}_1$ and measurements $\{S_{+a}, S_{-a}\}$ and $\{S_{+b}, S_{-b}\}$ on the subsystem $\mathcal{A}_2$, we take

$$\begin{aligned}
T_{+a} &= T_0, \qquad & T_{+b} &= T_\gamma, \\
S_{+a} &= T_\alpha, \qquad & S_{+b} &= T_{\alpha-\gamma},
\end{aligned}$$

and for T_{-a}, T_{-b}, S_{-a}, and S_{-b} be take the observations in the perpendicular directions. The expression in Theorem 6.2.1 then becomes

$$\big|\cos(2\alpha) + 2\cos(2\gamma) - \cos(4\gamma - 2\alpha)\big|.$$

We want to maximize this expression. Setting the derivatives with respect to α and γ to zero yields the equations

$$\begin{aligned}
-2\sin(2\alpha) - 2\sin(4\gamma - 2\alpha) &= 0, \\
-4\sin(2\gamma) + 4\sin(4\gamma - 2\alpha) &= 0.
\end{aligned}$$

It follows that $\sin(2\gamma) = \sin(4\gamma - 2\alpha) = -\sin(2\alpha)$. We choose

$$\gamma = \pi - \alpha.$$

The expression to be maximized then becomes

$$\cos(2\alpha) + 2\cos(2\pi - 2\alpha) - \cos(2\pi - 6\alpha) = 3\cos(2\alpha) - \cos(6\alpha).$$

Differentiating and setting to zero yields

$$-6\sin(2\alpha) + 6\sin(6\alpha) = 0 \ \Rightarrow\ \sin(2\alpha) = \sin(6\alpha).$$

Setting $z = e^{i2\alpha}$, we need to solve

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

Setting $y = z^2 = e^{i4\alpha}$ we get the cubic equation

$$y^3 - y^2 + y - 1 = 0.$$

We know that $y = e^{i2\pi} = 1$ 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 Theorem 6.2.1 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].

The strength of the Bell inequality lies in the fact that we have not made any assumptions on how the measurement process works locally. It holds whenever the following assumptions are fulfilled.

1° Our probabilities come from a classical probability space $(\mathcal{A}, \rho)$.

2° The algebra $\mathcal{A}$ is of the form $\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$, where $\mathcal{A}_1$ describes the first subsystem and $\mathcal{A}_2$ the second subsystem.

3° An observation on the first subsystem is described by an operator of the form $T_x \otimes 1$ where T_x is an observation on $\mathcal{A}_1$. Likewise, observations on the second subsystem are of the form $1 \otimes S_y$.

It seems that dropping the assumptions 2° or 3° would necessarily introduce some form of nonlocality into our system. Assumption 2° says that our subsystems are logically independent, and assumption 3° says that our measurements are restricted to one subsystem. In particular, these assumptions guarantee that performing a measurement on one subsystem does not change some (hidden) variables referring to the other subsystem. In view of this, many people prefer to drop the first assumption, rather than the other two.

6.3 Cloning, coding, and teleportation

In this section we will meet two surprising quantum impossibilities, and one possibility. The first theorem that we will prove says that it is not possibly to *copy* a quantum system. This result goes under the fancy name *'no cloning'*.

Theorem 6.3.1 (No cloning) *Let $\mathcal{A}$ be a positive *-algebra and let $T^* : \mathcal{A}' \to (\mathcal{A} \otimes \mathcal{A})'$ be an operation such that*

$$(T^*\rho)(A \otimes 1) = (T^*\rho)(1 \otimes A) = \rho(A) \qquad (A \in \mathcal{A},\ \rho \in \mathcal{A}'_{\rm prob}). \tag{6.2}$$

Then $\mathcal{A}$ is abelian.

Note that (6.2) says that the operation T^* takes a single system in the state ρ and produces two logically independent (but not necessarily independent) systems, both in the state ρ. The claim is that if the algebra is not abelian, then there is no operation that does this for any state ρ. On the other hand, in classical probability, copying is always possible:

Exercise 6.3.2 Show that if $\mathcal{A}$ is abelian, there exists an operation T^* such that (6.2) holds. Show that T^* can be chosen in such a way that

$$(T^*\rho)(P \otimes 1) = (T^*\rho)(P \otimes P) = (T^*\rho)(1 \otimes P)$$

for every projection $P \in \mathcal{A}$ and $\rho \in \mathcal{A}'_{\rm prob}$. If T^* is chosen in this way, then are the two subsystems in general independent under $T^*\rho$?

We postpone the proof of Theorem 6.3.1 and first state another, similar theorem. This theorem says that it is not possible to extract all information from a nonabelian algebra, and 'write it down' in an abelian algebra, such that with this information the original nonablian system can later be reconstructed. This result goes under the name *'no coding'*.

Theorem 6.3.3 (No coding) *Let $\mathcal{A}, \mathcal{B}$ be a positive *-algebras and let $T^* : \mathcal{A}' \to \mathcal{B}'$ and $S^* : \mathcal{B}' \to \mathcal{A}'$ be operations such that*

$$S^*T^*\rho = \rho \qquad (\rho \in \mathcal{A}'_{\rm prob}). \tag{6.3}$$

Then if $\mathcal{B}$ is abelian, so is $\mathcal{A}$.

Remark As in the previous theorem, it is important that S^* and T^* work for *any* $\rho \in \mathcal{A}'_{\rm prob}$.

Again we postpone the proof, and first state another theorem. This theorem says that it is possible to extract information from a quantum system, 'write it down' in an abelian algebra, and then send it to someone else, so that he can reconstruct the original system. This seems to contradict our previous theorem, but the trick is that the two people who want to send information to each other have prepared two entangled particles, and each of them keeps one particle with them. By making use of these entangled particles, they can send the quantum system. This result goes under the science fiction-like name *'teleportation'* [B–W93]. For simplicity, we will only teleport states on an algebra of the form $\mathcal{L}(\mathcal{H})$ where $\mathcal{H}$ is two-dimensional. As is standard in this sort of communication problems, the sender will be called *Alice* and the recipient will be called *Bob*. (Which explains why I wrote that '*he* can reconstruct the original system'.)

Theorem 6.3.4 (Quantum teleportation) *Let $\mathcal{A}$ and $\mathcal{C}$ be positive-*-algebras of the form $\mathcal{L}(\mathcal{H})$, where $\mathcal{H}$ is two-dimensional, and let $\mathcal{B}$ be a four-dimensional abelian positive-*-algebra. Then there exists a state $\eta \in (\mathcal{C}\otimes\mathcal{C})'_{\rm prob}$ and operations $T^* : (\mathcal{A}\otimes\mathcal{C})' \to \mathcal{B}'$ and $S^* : (\mathcal{B}\otimes\mathcal{C})' \to \mathcal{A}'$, such that*

$$S^* \circ (T^* \otimes 1)\, \rho \otimes \eta = \rho \qquad (\rho \in \mathcal{A}'_{\rm prob}). \tag{6.4}$$

Remark 1 In (6.4), $(T^* \otimes 1)$ is map from $(\mathcal{A}\otimes\mathcal{C}\otimes\mathcal{C})'_{\rm prob}$ to $(\mathcal{B}\otimes\mathcal{C})'_{\rm prob}$, so that the composition of operations $S^* \circ (T^* \otimes 1)$ is a map from $(\mathcal{A}\otimes\mathcal{C}\otimes\mathcal{C})'_{\rm prob}$ to $\mathcal{A}'_{\rm prob}$. The abelian algebra $\mathcal{B}$ contains all information that is sent from Alice to Bob. Therefore, the operation $T^* \otimes 1$ acts only on objects that are under Alice's control and S^* acts only on objects that are under the control of Bob.

Remark 2 It follows from Lemma 5.3.3 that entangled states cannot be used to send information from Alice to Bob, in spite of their seemingly nonlocal behavior in relation with the Bell inequality. Nevertheless, quantum teleportation shows that entangled states can be used to upgrade 'classical information' to 'quantum information'.

We now set out to prove Theorems 6.3.1–6.3.4. We start with a preparatory result. Recall the partial order for hermitian operators defined on page 63.

Theorem 6.3.5 (Cauchy-Schwarz for operations) *Let $\mathcal{A}, \mathcal{B}$ be positive *-algebras and let $T^* : \mathcal{A}' \to \mathcal{B}'$ be an operation. Then*

$$T(B^*B) \geq T(B)^*T(B) \qquad (B \in \mathcal{B}).$$

If equality holds for some $B_1 \in \mathcal{B}$, then

$$T(B_1^*B_2) = T(B_1)^*T(B_2) \quad \textit{and} \quad T(B_2^*B_1) = T(B_2)^*T(B_1).$$

Remark 1 The inequality is called 'Cauchy-Schwarz' since we can view $(B_1, B_2) \mapsto T(B_1^* B_2)$ as a sort of $\mathcal{A}$-valued (!) 'inner product' on $\mathcal{B}$. Then we can rewrite the inequality above as $T(B^*B)T(1^*1) \geq T(1^*B)^*T(1^*B)$ which looks like $\langle\phi|\phi\rangle\langle\psi|\psi\rangle \geq |\langle\psi|\phi\rangle|^2$.

Remark 2 Note that if a Cauchy-Schwarz *equality* holds for *all* $B \in \mathcal{B}$, then T is a *-algebra homomorphism.

Proof The operator

$$\begin{pmatrix} B^*B & -B^* \\ -B & 1 \end{pmatrix} = \begin{pmatrix} B & -1 \\ 0 & 0 \end{pmatrix}^* \begin{pmatrix} B & -1 \\ 0 & 0 \end{pmatrix}$$

in $M_2(\mathcal{A})$ is positive. Therefore, since T is completely positive, also

$$\begin{pmatrix} T(B^*B) & -T(B)^* \\ -T(B) & 1 \end{pmatrix}$$

is positive. If $\mathcal{H}$ is a representation for $\mathcal{A}$ then $\mathcal{H} \oplus \mathcal{H}$ is a representation for $M_2(\mathcal{A})$ and therefore, for each $\psi \in \mathcal{H}$,

$$\left\langle \begin{pmatrix} \psi \\ T(B)\psi \end{pmatrix} \middle| \begin{pmatrix} T(B^*B) & -T(B)^* \\ -T(B) & 1 \end{pmatrix} \begin{pmatrix} \psi \\ T(B)\psi \end{pmatrix} \right\rangle \geq 0.$$

This means that

$$\left\langle \begin{pmatrix} \psi \\ T(B)\psi \end{pmatrix} \middle| \begin{pmatrix} T(B^*B)\psi - T(B)^*T(B)\psi \\ 0 \end{pmatrix} \right\rangle \geq 0,$$

and therefore

$$\langle\psi|T(B^*B) - T(B)^*T(B)|\psi\rangle \geq 0 \qquad (\psi \in \mathcal{H}),$$

so $T(B^*B) - T(B)^*T(B)$ is a positive operator, i.e., $T(B^*B) \geq T(B)^*T(B)$.

Now assume that $T(B_1^*B_1) = T(B_1)^*T(B_1)$ for some $B_1 \in \mathcal{B}$. Then, for all $B_2 \in \mathcal{B}$ and $\lambda_1, \lambda_2 \in \mathbb{C}$,

$$\begin{aligned} &|\lambda_1|^2 T(B_1^*B_1) + \overline{\lambda}_1\lambda_2 T(B_1^*B_2) + \overline{\lambda}_2\lambda_1 T(B_2^*B_1) + |\lambda_2|^2 T(B_2^*B_2) \\ &\quad = T\big((\lambda_1 B_1 + \lambda_2 B_2)^*(\lambda_1 B_1 + \lambda_2 B_2)\big)\big) \\ &\quad \geq |\lambda_1|^2 T(B_1)^*T(B_1) + \overline{\lambda}_1\lambda_2 T(B_1)^*T(B_2) \\ &\qquad + \overline{\lambda}_2\lambda_1 T(B_2)^*T(B_1) + |\lambda_2|^2 T(B_2)^*T(B_2). \end{aligned}$$

Using our assumption that $T(B_1^*B_1) = T(B_1)^*T(B_1)$ this implies that

$$\begin{aligned}&\overline{\lambda}_1\lambda_2 T(B_1^*B_2) + \overline{\lambda}_2\lambda_1 T(B_2^*B_1) + |\lambda_2|^2 T(B_2^*B_2) \\ &\qquad \geq \overline{\lambda}_1\lambda_2 T(B_1)^*T(B_2) + \overline{\lambda}_2\lambda_1 T(B_2)^*T(B_1) + |\lambda_2|^2 T(B_2)^*T(B_2).\end{aligned}$$

In particular, setting $\lambda_1 = 1$ and $\lambda_2 = \lambda$ where λ is real, we get

$$\begin{aligned}&\lambda\big(T(B_1^*B_2) + T(B_2^*B_1)\big) + \lambda^2 T(B_2^*B_2) \\ &\qquad \geq \lambda\big(T(B_1)^*T(B_2) + T(B_2)^*T(B_1)\big) + \lambda^2 T(B_2)^*T(B_2).\end{aligned}$$

This can only be true for all $\lambda \in \mathbb{R}$ (in particular, for λ very small), if

$$T(B_1^*B_2) + T(B_2^*B_1) \geq T(B_1)^*T(B_2) + T(B_2)^*T(B_1).$$

In the same way, setting $\lambda_1 = -i$ and $\lambda_2 = i\lambda$, where λ is real, we get

$$\begin{aligned}&-\lambda\big(T(B_1^*B_2) + T(B_2^*B_1)\big) + \lambda^2 T(B_2^*B_2) \\ &\qquad \geq -\lambda\big(T(B_1)^*T(B_2) + T(B_2)^*T(B_1)\big) + \lambda^2 T(B_2)^*T(B_2)\end{aligned}$$

which together with our previous inequality yields

$$T(B_1^*B_2) + T(B_2^*B_1) = T(B_1)^*T(B_2) + T(B_2)^*T(B_1).$$

In the same way, setting $\lambda_1 = 1$ and $\lambda_2 = \pm i\lambda$ gives

$$T(B_1^*B_2) - T(B_2^*B_1) = T(B_1)^*T(B_2) - T(B_2)^*T(B_1).$$

Therefore

$$T(B_1^*B_2) = T(B_1)^*T(B_2) \quad \text{and} \quad T(B_2^*B_1) = T(B_2)^*T(B_1).$$

Note that these two equalities are actually equivalent, since one can be obtained from the other by taking adjoints. ∎

Proof of Theorem 6.3.1 If T^* satisfies (6.2) then $T : \mathcal{A} \otimes \mathcal{A} \to \mathcal{A}$ satisfies

$$T(A \otimes 1) = T(1 \otimes A) = A \qquad (A \in \mathcal{A}).$$

It follows that for any $A \in \mathcal{A}$

$$T((1 \otimes A)^*(1 \otimes A)) = T(1 \otimes A^*A) = A^*A = T(1 \otimes A)^*T(1 \otimes A),$$

i.e., a Cauchy-Schwarz equality holds for operators of the form $1 \otimes A$. Therefore, Theorem 6.3.5 tells us that for any $A, B \in \mathcal{A}$

$$\begin{aligned} AB &= T(A \otimes 1)T(1 \otimes B) = T\big((A \otimes 1)(1 \otimes B)\big) \\ &= T\big((1 \otimes B)(A \otimes 1)\big) = T(1 \otimes B)T(A \otimes 1) = BA, \end{aligned}$$

which shows that $\mathcal{A}$ is abelian. ∎

Proof of Theorem 6.3.3 Since cloning is possible for states on abelian algebras, one can show that if states on nonabelian algebras could be completely 'written down' in abelian algebras, then they could also be cloned. In this way, it is possible to derive Theorem 6.3.3 from Theorem 6.3.1. However, we will give a shorter, independent proof.

If T^* and S^* satisfy (6.3) then

$$T \circ S(A) = A \qquad (A \in \mathcal{A}).$$

Cauchy-Schwarz gives

$$A^*A = T \circ S(A^*A) \geq T\big(S(A)^*S(A)\big) \geq T\big(S(A)\big)^* T(S(A)) = A^*A \qquad (A \in \mathcal{A}),$$

so we must have two times equality here. In particular, by Theorem 6.3.5, the second equality tells us that $T(S(A)S(B)) = T(S(A))T(S(B))$ for all $A, B \in \mathcal{A}$. Therefore, since $\mathcal{B}$ is abelian,

$$AB = T(S(A))T(S(B)) = T\big(S(A)S(B)\big) = T\big(S(B)S(A)\big) = BA \qquad (A, B \in \mathcal{A}),$$

which shows that $\mathcal{A}$ is abelian. ∎

Proof of Theorem 6.3.4 We take $\mathcal{A} = \mathcal{C} = \mathcal{L}(\mathcal{H})$, where $\mathcal{H}$ is a two-dimensional inner product space with orthonormal basis $\{e_1, e_2\}$. For η, we take the entangled pure state described by the state vector $\chi \in \mathcal{H} \otimes \mathcal{H}$ given by

$$\chi := \tfrac{1}{\sqrt{2}}\, e_1 \otimes e_1 + \tfrac{1}{\sqrt{2}}\, e_2 \otimes e_2.$$

It suffices to prove the theorem for the case that ρ is a pure state. (The general case then follows by linearity.) In that case, the initial state $\rho \otimes \eta \in (\mathcal{A} \otimes \mathcal{C} \otimes \mathcal{C})'_{\rm prob}$ is described by a state vector in $\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}$ of the form

$$y \otimes \chi = (y_1 e_1 + y_2 e_2) \otimes \chi,$$

where y_1, y_2 are complex numbers (unknown to Alice and Bob) such that $|y_1|^2 + |y_2|^2 = 1$. Now Alice performs an ideal measurement on the joint system consisting of the state she wants to send and her particle from the entangled pair. This ideal measurement is described by the partition of the identity $\{P_1, \ldots, P_4\}$, where

$$P_i = |\psi_i\rangle\langle\psi_i| \qquad (i = 1, \ldots, 4),$$

and

$$\begin{aligned}
\psi_1 &:= \tfrac{1}{\sqrt{2}} e_1 \otimes e_1 + \tfrac{1}{\sqrt{2}} e_2 \otimes e_2, \\
\psi_2 &:= \tfrac{1}{\sqrt{2}} e_1 \otimes e_1 - \tfrac{1}{\sqrt{2}} e_2 \otimes e_2, \\
\psi_3 &:= \tfrac{1}{\sqrt{2}} e_1 \otimes e_2 + \tfrac{1}{\sqrt{2}} e_2 \otimes e_1, \\
\psi_4 &:= \tfrac{1}{\sqrt{2}} e_1 \otimes e_2 - \tfrac{1}{\sqrt{2}} e_2 \otimes e_1.
\end{aligned}$$

We rewrite the state $y \otimes \chi$ as

$$\begin{aligned}
y \otimes \chi =& \,(y_1 e_1 + y_2 e_2) \otimes \big(\tfrac{1}{\sqrt{2}}\, e_1 \otimes e_1 + \tfrac{1}{\sqrt{2}}\, e_2 \otimes e_2\big) \\
=& \,\tfrac{1}{\sqrt{2}} y_1 e_1 \otimes e_1 \otimes e_1 + \tfrac{1}{\sqrt{2}} y_1 e_1 \otimes e_2 \otimes e_2 \\
& + \tfrac{1}{\sqrt{2}} y_2 e_2 \otimes e_1 \otimes e_1 + \tfrac{1}{\sqrt{2}} y_2 e_2 \otimes e_2 \otimes e_2 \\
=& \,y_1 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_1 + \tfrac{1}{\sqrt{2}} e_2 \otimes e_2\big) \otimes e_1 + y_1 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_1 - \tfrac{1}{\sqrt{2}} e_2 \otimes e_2\big) \otimes e_1 \\
& + y_1 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_2 + \tfrac{1}{\sqrt{2}} e_2 \otimes e_1\big) \otimes e_2 + y_1 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_2 - \tfrac{1}{\sqrt{2}} e_2 \otimes e_1\big) \otimes e_2 \\
& + y_2 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_2 + \tfrac{1}{\sqrt{2}} e_2 \otimes e_1\big) \otimes e_1 - y_2 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_2 - \tfrac{1}{\sqrt{2}} e_2 \otimes e_1\big) \otimes e_1 \\
& + y_2 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_1 + \tfrac{1}{\sqrt{2}} e_2 \otimes e_2\big) \otimes e_2 - y_2 \tfrac{1}{2}\big(\tfrac{1}{\sqrt{2}} e_1 \otimes e_1 - \tfrac{1}{\sqrt{2}} e_2 \otimes e_2\big) \otimes e_2 \\
=& \,\tfrac{1}{2}\psi_1 \otimes \big(y_1 e_1 + y_2 e_2\big) + \tfrac{1}{2}\psi_2 \otimes \big(y_1 e_1 - y_2 e_2\big) \\
& + \tfrac{1}{2}\psi_3 \otimes \big(y_1 e_2 + y_2 e_1\big) + \tfrac{1}{2}\psi_4 \otimes \big(y_1 e_2 - y_2 e_1\big).
\end{aligned}$$

The measurement of Alice produces with equal probabilities any of the outcomes $1, \ldots, 4$. If the outcome is 1, then Bob's particle of the entangled pair is in the state $y_1 e_1 + y_2 e_2$. If the outcome of Alice's measurement is 2, then Bob's particle is in the state $y_1 e_1 - y_2 e_2$, and so on.

If Bob learns about the outcome of Alice's experiment, then he perform a deterministic operation on his particle from the entangled pair, so that after this operation, this particle is in the state y that Alice wanted to send. (Note that y is still unknown to Alice and Bob!) If the outcome of Alice's experiment is i, then Bob performs the deterministic operation described by the unitary map U_i, where the matrices of $U_1, \ldots, U_4 \in \mathcal{L}(\mathcal{H})$ with respect to the basis $\{e_1, e_2\}$ are given by

$$U_1 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad U_2 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad U_3 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad U_4 := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

For example, if $i = 4$, then Bob's particle is in the state $(y_1 e_2 - y_2 e_1)$, which under the unitary transformation U_4 becomes

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} -y_2 \\ y_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

It is easy to see that in any of the four cases, Bob ends up with a particle in the pure state $y_1 e_1 + y_2 e_2$, which is the state Alice wanted to send.

More formally, the operations of Alice and Bob can be described as follows.

Alice's operation is a map $T^* : \mathcal{L}(\mathcal{H} \otimes \mathcal{H})' \to \mathcal{B}'$. Here $\mathcal{B}$ is of the form $\mathcal{B} \cong \mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C} = \{b = (b_1, \ldots, b_4) : b_i \in \mathbb{C}\}$ and $\mathcal{A} \otimes \mathcal{C} \cong \mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{H}) \cong \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$. Let π_i denote the pure state on $\mathcal{B}$ defined as $\pi_i(b) := b_i$. Then $T^* : \mathcal{L}(\mathcal{H} \otimes \mathcal{H})' \to \mathcal{B}'$ maps the pure state described by the state vector ψ_i to the pure state π_i. It is not hard to see that this is achieved by the operator $T : \mathcal{B} \to \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$ given by

$$T(b_1, \ldots, b_4) = \sum_{i=1}^{4} V_i b_i V_i^*,$$

where $V_i : \mathbb{C} \to (\mathcal{H} \otimes \mathcal{H})$ is defined as

$$V_i := |\psi_i\rangle \qquad (i = 1, \ldots, 4).$$

It follows from Theorem 5.3.9 that T is completely positive. Moreover, $T(1) = \sum_i |\psi_i\rangle\langle\psi_i| = 1$.

Bob's operation is a map $S^* : (\mathcal{B} \otimes \mathcal{C})' \to \mathcal{C}'$. Here $\mathcal{C} \cong \mathcal{L}(\mathcal{H})$ and $\mathcal{B} \otimes \mathcal{C} \cong \big(\mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C}\big) \otimes \mathcal{L}(\mathcal{H}) \cong \mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{H})$. It is not hard to see that Bob's operation is achieved by the operator $S : \mathcal{C} \to (\mathcal{B} \otimes \mathcal{C})$ given by

$$S(A) = \big(U_1^* A U_1, \ldots, U_4^* A U_4\big) \qquad (A \in \mathcal{L}(\mathcal{H})),$$

where $U_1, \ldots, U_n$ are the unitary operators defined above. Comparing this with Proposition 5.3.11 we see that S^* is a deterministic operation.

Define $(1 \otimes \eta) : \mathcal{A} \otimes \mathcal{C} \otimes \mathcal{C} \to \mathcal{A}$ by

$$\eta(A \otimes B) := \eta(B) A \qquad (A \in \mathcal{A},\ B \in \mathcal{C} \otimes \mathcal{C}).$$

Then the operations T^* and S^* and the state η satisfy (6.4) for all $\rho \in \mathcal{A}'_{\rm prob}$ if and only if

$$\rho \circ (1 \otimes \eta) \circ (T \otimes 1) \circ S(A) = \rho(A) \qquad (A \in \mathcal{A},\ \rho \in \mathcal{A}'_{\rm prob}),$$

which is equivalent to

$$(1 \otimes \eta) \circ (T \otimes 1) \circ S(A) = \rho(A) \qquad (A \in \mathcal{A}).$$

This formula can be checked by straightforward calculation. In fact, we have already seen all the essential ingredients of this calculation in our informal discussion of Alice's and Bob's operations, so we do not go into details. ∎

6.4 Quantum cryptography

In the previous section, we saw that quantum probability leads to some surprising impossibilities: 'no cloning' and 'no coding'. In view of these impossibilities, the possibility of 'quantum teleportation' is surprising, but from the classical point of view, where copying is possible, this is nothing new. In this section, we will see that the peculiarities of quantum probability also open some new possibilities that are not present in classical probability.

Suppose that Alice wants to send a (classical) message to Bob, say, a sequence $(x_1, \ldots, x_n)$ of zeros and ones, while making sure that a third party, called *Eve*, is not eavesdropping. Alice can do this if she has another, random sequence $(y_1, \ldots, y_n)$ of zeros and ones, that is known to her and Bob, but to no one else. With such a sequence, she simply sends Bob the sequence $(z_1, \ldots, z_n)$ given by

$$z_i = x_i + y_i \mod 1.$$

To anyone who does not know the code $(y_1, \ldots, y_n)$, the message $(z_1, \ldots, z_n)$ is just a random sequence of zeros and ones, but Bob, who knows the code, calculates

$$x_i = z_i + y_i \mod 1$$

to get Alice's message $(x_1, \ldots, x_n)$.

Now Alice and Bob may have prepared their code before, when they were together, but since each code may be used only once[1] they may run out of code after a while. This leads to the following question: can Alice send Bob a secret code of independent zeros and ones, while being sure that Eve is not listening somewhere on the telephone line?

In classical probability, this is not possible, since Eve can perform a nonperturbing measurement on the signal passing through the telephone line. But in quantum

[1] If the same code is use repeatedly, then the coded messages are no longer sequences of independent random variables, and this dependence may be used to crack the code.

probability, the situation is quite different. In [J–Z00], a team around professor Zeilinger from Vienna shows how Alice can send a code to Bob while making sure that Eve is not eavesdropping. What is more, they report on an experiment that shows that this form of communication is possible in practice. On April 27, 2004, the *Süddeutsche Zeitung* reported how in Vienna, 3000 euro were transferred from the town hall to the bank account of the university, using this form of quantum cryptography.

This is how it works. Alice prepares pairs entangled photons, of which she keeps one for herself, and sends the other one to Bob through a glass fiber cable. On the photons that Alice keeps for herself, she performs at random, with equal probabilities, either the ideal measurement $\{P_0, P_{\pi/2}\}$ or the ideal measurement $\{P_\gamma, P_{\gamma+\pi/2}\}$, where P_γ is the projection defined on page 112 and γ is an angle that we will choose later. Likewise, Bob performs on his photons with equal probabilities either the measurement $\{P_0, P_{\pi/2}\}$ or $\{P_{-\gamma}, P_{-\gamma+\pi/2}\}$. After sending as many photons as she needs, Alice tells Bob over a regular telephone line which measurements she used for her photons. Bob then tells Alice which measurements he used, and he tells the outcomes of all his measurements in those cases where they did not both perform the measurement $\{P_0, P_{\pi/2}\}$. As we will see in a moment, Alice can see from this information whether Eve was eavesdropping or not. If she sees that Eve was not tapping the phone, Alice sends Bob the message that she wanted to send, using as a code the outcome of those measurements where she and Bob both performed the measurement $\{P_0, P_{\pi/2}\}$.

Let ψ be the entangled state from page 112 and let

$$\pi_{\alpha,\beta}(+,+) := \rho_\psi(P_\alpha \otimes P_\beta)$$

denote the probability that Alice and Bob do the observations P_α and P_β, if they perform the ideal measurements $\{P_\alpha, P_{\alpha+\pi/2}\}$ and $\{P_\beta, P_{\beta+\pi/2}\}$, respectively. Likewise, let

$$\begin{aligned}
\pi_{\alpha,\beta}(+,-) &:= \rho_\psi(P_\alpha \otimes P_{\beta+\pi/2}),\\
\pi_{\alpha,\beta}(-,+) &:= \rho_\psi(P_{\alpha+\pi/2} \otimes P_\beta),\\
\pi_{\alpha,\beta}(-,-) &:= \rho_\psi(P_{\alpha+\pi/2} \otimes P_{\beta+\pi/2}),
\end{aligned}$$

denote the probabilities that one of them, or both, perform the complementary observation. We calculated on page 112 that

$$\begin{aligned}
\pi_{\alpha,\beta}(+,+) = \pi_{\alpha,\beta}(-,-) = \tfrac{1}{2}\cos(\beta-\alpha)^2,\\
\pi_{\alpha,\beta}(+,-) = \pi_{\alpha,\beta}(-,+) = \tfrac{1}{2}\sin(\beta-\alpha)^2.
\end{aligned}$$

In particular, $\pi_{0,0}(+,+) = \pi_{0,0}(-,-) = \frac{1}{2}$ and $\pi_{0,0}(+,-) = \pi_{0,0}(-,+) = 0$, so the pairs of photons on which Alice and Bob both performed the measurement $\{P_0, P_{\pi/2}\}$ can be used as a secret code.

If Eve wants to find out this secret code, she has no other choice than to perform a measurement on all photons that pass through the glass fiber cable on their way to Bob, since she cannot know in advance which photons are going to be used for the secret code. We claim that if Eve fiddles with these photons in any way such that she gets to know the secret code, then she cannot avoid changing the probabilities in such a way that

$$\pi_{\gamma,0}(+,-) + \pi_{0,-\gamma}(+,-) - \pi_{\gamma,-\gamma}(+,-) \geq 0. \tag{6.5}$$

This is called *Wigner's inequality*. It follows from the assumptions of anticorrelations and nonentanglement:

Lemma 6.4.1 (Wigner's inequality). *Assume that probabilities $\pi_{\alpha,\beta}(\pm,\pm)$ satisfy $\pi_{0,0}(+,-) = \pi_{0,0}(-,+) = 0$ and*

$$\pi_{\alpha,\beta}(\sigma,\tau) = \sum_k p_k\, \pi^{1,k}_\alpha(\sigma)\pi^{2,k}_\beta(\tau) \qquad (\sigma,\tau = +,-),$$

where the $\pi^{i,k}_\alpha(\pm)$ are nonnegative numbers such that $\pi^{i,k}_\alpha(+) + \pi^{i,k}_\alpha(-) = 1$ and the p_k are positive numbers summing up to one. Then (6.5) holds.

Proof Write

$$\pi_{\alpha,\beta}(\sigma,\tau) = \sum_k p_k\, \pi^k_{\alpha,\beta}(\sigma,\tau),$$

where $\pi^k_{\alpha,\beta}(\sigma,\tau) = \pi^{1,k}_\alpha(\sigma)\pi^{2,k}_\beta(\tau)$. Since $\pi_{0,0}(+,-) = \pi_{0,0}(-,+) = 0$ we must have that for each k, either $\pi^{1,k}_0(+) = \pi^{2,k}_0(+) = 1$ or $\pi^{1,k}_0(-) = \pi^{2,k}_0(-) = 1$. In the first case,

$$\begin{aligned}
&\pi^k_{\gamma,0}(+,-) + \pi^k_{0,-\gamma}(+,-) - \pi^k_{\gamma,-\gamma}(+,-) \\
&\quad = \pi^{1,k}_\gamma(+)\cdot 0 + 1\cdot \pi^{2,k}_{-\gamma}(-) - \pi^{1,k}_\gamma(+)\pi^{2,k}_{-\gamma}(-) \\
&\quad = \big(1 - \pi^{1,k}_\gamma(+)\big)\pi^{2,k}_{-\gamma}(-) \geq 0,
\end{aligned}$$

while in the second case

$$\pi^k_{\gamma,0}(+,-) + \pi^k_{0,-\gamma}(+,-) - \pi^k_{\gamma,-\gamma}(+,-) = \pi^{1,k}_\gamma(+)\big(1 - \pi^{2,k}_{-\gamma}(-)\big) \geq 0.$$

Summing up over k we arrive at (6.5). ∎

We claim that Eve's measurement of the secret code necessarily destroys the entanglement of the two photons. Indeed, if Eve wants to be sure that she gets the same code as Alice, she has no other option than to perform the ideal measurement $\{P_0, P_{\pi/2}\}$ on the photon that is on its way to Bob. If she also wants Bob to recieve the secret code, she must send the photon on to Bob after she has performed her measurement, or she must send another photon that is polarized in the direction that she measured. In any case, in doing so, she will have changed the system from the pure state ρ_ψ with

$$\psi = \tfrac{1}{\sqrt{2}}\,\phi_1 \otimes \phi_1 + \tfrac{1}{\sqrt{2}}\,\phi_2 \otimes \phi_2$$

to the mixed state

$$\tfrac{1}{2}\rho_{\phi_1 \otimes \phi_1} + \tfrac{1}{2}\rho_{\phi_2 \otimes \phi_2},$$

which is not entangled, and therefore satisfies Wigner's inequality.

We now know how Alice can find out from the information that Bob sends her whether Eve has been eavesdropping. If Eve has not interfered with the signal, the quantity in Wigner's inequality is

$$\tfrac{1}{2}\sin(\gamma)^2 + \tfrac{1}{2}\sin(\gamma)^2 - \tfrac{1}{2}\sin(2\gamma)^2,$$

which reaches a minimal value of $-\frac{1}{8}$ at $\gamma = \frac{1}{6}\pi$. From the relative frequencies of outcomes in those measurements where she and Bob did not both perform the measurement $\{P_0, P_{\pi/2}\}$, Alice can check whether Wigner's inequality is violated. Indeed, Wigner's inequality holds also with $+$ and $-$ reversed, so for greater statistical precision, Alice checks whether the quantity

$$\begin{aligned}&\pi_{\gamma,0}(+,-) + \pi_{0,-\gamma}(+,-) - \pi_{\gamma,-\gamma}(+,-)\\ &\quad + \pi_{\gamma,0}(-,+) + \pi_{0,-\gamma}(-,+) - \pi_{\gamma,-\gamma}(-,+)\end{aligned}$$

is close enough to $-\frac{1}{4}$. If she is satisfied with the answer, she knows that the outcomes of the experiments where she and Bob both performed the measurement $\{P_0, P_{\pi/2}\}$ are not known to Eve, and she uses these as a secret code to send her message to Bob.

Exercise 6.4.2 Eve decides to use the following tactic: She cuts the glass fiber between Alice and Bob, blocking all direct communication between them. Instead, Eve communicates now with both Alice and Bob, pretending to be Bob when she communicates with Alice, and pretending to be Alice when she communicates with Bob. In this way, she builds up a secret code with Alice and another secret code with Bob. When Alice sends the coded signal, Eve decodes it using the code she shares with Alice and then codes it with the code she shares with Bob, before passing it on to Bob. Can Alice and Bob do anything to detect this kind of eavesdropping?

German-English glossary

Chinese remainder theorem	Chinesischer Restsatz
conditional probability	bedingte Wahrscheinlichkeit
density	Dichte
entanglement	Verschränktheit
event	Ereignis
expectation	Erwartung
faithful representation	treue Darstellung
field	Körper
fully reducible	völlig auflösbar
functional calculus	Funktionalkalkül
Heisenberg picture	Heisenbergbild
hermitian	hermitesch
irreducible	irreduzibel
mixed state	gemischter Zustand
momentum	Impuls
multiplicity	Vielfachheit
observation	Beobachtung
operation	Operation, Eingriff
physical quantity	physikalische Größe
probability	Wahrscheinlichkeit
proper subspace	echter Unterraum
pseudotrace	verallgemeinerte Spur
pure state	reiner Zustand
reducible	auflösbar
relative frequencies	relative Häufigkeiten
representation	Darstellung
semisimple	halbeinfach
simple algebra	einfache Algebra
span	Spann, Einhüllende
state	Zustand
super selection rule	Superauswahlregel
tensor calculus	Tensorrechnung
time evolution	Zeitentwicklung
trace	Spur
wave function	Wellenfunktion

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.

[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.

[CS78] J.F. Clauser and A. Shimony. Bell's theorem: experimental tests and implications. *Reports on Progress in Physics* 41, 1881-1927, 1978.

[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.

[Tak79] M. Takesaki. *Theory of Operator Algebras I.* Springer, New York, 1979.

Index