Chapter 2
Vector and Hilbert Spaces

2.1 Introduction

The purpose of this chapter is to introduce Hilbert spaces, and more precisely the *Hilbert spaces* on the field of complex numbers, which represent the abstract environment in which Quantum Mechanics is developed.

To arrive at Hilbert spaces, we proceed gradually, beginning with spaces mathematically less structured, to move toward more and more structured ones, considering, in order of complexity:

(1) *linear* or *vector* spaces, in which the points of the space are called *vectors*, and the operations are the *sum* between two vectors and the *multiplication by a scalar*;
(2) *normed vector spaces*, in which the concept of *norm* of a vector x is introduced, indicated by $||x||$, from which one can obtain the distance between two vectors x and y as $d(x, y) = ||x - y||$;
(3) *vector spaces with inner product*, in which the concept of *inner product* between two vectors x, y is introduced, and indicated in the form (x, y), from which the norm can be obtained as $||x|| = (x, x)^{1/2}$, and then also the distance $d(x, y)$;
(4) *Hilbert spaces*, which are vector spaces with inner product, with the additional property of *completeness*.

We will start from vector spaces, then we will move on directly to vector spaces with inner product and, eventually, to Hilbert spaces. For vectors, we will initially adopt the standard notation (x, y, etc.), and subsequently we will switch to Dirac's notation, which has the form $|x\rangle$, $|y\rangle$, etc., universally used in Quantum Mechanics.

G. Cariolaro, *Quantum Communications*, Signals and Communication Technology,
DOI 10.1007/978-3-319-15600-2_2

2.2 Vector Spaces

2.2.1 Definition of Vector Space

A vector space on a field $\mathbb{F}$ is essentially an Abelian group, and therefore a set provided with the addition operation $+$, but completed with the operation of multiplication by a scalar belonging to $\mathbb{F}$.

Here we give the definition of vector space in the field of complex numbers $\mathbb{C}$, as it is of interest to Quantum Mechanics.

Definition 2.1 A *vector space* in the field of complex numbers $\mathbb{C}$ is a nonempty set $\mathcal{V}$, whose elements are called *vectors*, for which two operations are defined. The first operation, *addition*, is indicated by $+$ and assigns to each pair $(x, y) \in \mathcal{V} \times \mathcal{V}$ a vector $x + y \in \mathcal{V}$. The second operation, called multiplication by a scalar or simply *scalar multiplication*, assigns to each pair $(a, x) \in \mathbb{C} \times \mathcal{V}$ a vector $ax \in \mathcal{V}$. These operations must satisfy the following properties, for $x, y, z \in \mathcal{V}$ and $a, b \in \mathbb{C}$:

(1) $x + (y + z) = (x + y) + z$ (associative property),
(2) $x + y = y + x$ (commutative property),
(3) $\mathcal{V}$ contains an identity element 0 with the property $0 + x = x, \ \forall x \in \mathcal{V}$,
(4) $\mathcal{V}$ contains the opposite (or inverse) vector $-x$ such that $-x + x = 0, \ \forall x \in \mathcal{V}$,
(5) $a(x + y) = ax + ay$,
(6) $(a + b)x = ax + bx$. □

Notice that the first four properties assure that $\mathcal{V}$ is an *Abelian group* or commutative group, and, globally, the properties make sure that every linear combination

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n \qquad a_i \in \mathbb{C}, \ x_i \in \mathcal{V}$$

is also a vector of $\mathcal{V}$.

2.2.2 Examples of Vector Spaces

A first example of a vector space on $\mathbb{C}$ is given by $\mathbb{C}^n$, that is, by the set of the n-tuples of complex numbers,

$$x = (x_1, x_2, \ldots, x_n) \quad \text{with} \quad x_i \in \mathbb{C}$$

where scalar multiplication and addition must be intended in the usual sense, that is,

$$ax = (ax_1, ax_2, \ldots, ax_n), \qquad \forall a \in \mathbb{C}$$

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).$$

A second example is given by the sequence of complex numbers

$$x = (x_1, x_2, \dots, x_i, \dots) \quad \text{with} \quad x_i \in \mathbb{C}.$$

In the first example, the vector space is *finite dimensional*, in the second, it is *infinite dimensional* (further on, the concept of dimension of a vector space will be formalized in general).

A third example of vector space is given by the class of continuous-time or discrete-time, and also multidimensional, signals (complex functions). We will return to this example with more details in the following section.

2.2.3 Definitions on Vector Spaces and Properties

We will now introduce the main definitions and establish a few properties of vector spaces, following Roman's textbook [1].

Vector Subspaces

A nonempty subset $\mathcal{S}$ of a vector space $\mathcal{V}$, itself a vector space provided with the same two operations on $\mathcal{V}$, is called a *subspace* of $\mathcal{V}$. Therefore, by definition, $\mathcal{S}$ is *closed* with respect to the linear combinations of vectors of $\mathcal{S}$.

Notice that $\{0\}$, where 0 is the identity element of $\mathcal{V}$, is a subspace of $\mathcal{V}$.

Generator Sets and Linear Independence

Let $\mathcal{S}_0$ be a nonempty subset of $\mathcal{V}$, not necessarily a subspace; then the set of all the linear combinations of vectors of $\mathcal{S}_0$ *generates a subspace* $\mathcal{S}$ of $\mathcal{V}$, indicated in the form

$$\mathcal{S} = \text{span}\,(\mathcal{S}_0) = \{a_1 x_1 + a_2 x_2 + \cdots + a_n x_n \mid a_i \in \mathbb{C},\ x_i \in \mathcal{S}_0\}. \tag{2.1}$$

In particular, the generator set $\mathcal{S}_0$ can consist of a single point of $\mathcal{V}$. For example, in $\mathbb{C}^2$, the set $\mathcal{S}_0 = \{(1, 2)\}$ consisting of the vector $(1, 2)$, generates $\mathcal{S} = \text{span}\,(\mathcal{S}_0) = \{a(1, 2)|a \in \mathbb{C}\} = \{(a, 2a)|a \in \mathbb{C}\}$, which represents a straight line passing through the origin (Fig. 2.1); it can be verified that $\mathcal{S}$ is a subspace of $\mathbb{C}^2$. The set $\mathcal{S}_0 = \{(1, 2), (3, 0)\}$ generates the entire $\mathbb{C}^2$, that is,[1]

$$\text{span}\,((1, 2), (3, 0)) = \mathbb{C}^2.$$

The concept of linear independence of a vector space is the usual one. A set $\mathcal{S}_0 = \{x_1, x_2, \dots, x_n\}$ of vectors of $\mathcal{V}$ is *linearly independent*, if the equality

[1] If $\mathcal{S}_0$ is constituted by some points, for example $\mathcal{S}_0 = \{x_1, x_2, x_3\}$, the notation $\text{span}(\mathcal{S}_0) = \text{span}(\{x_1, x_2, x_3\})$ is simplified to $\text{span}(x_1, x_2, x_3)$.

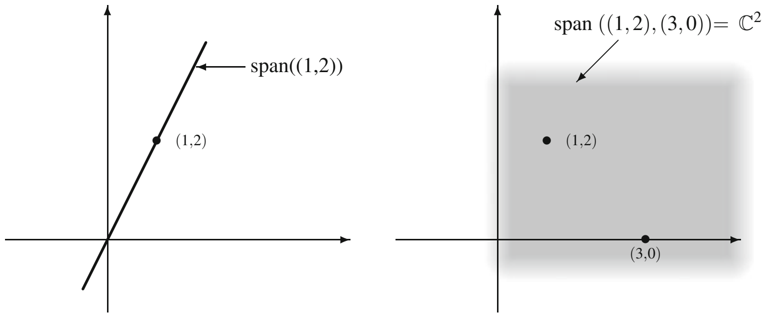

Fig. 2.1 The set $\{(1, 2)\}$ of $\mathbb{C}^2$ generates a straight line trough the origin, while the set $\{(3, 0), (1, 2)\}$ generates $\mathbb{C}^2$ (for graphical reason the representation is limited to $\mathbb{R}^2$)

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0 \tag{2.2}$$

implies

$$a_1 = 0,\ a_2 = 0,\ \ldots,\ a_n = 0.$$

Otherwise, the set is *linearly dependent*. For example, in $\mathbb{C}^2$ the set $\{(1, 2), (0, 3)\}$ is constituted by two linear independent vectors, whereas the set $\{(1, 2), (2, 4)\}$ is linearly dependent because

$$a_1(1, 2) + a_2(2, 4) = (0, 0) \quad \text{for} \quad a_1 = 2 \quad \text{e} \quad a_2 = -1.$$

2.2.4 Bases and Dimensions of a Vector Space

A subset $\mathcal{B}$ of a vector space $\mathcal{V}$ constituted by linearly independent vectors is *a basis of* $\mathcal{V}$ if $\mathcal{B}$ generates $\mathcal{V}$, that is, if two conditions are met:

(1) $\mathcal{B} \subset \mathcal{V}$ is formed by linearly independent vectors,
(2) span $(\mathcal{B}) = \mathcal{V}$.

It can be proved that [1, Chap.1]:

(a) Every vector space $\mathcal{V}$, except the degenerate space $\{0\}$, admits a basis $\mathcal{B}$.
(b) If $b_1, b_2, \ldots, b_n$ are vectors of a basis $\mathcal{B}$ of $\mathcal{V}$, the linear combination

$$a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = x \tag{2.3}$$

is unique, i.e., the coefficients $a_1, a_2, \ldots, a_n$, are uniquely identified by x.
(c) All the bases of a vector space have the same *cardinality*. Therefore, if $\mathcal{B}_1$ and $\mathcal{B}_2$ are two bases of $\mathcal{V}$, it follows that $|\mathcal{B}_1| = |\mathcal{B}_2|$.

The property (c) is used to define the *dimension* of a vector space $\mathcal{V}$, letting

$$\dim \mathcal{V} := |\mathcal{B}|. \tag{2.4}$$

Then the dimension of a vector space is given by the common cardinality of its bases. In particular, if $\mathcal{B}$ is finite, the vector space $\mathcal{V}$ is of *finite dimension*; otherwise $\mathcal{V}$ is of *infinite dimension*.

In $\mathbb{C}^n$ the *standard basis* is given by the n vectors

$$(1, 0, \ldots, 0),\ (0, 1, \ldots, 0),\ \ldots,\ (0, 0, \ldots, 1). \tag{2.5}$$

Therefore, $\dim \mathbb{C}^n = n$. We must observe that in $\mathbb{C}^n$ there are infinitely many other bases, all of cardinality n.

In the vector space consisting of the sequences $(x_1, x_2, \ldots)$ of complex numbers, the standard basis is given by the vectors

$$(1, 0, 0, \ldots),\ (0, 1, 0, \ldots),\ (0, 0, 1, \ldots),\ \ldots \tag{2.6}$$

which are infinite. Therefore this space is of infinite dimension.

2.3 Inner-Product Vector Spaces

2.3.1 Definition of Inner Product

In a vector space $\mathcal{V}$ on complex numbers, the *inner product*, here indicated by the symbol $\langle\cdot,\ \cdot\rangle$, is a function

$$\langle\cdot,\ \cdot\rangle\ :\ \mathcal{V}\times\mathcal{V}\ \to\ \mathbb{C}$$

with the following properties, for $x, y, z \in \mathcal{V}$ and $a, b \in \mathbb{C}$:

(1) it is a positive definite function, that is,

$$\langle x, x\rangle \geq 0 \quad \text{and} \quad \langle x, x\rangle = 0 \quad \text{if and only if} \quad x = 0;$$

(2) it enjoys the Hermitian symmetry

$$\langle x, y\rangle = \langle y, x\rangle^*;$$

(3) it is linear with respect to the first argument

$$\langle ax + by, z\rangle = a\langle x, z\rangle + b\langle y, z\rangle.$$

□

From properties (2) and (3) it follows that with respect to the second argument the so-called *conjugate linearity* holds, namely

$$\langle z, ax + by\rangle = a^*\langle z, x\rangle + b^*\langle z, y\rangle.$$

We observe that within the same vector space $\mathcal{V}$ it is possible to introduce different inner products, and the choice must be made according to the application of interest.

2.3.2 Examples

In $\mathbb{C}^n$, the standard form of inner product of two vectors $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ is defined as follows:

$$\langle x, y\rangle = x_1\, y_1^* + \cdots + x_n\, y_n^* = \sum_{i=1}^{n} x_i\, y_i^* \tag{2.7a}$$

and it can be easily seen that such expression satisfies the properties (1), (2), and (3). Interpreting the vectors $x \in \mathbb{C}^n$ as column vectors ($n \times 1$ matrices), and indicating with y^* the conjugate transpose of y ($1 \times n$ matrix), that is,

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad y^* = \begin{bmatrix} y_1^*, & \ldots, & y_n^* \end{bmatrix} \tag{2.7b}$$

and applying the usual matrix product, we obtain

$$\langle x, y\rangle = y^* x = x_1\, y_1^* + \cdots + x_n\, y_n^* \tag{2.7c}$$

a very handy expression for algebraic manipulations.

The most classic example of infinite-dimensional inner-product vector space, introduced by Hilbert himself, is the space ℓ_2 of the square-summable complex sequences $x = (x_1, x_2, \ldots)$, that is, with

$$\sum_{i=1}^{\infty} |x_i|^2 < \infty \tag{2.8}$$

where the standard inner product is defined by

$$\langle x, y\rangle = \sum_{i=1}^{\infty} x_i\, y_i^* = \lim_{n\to\infty} \sum_{i=1}^{n} x_i\, y_i^*.$$

The existence of this limit is ensured by Schwartz's inequality (see (2.12)), where (2.8) is used.

Another example of inner-product vector space is given by the continuous functions over an interval $[a,\ b]$, where the standard inner product is defined by

$$\langle x,\ y\rangle = \int_a^b x(t)\, y^*(t)\, \mathrm{d}t.$$

2.3.3 Examples from Signal Theory

These examples are proposed because they will allow us to illustrate some concepts on vector spaces, in view of the reader's familiarity with the subject.

We have seen that the class of signals $s(t)$, $t \in I$, defined on a domain I, form a vector space. If we limit ourselves to the signals $\mathcal{L}_2(I)$, for which it holds that [2]

$$\int_I \mathrm{d}t\, |s(t)|^2 < \infty, \tag{2.9}$$

we can obtain a space with inner product defined by

$$\langle x,\ y\rangle = \int_I \mathrm{d}t\, x(t)\, y^*(t) \tag{2.10}$$

which verifies conditions (1), (2), and (3).

A first concept that can be exemplified through signals is that of a *subspace*. In the space $\mathcal{L}_2(I)$, let us consider the subspace $\mathcal{E}(I)$ formed by the *even* signals. Is $\mathcal{E}(I)$ a subspace? The answer is yes, because every linear combination of even signals is an even signal: therefore $\mathcal{E}(I)$ is a subspace of $\mathcal{L}_2(I)$. The same conclusion applies to the class $\mathcal{O}(I)$ of odd signals. These two subspaces are illustrated in Fig. 2.2.

2.3.4 Norm and Distance. Convergence

From the inner product it is possible to define the norm $||x||$ of a vector $x \in \mathcal{V}$ through the relation

$$||x|| = \sqrt{\langle x,\ x\rangle}. \tag{2.11}$$

Intuitively, the norm may be thought of as representing the *length* of the vector. A vector with unit norm $||x|| = 1$, is called *unit vector* (we anticipate that in Quantum Mechanics only unit vectors are used). In terms of inner product and norm, we can

[2] To proceed in unified form, valid for all the classes $\mathcal{L}_2(I)$, we use the Haar integral (see [2]).

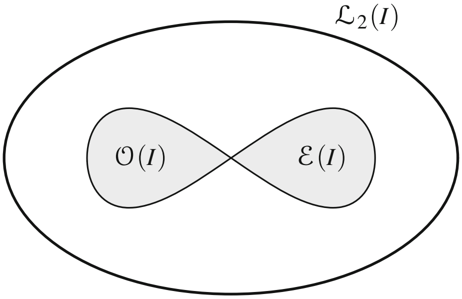

Fig. 2.2 Examples of *vector subspaces* of the signal class $\mathcal{L}_2(I)$

write the important *Schwartz's inequality*

$$|\langle x,\ y\rangle| \le ||x||\,||y|| \tag{2.12}$$

where the equal sign holds if and only if y is proportional to x, that is, $y = kx$ for an appropriate $k \in \mathbb{C}$.

From (2.11) it follows that an inner-product vector space is also a normed space, with the norm introduced by the inner product.

In an inner-product vector space we can also introduce the distance $d(x, y)$ between two points $x, y \in \mathcal{V}$, through the relation

$$d(x, y) = ||x - y|| \tag{2.13}$$

and we can verify that this parameter has the properties required by distance in metric spaces, in particular the triangular inequality holds

$$d(x, y) \le d(x, z) + d(y, z). \tag{2.14}$$

So an inner-product vector space is also a metric space.

Finally, the inner product allows us to introduce the concept of *convergence*. A sequence $\{x_n\}$ of vectors of $\mathcal{V}$ *converges* to the vector x if

$$\lim_{n\to\infty} d(x_n, x) = \lim_{n\to\infty} ||x_n - x|| = 0. \tag{2.15}$$

Now, suppose that a sequence $\{x_n\}$ has the property (*Cauchy's sequence* or *fundamental sequence*)

$$d(x_m, x_n) \to 0 \quad \text{for} \quad m, n \to \infty.$$

In general, for such a sequence, the limit (2.15) is not guaranteed to exist, and, if it exists, it is not guaranteed that the limit x is a vector of $\mathcal{V}$. So, an inner-product vector space in which all the Cauchy sequences converge to a vector of $\mathcal{V}$ is said to be *complete*. At this point we have all we need to define a Hilbert space.

2.4 Definition of Hilbert Space

Definition 2.2 A Hilbert space is a *complete* inner-product vector space.

It must be observed that a *finite* dimensional vector space is always complete, as it is closed with respect to all its sequences, and therefore it is always a Hilbert space. Instead, if the space is *infinite* dimensional, the completeness is not ensured, and therefore it must be added as a hypothesis, in order for the inner-product vector space to become a Hilbert space.

At this point, we want to reassure the reader: the theory of optical quantum communications will be developed at a level that will not fully require the concept of a Hilbert space, but the concept of inner-product vector space will suffice. Nonetheless, the introduction of the Hilbert space is still done here for consistency with the Quantum Mechanics literature.

From now on, we will assume to operate on a Hilbert space, but, for what we just said, we can refer to an inner-product vector space.

2.4.1 Orthogonality, Bases, and Coordinate Systems

In a Hilbert space, the basic concepts, introduced for vector spaces, can be expressed by using orthogonality.

Let $\mathcal{H}$ be a Hilbert space. Then two vectors $x, y \in \mathcal{H}$ are *orthogonal* if

$$\langle x, \ y \rangle = 0. \tag{2.16}$$

Extending what was seen in Sect. 2.2, we have that a Hilbert space admits orthogonal bases, where each basis

$$\mathcal{B} = \{b_i, i \in I\} \tag{2.17}$$

is formed by pairwise orthogonal vectors, that is,

$$\langle b_i, b_j \rangle = 0 \qquad i, j \in I, \ i \neq j$$

and furthermore, $\mathcal{B}$ generates $\mathcal{H}$

$$\text{span}(\mathcal{B}) = \mathcal{H}.$$

The set I in (2.17) is finite, $I = \{1, 2, \ldots, n\}$, or countably infinite, $I = \{1, 2, \ldots\}$, and may even be a continuum (but not considered in this book until Chap. 11).

Remembering that a vector b is a *unit vector* if $||b||^2 = \langle b, b\rangle = 1$, a basis becomes *orthonormal*, if it is formed by unit vectors. The *orthonormality condition* of a basis can be written in the compact form

$$\boxed{\langle b_i, b_j\rangle = \delta_{ij},} \tag{2.18}$$

where δ_{ij} is Kronecker's symbol, defined as $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$. In general, a Hilbert space admits infinite orthonormal bases, all, obviously, with the same cardinality.

For a fixed orthonormal basis $\mathcal{B} = \{b_i, i \in I\}$, every vector x of $\mathcal{H}$ can be uniquely written as a linear combination of the vectors of the basis

$$\boxed{x = \sum_{i \in I} a_i\, b_i} \tag{2.19}$$

where the coefficients are given by the inner products

$$\boxed{a_i = \langle x, b_i\rangle.} \tag{2.20}$$

In fact, we obtain

$$\langle x, b_j\rangle = \left\langle \sum_i a_i\, b_i, b_j \right\rangle = \sum_i a_i\, \langle b_i, b_j\rangle = a_j$$

where in the last equality we used orthonormality condition (2.18).

The expansion (2.19) is called *Fourier expansion* of the vector x and the coefficients a_i the *Fourier coefficients* of x, obtained with the basis $\mathcal{B}$.

Through Fourier expansion, every orthonormal basis $\mathcal{B} = \{b_i, i \in I\}$ defines a *coordinate system* in the Hilbert space. In fact, according to (2.19) and (2.20), a vector x uniquely identifies its Fourier coefficients $\{a_i, i \in I\}$, which are the *coordinates* of x obtained with the basis $\mathcal{B}$. Of course, if the basis is changed, the coordinate system changes too, and so do the coordinates $\{a_i, i \in I\}$. Sometimes, to remark the dependence on $\mathcal{B}$, we write $(a_i)_{\mathcal{B}}$.

For a Hilbert space $\mathcal{H}$ with finite dimension n, a basis and the corresponding coordinate system establish a one-to-one correspondence between $\mathcal{H}$ and $\mathbb{C}^n$: the vectors x of $\mathcal{H}$ become the vectors of $\mathbb{C}^n$ composed by the Fourier coefficients of a_i, that is,

$$x \in \mathcal{H} \xrightarrow{\text{coordinates}} x_{\mathcal{B}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{C}^n. \tag{2.21}$$

Example 2.1 (Periodic discrete signals) Consider the vector space $\mathcal{L}_2 = \mathcal{L}_2(\mathbb{Z}(T)/\mathbb{Z}(NT))$ constituted by periodic discrete signals (with spacing T and period NT); $\mathbb{Z}(T) := \{nT|n \in \mathbb{Z}\}$ is the set of multiples of T. A basis for this space is formed by the signals

$$b_i = b_i(t) = \frac{1}{T}\delta_{\mathbb{Z}(T)/\mathbb{Z}(NT)}(t - iT), \qquad i = 0,\ 1,\ \ldots,\ N-1,$$

where $\delta_{\mathbb{Z}(T)/\mathbb{Z}(NT)}$ is the periodic discrete impulse [2]

$$\delta_{\mathbb{Z}(T)/\mathbb{Z}(NT)}(t) = \begin{cases} 1/T & t \in \mathbb{Z}(NT) \\ 0 & t \notin \mathbb{Z}(NT) \end{cases} \qquad t \in \mathbb{Z}(T).$$

This basis is orthonormal because

$$\langle b_i, b_j \rangle = \int_{\mathbb{Z}(T)/\mathbb{Z}(NT)} \mathrm{d}t\ b_i(t)\, b_j^*(t) = \delta_{ij}.$$

A first conclusion is that this vector space has finite dimension N.

For a generic signal $x = x(t)$, coefficients (2.20) provide

$$a_i = \langle x, b_i \rangle = \int_{\mathbb{Z}(T)/\mathbb{Z}(T_p)} \mathrm{d}t\ x(t)\, b_i^*(t) = \frac{1}{T}x(iT),$$

and therefore the signal coordinates are given by a vector collecting the values in one period, divided by T.

2.4.2 Dirac's Notation

In Quantum Mechanics, where systems are defined on a Hilbert space, vectors are indicated with a special notation, introduced by Dirac [3]. This notation, although apparently obscure, is actually very useful, and will be adopted from now on.

A vector x of a Hilbert space $\mathcal{H}$ is interpreted as a *column vector*, of possibly infinite dimension, and is indicated by the symbol

$$|x\rangle \tag{2.22a}$$

which is called *ket*. Its transpose conjugate $|x\rangle^*$ should be interpreted as a *row vector*, and is indicated by the symbol

$$\langle x| = |x\rangle^* \tag{2.22b}$$

which is called *bra*.[3] As a consequence, the inner product of two vectors $|x\rangle$ and $|y\rangle$ is indicated in the form

$$\langle x|y\rangle. \tag{2.22c}$$

We now exemplify this notation for the Hilbert space $\mathbb{C}^n$, comparing it to the standard notation

$$\begin{aligned} x &= \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{becomes} \quad |x\rangle = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ x^* &= \left[x_1^*, \ \ldots, \ x_n^*\right] \quad \text{becomes} \quad \langle x| = |x\rangle^* = \left[x_1^*, \ \ldots, \ x_n^*\right] \\ \langle x, y\rangle &= y^* x \quad \text{becomes} \quad \langle y|x\rangle = x_1\, y_1^* + \cdots + x_n\, y_n^*. \end{aligned} \tag{2.23}$$

Again, to become familiar with Dirac's notation, we also rewrite some relations, previously formulated with the conventional notation. A linear combination of vectors is written in the form

$$|x\rangle = a_1|x_1\rangle + a_2|x_2\rangle + \cdots + a_n|x_n\rangle.$$

The norm of a vector is written as $||x|| = \sqrt{\langle x|x\rangle}$. The orthogonality condition between two vectors $|x\rangle$ and $|y\rangle$ is now written as

$$\langle x|y\rangle = 0,$$

and the orthonormality of a basis $\mathcal{B} = \{|b_i\rangle, i \in I\}$ is written in the form

$$\langle b_i|b_j\rangle = \delta_{ij}.$$

The Fourier expansion with a finite-dimensional orthonormal basis $\mathcal{B} = \{|b_i\rangle | i = 1, \ldots, n\}$ becomes

$$|x\rangle = a_1|b_1\rangle + \cdots + a_n|b_n\rangle \tag{2.24}$$

where

$$a_i = \langle b_i|x\rangle, \tag{2.24a}$$

and can also be written in the form

$$|x\rangle = (\langle b_1|x\rangle)\ |b_1\rangle + \cdots + (\langle b_n|x\rangle)\ b_n\rangle. \tag{2.25}$$

[3] These names are obtained by splitting up the word "bracket"; in the specific case, the brackets are $\langle\ \rangle$.

Schwartz's inequality (2.12) becomes

$$|\langle x|y\rangle|^2 \leq \langle x|x\rangle\langle y|y\rangle \quad \text{or} \quad \langle x|y\rangle\langle y|x\rangle \leq \langle x|x\rangle\langle y|y\rangle. \tag{2.26}$$

Problem 2.1 ⋆ A basis in $\mathcal{H} = \mathbb{C}^2$ is usually denoted by $\{|0\rangle, |1\rangle\}$. Write the standard basis and a nonorthogonal basis.

Problem 2.2 ⋆⋆ An important basis in $\mathcal{H} = \mathbb{C}^n$ is given by the columns of the Discrete Fourier Transform (DFT) matrix of order n, given by

$$|w_i\rangle = \frac{1}{\sqrt{n}}\left[1, W_n^{-i}, W_n^{-2i}, \ldots, W_n^{-i(n-1)}\right]^{\mathrm{T}}, \qquad i = 0, 1, \ldots, n-1 \tag{E1}$$

where $W_n := \exp(\mathrm{i}2\pi/n)$ is the nth root of 1. Prove that this basis is orthonormal.

Problem 2.3 ⋆ Find the Fourier coefficients of ket

$$|x\rangle = \begin{bmatrix} 1 \\ \mathrm{i} \\ 2 \end{bmatrix} \in \mathbb{C}^3$$

with respect to the orthonormal basis (E1).

Problem 2.4 ⋆ Write the Fourier expansion (2.24) and (2.25) with a general orthonormal basis $\mathcal{B} = \{|b_i\rangle | i \in I\}$.

2.5 Linear Operators

2.5.1 Definition

An *operator* A from the Hilbert space $\mathcal{H}$ to the same space $\mathcal{H}$ is defined as a function

$$A : \mathcal{H} \to \mathcal{H}. \tag{2.27}$$

If $|x\rangle \in \mathcal{H}$, the operator A returns the vector

$$|y\rangle = A|x\rangle \quad \text{with} \quad |y\rangle \in \mathcal{H}. \tag{2.28}$$

To represent graphically the operator A, we can introduce a block (Fig. 2.3) containing the symbol of the operator, and in (2.28) $|x\rangle$ is interpreted as *input* and $|y\rangle$ as *output*.

The operator $A : \mathcal{H} \to \mathcal{H}$ is *linear* if the superposition principle holds, that is, if

$$A(a_1|x_1\rangle + a_2|x_2\rangle) = a_1 A|x_1\rangle + a_2 A|x_2\rangle$$

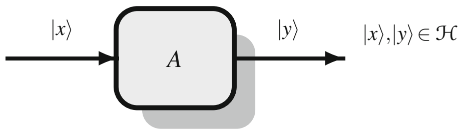

Fig. 2.3 Graphical representation of a linear operator

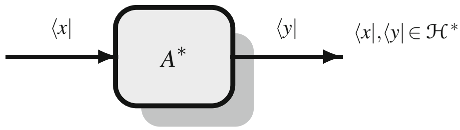

Fig. 2.4 The linear operator for "bras"; A^* is the *adjoint* of A

for every $|x_1\rangle, |x_2\rangle \in \mathcal{H}$ and $a_1, a_2 \in \mathbb{C}$.

A trivial linear operator is the *identity operator* $I_{\mathcal{H}}$ on $\mathcal{H}$ defined by the relation $I_{\mathcal{H}}\ |x\rangle \equiv |x\rangle$, for any vector $|x\rangle \in \mathcal{H}$. Another trivial linear operator is the *zero operator*, $0_{\mathcal{H}}$, which maps any vector onto the zero vector, $0_{\mathcal{H}}|x\rangle \equiv 0$.

In the interpretation of Fig. 2.3 the operator A acts on the kets (column vectors) of $\mathcal{H}$: assuming as input the ket $|x\rangle$, the operator outputs the ket $|y\rangle = A|x\rangle$. It is possible to associate to $\mathcal{H}$ a Hilbert space $\mathcal{H}^*$ (*dual* space) creating a correspondence between each ket $|x\rangle \in \mathcal{H}$ and its bra $\langle x|$ in $\mathcal{H}^*$. In this way, to each linear operator A of $\mathcal{H}$ a corresponding A^* of $\mathcal{H}^*$ can be associated, and the relation (2.28) becomes (Fig. 2.4)

$$\langle y| = \langle x|A^*.$$

The operator A^* is called the *adjoint*[4] of A. In particular, if $A = [a_{ij}]_{i,j=1,\ldots,n}$ is a square matrix, it results that $A^* = [a^*_{ji}]_{i,j=1,\ldots,n}$ is the *conjugate transpose*.

2.5.2 Composition of Operators and Commutability

The composition (*product*)[5] AB of two linear operators A and B is defined as the linear operator that, applied to a generic ket $|x\rangle$, gives the same result as would be obtained from the successive application of B followed by A, that is,

$$\{AB\}|x\rangle = A\{B|x\rangle\}. \tag{2.29}$$

In the graphical representation the product must be seen as a cascade of blocks (Fig. 2.5).

In general, like for matrices, the commutative property $AB = BA$ does not hold. Instead, to account for noncommutativity, the **commutator** between two operators

---

[4] This is not the ordinary definition of adjoint operator, but it is an equivalent definition, deriving from the relation $|x\rangle^* = \langle x|$ (see Sect. 2.8).

[5] We take for granted the definition of *sum* $A + B$ of two operators, and of *multiplication of an operator by a scalar* kA.

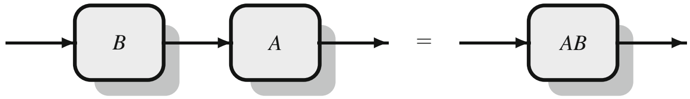

Fig. 2.5 Cascade connection of two operators

A and B is introduced, defined by

$$[A,\ B] := AB - BA. \tag{2.30}$$

In particular, if two operators commute, that is, if $AB = BA$, the commutator results in $[A,\ B] = 0$. Also an **anticommutator** is defined as

$$\{A,\ B\} := AB + BA. \tag{2.31}$$

Clearly, it is possible to express the product between two operators A and B in terms of the commutator and anticommutator

$$AB = \frac{1}{2}[A,\ B] + \frac{1}{2}\{A,\ B\}. \tag{2.31a}$$

In Quantum Mechanics, the commutator and the anticommutator are extensively used, e.g., to establish Heisenberg's uncertainty principle (see Sect. 3.9). Since most operator pairs do not commute, specific *commutation relations* are introduced through the commutator (see Chap. 11).

2.5.3 Matrix Representation of an Operator

As we have seen, a linear operator has properties very similar to those of a square matrix and, more precisely, to the ones that are obtained with a linear transformation of the kind $y = Ax$, where x and y are column vectors, and A is a square matrix, and it can be stated that linear operators are a generalization of square matrices. Also, it is possible to associate to each linear operator A a square matrix $A_{\mathcal{B}}$ of appropriate dimensions, $n \times n$, if the Hilbert space has dimension n, or of infinite dimensions if $\mathcal{H}$ has infinite dimension.

To associate a matrix to an operator A, we must fix an orthonormal basis $\mathcal{B} = \{|b_i\rangle,\ i \in I\}$ of $\mathcal{H}$. The relation

$$a_{ij} = \langle b_i|A|b_j\rangle, \qquad |b_i\rangle, |b_j\rangle \in \mathcal{B} \tag{2.32}$$

allows us to define the elements a_{ij} of a complex matrix $A_{\mathcal{B}} = [a_{ij}]$. In (2.32), the expression $\langle b_i|A|b_j\rangle$ must be intended as $\langle b_i|\{A|b_j\rangle\}$, that is, as the inner product of the bra $\langle b_i|$ and the ket $|A|b_j\rangle$ that is obtained by applying the operator A to the ket $|b_j\rangle$. Clearly, the matrix $A_{\mathcal{B}} = [a_{ij}]$, obtained from (2.32), depends on the basis chosen, and sometimes the elements of the matrix are indicated in the form $a_{ij\,\mathcal{B}}$ to stress such dependence on $\mathcal{B}$. Because all the bases of $\mathcal{H}$ have the same cardinality, all the matrices that can be associated to an operator have the same dimension.

From the matrix representation $A_{\mathcal{B}} = [a_{ij\,\mathcal{B}}]$ we can obtain the operator A using the *outer product* $|b_i\rangle\langle b_j|$, which will be introduced later on. The relation is

$$A = \sum_i \sum_j a_{ij\,\mathcal{B}} |b_i\rangle\langle b_j| \tag{2.33}$$

and will be proved in Sect. 2.7.

The matrix representation of an operator turns out to be useful as long as it allows us to interpret relations between operators as relations between matrices, with which we are usually more familiar. It is interesting to remark that an appropriate choice of a basis for $\mathcal{H}$ can lead to an "equivalent" matrix representation, simpler with respect to a generic choice of the basis. For example, we will see that a Hermitian operator admits a diagonal matrix representation with respect to a basis given by the eigenvectors of the operator itself.

In practice, as previously mentioned, in the calculations we will always refer to the Hilbert space $\mathcal{H} = \mathbb{C}^n$, where the operators can be interpreted as $n \times n$ square matrices with complex elements (the dimension of $\mathcal{H}$ could be infinite), keeping in mind anyhow that matrix representations with different bases correspond to the usual basis changes in normed vector spaces.

2.5.4 Trace of an Operator

An important parameter of an operator A is its *trace*, given by the sum of the diagonal elements of its matrix representation, namely

$$\boxed{\mathrm{Tr}[A] = \sum_i \langle b_i|A|b_i\rangle.} \tag{2.34}$$

The operation $\mathrm{Tr}[\cdot]$ appears in the formulation of the third postulate of Quantum Mechanics, and it is widely used in quantum decision.

The trace of an operator has the following properties, which will be often used in the following:

(1) The trace of A is independent of the basis with respect to which it is calculated, and therefore it is a characteristic parameter of the operator.
(2) The trace has the *cyclic* property

$$\text{Tr}[AB] = \text{Tr}[BA] \tag{2.35}$$

which holds even if the operators A and B are not commutable; such property holds also for rectangular matrices, providing that the products AB and BA make sense.

(3) The trace is *linear*, that is,

$$\text{Tr}[aA + bB] = a\,\text{Tr}[A] + b\,\text{Tr}[B], \qquad a, b \in \mathbb{C}. \tag{2.36}$$

For completeness we recall the important identity

$$\boxed{\langle u|A|u\rangle = \text{Tr}[A|u\rangle\langle u|]} \tag{2.37}$$

where $|u\rangle$ is an arbitrary vector, and $|u\rangle\langle u|$ is the operator given by the outer product, which will be introduced later.

2.5.5 Image and Rank of an Operator

The *image* of an operator A of $\mathcal{H}$ is the set

$$\text{im}(A) := A\mathcal{H} = \{A|x\rangle \mid |x\rangle \in \mathcal{H}\}. \tag{2.38}$$

It can be easily proved that $\text{im}(A)$ is *a subspace of* $\mathcal{H}$ (see Problem 2.6).

The dimension of this subspace defines the *rank* of the operator

$$\text{rank}(A) = \dim \text{im}(A) = |A\,\mathcal{H}|. \tag{2.39}$$

This definition can be seen as the extension to operators of the concept of rank of a matrix. As it appears from (2.38), to indicate the image of an operator of $\mathcal{H}$ we use the compact symbol $A\mathcal{H}$.

Problem 2.5 ⋆ Prove that the image of an operator on $\mathcal{H}$ is a subspace of $\mathcal{H}$.

Problem 2.6 ⋆ Define the 2D operator that inverts the entries of a ket and write its matrix representation with respect to the standard basis.

Problem 2.7 ⋆⋆ Find the matrix representation of the operator of the previous problem with respect to the DFT basis.

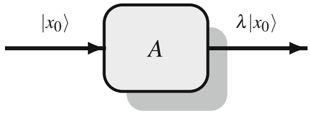

Fig. 2.6 Interpretation of eigenvalue and eigenvector of a linear operator A

2.6 Eigenvalues and Eigenvectors

An *eigenvalue* λ of a given operator A is a complex number such that a vector $|x_0\rangle \in \mathcal{H}$ exists, different from zero, satisfying the following equation:

$$\boxed{A|x_0\rangle = \lambda|x_0\rangle \qquad |x_0\rangle \neq 0\ .} \tag{2.40}$$

The vector $|x_0\rangle$ is called *eigenvector corresponding to the eigenvalue* λ.[6] The interpretation of relation (2.40) is illustrated in Fig. 2.6.

The set of all the eigenvalues is called *spectrum of the operator* and it will be indicated by the symbol $\sigma(A)$.

From the definition it results that the eigenvector $|x_0\rangle$ associated to a given eigenvalue is not unique, and in fact from (2.40) it results that also $2|x_0\rangle$, or $\mathrm{i}|x_0\rangle$ with i the imaginary unit, are eigenvectors of λ. The set of all the eigenvectors associated to the same eigenvalue

$$\mathcal{E}_\lambda = \{|x_0\rangle \;\big|\; A|x_0\rangle = \lambda|x_0\rangle\} \tag{2.41}$$

is always a subspace,[7] which is called *eigenspace* associated to the eigenvalue λ.

2.6.1 Computing the Eigenvalues

In the space $\mathcal{H} = \mathbb{C}^n$, where the operator A can be interpreted as an $n \times n$ matrix, the eigenvalue computation becomes the procedure usually followed with complex square matrices, consisting in the evaluation of the solutions to the *characteristic equation*

$$c(\lambda) = \det[A - \lambda\, I_{\mathcal{H}}] = 0$$

where $c(\lambda)$ is a polynomial. Then, for the fundamental theorem of Algebra, the number $r \leq n$ of *distinct* solutions is found: $\lambda_1, \lambda_2, \ldots, \lambda_r$, forming the spectrum of A

$$\sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_r\}.$$

[6] The eigenvector corresponding to the eigenvalue λ is often indicated by the symbol $|\lambda\rangle$.

[7] As we assume $|x_0\rangle \neq 0$, to complete $\mathcal{E}_\lambda$ as a subspace, the vector 0 of $\mathcal{H}$ must be added.

The solutions allow us to write $c(\lambda)$ in the form

$$c(\lambda) = a_0(\lambda - \lambda_1)^{p_1}(\lambda - \lambda_2)^{p_2} \dots (\lambda - \lambda_r)^{p_r}, \qquad a_0 \neq 0$$

where $p_i \geq 1$, $p_1+p_2+\cdots+p_r = n$, and p_i is called the *multiplicity* of the eigenvalue λ_i. Then we can state that the characteristic equation has always n solutions, counting the multiplicities.

As it is fundamental to distinguish whether we refer to distinct or to multiple solutions, we will use different notations in the two cases

$$\begin{array}{ll} \lambda_1, \lambda_2, \dots, \lambda_r & \text{for distinct eigenvalues} \\ \tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_n & \text{for eigenvalues with repetitions} \end{array} \tag{2.42}$$

Example 2.2 The 4×4 complex matrix

$$A = \frac{1}{4}\begin{bmatrix} 7 & -1+2\,\mathrm{i} & -1 & -1-2\,\mathrm{i} \\ -1-2\,\mathrm{i} & 7 & -1+2\,\mathrm{i} & -1 \\ -1 & -1-2\,\mathrm{i} & 7 & -1+2\,\mathrm{i} \\ -1+2\,\mathrm{i} & -1 & -1-2\,\mathrm{i} & 7 \end{bmatrix} \tag{2.43}$$

has the characteristic polynomial

$$c(\lambda) = 6 - 17\,\lambda + 17\,\lambda^2 - 7\,\lambda^3 + \lambda^4$$

which has solutions $\lambda_1 = 1$ with multiplicity 2, and $\lambda_2 = 2$ and $\lambda_3 = 3$ with multiplicity 1. Therefore, the distinct eigenvalues are $\lambda_1 = 1$, $\lambda_2 = 2$ and $\lambda_3 = 3$, whereas the eigenvalues with repetition are

$$\tilde{\lambda}_1 = 1, \qquad \tilde{\lambda}_2 = 1, \qquad \tilde{\lambda}_3 = 2, \qquad \tilde{\lambda}_4 = 3.$$

The corresponding eigenvectors are, for example,

$$|\tilde{\lambda}_1\rangle = \begin{bmatrix} 1+\mathrm{i} \\ \mathrm{i} \\ 0 \\ 1 \end{bmatrix} \quad |\tilde{\lambda}_2\rangle = \begin{bmatrix} -\mathrm{i} \\ 1-\mathrm{i} \\ 1 \\ 0 \end{bmatrix} \quad |\tilde{\lambda}_3\rangle = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \quad |\tilde{\lambda}_4\rangle = \begin{bmatrix} -\mathrm{i} \\ -1 \\ \mathrm{i} \\ 1 \end{bmatrix}. \tag{2.44}$$

As we will see with the spectral decomposition theorem, it is possible to associate different eigenvectors to coincident, or even orthogonal, eigenvalues. In (2.44) $|\tilde{\lambda}_1\rangle$ and $|\tilde{\lambda}_2\rangle$ are orthogonal, namely, $\langle\tilde{\lambda}_1|\tilde{\lambda}_2\rangle = 0$.

What was stated above for $\mathbb{C}^n$ can apply to any finite dimensional space n, using matrix representation. For an infinite dimensional space the spectrum can have infinite cardinality, but not necessarily. In any case it seems that no general procedures exist to compute the eigenvalues for the operators in an infinite dimensional space.

Trace of an operator from the eigenvalues It can be proved that the sum of the eigenvalues with coincidences gives the trace of the operator

$$\sum_{i=1}^{n} \tilde{\lambda}_i = \sum_{i=1}^{r} p_i \, \lambda_i = \mathrm{Tr}[A]. \tag{2.45}$$

It is also worthwhile to observe that the product of the eigenvalues $\tilde{\lambda}_i$ gives the determinant of the operator

$$\tilde{\lambda}_1 \, \tilde{\lambda}_2 \ldots \tilde{\lambda}_n = \lambda_1^{p_1} \, \lambda_2^{p_2} \ldots \lambda_r^{p_r} = \det(A) \tag{2.46}$$

and that the rank of A is given by the sum of the multiplicities of the $\tilde{\lambda}_i$ different from zero.

2.7 Outer Product. Elementary Operators

The *outer product* of two vectors $|x\rangle$ and $|y\rangle$ in Dirac's notation is indicated in the form

$$|x\rangle\langle y|,$$

which may appear similar to the inner product notation $\langle x|y\rangle$, but with factors inverted. This is not the case: while $\langle x|y\rangle$ is a complex number, $|x\rangle\langle y|$ is an operator. This can be quickly seen if $|x\rangle$ is interpreted as a column vector and $\langle y|$ as a row vector, referring for simplicity to the space $\mathbb{C}^n$, where

$$|x\rangle = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad \langle y| = \left[y_1^*, \ldots, y_n^* \right].$$

Then, using the matrix product, we have

$$|x\rangle\langle y| = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \left[y_1^*, \ldots, y_n^* \right] = \begin{bmatrix} x_1 y_1^* & \cdots & x_1 y_n^* \\ \vdots & & \vdots \\ x_n y_1^* & \cdots & x_n y_n^* \end{bmatrix}$$

that is, $|x\rangle\langle y|$ is an $n \times n$ square matrix.

The outer product[8] makes it possible to formulate an important class of linear operators, called *elementary operators* (or rank 1 operators) in the following way

[8] Above, the outer product was defined in the Hilbert space $\mathbb{C}^n$. For the definition in a generic Hilbert space one can use the subsequent (2.48), which defines $C = |c_1\rangle\langle c_2|$ as a linear operator.

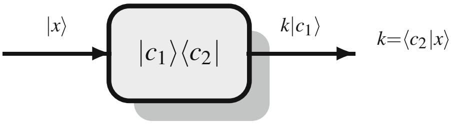

Fig. 2.7 Interpretation of the linear operator $C = |c_1\rangle\langle c_2|$

$$C = |c_1\rangle\langle c_2| \tag{2.47}$$

where $|c_1\rangle$ and $|c_2\rangle$ are two arbitrary vectors of the Hilbert space $\mathcal{H}$. To understand its meaning, let us apply to $C = |c_1\rangle\langle c_2|$ an arbitrary ket $|x\rangle \in \mathcal{H}$ (Fig. 2.7), which results in

$$C|x\rangle = (|c_1\rangle\langle c_2|)|x\rangle = (\langle c_2|x\rangle)|c_1\rangle, \qquad \forall\, |x\rangle \in \mathcal{H}, \tag{2.48}$$

namely, a vector proportional to $|c_1\rangle$, with a proportionality constant given by the complex number $k = \langle c_2|x\rangle$.

From this interpretation it is evident that the image of the elementary operator C is a straight line through the origin identified by the vector $|c_1\rangle$

$$\text{im}(|c_1\rangle\langle c_2|) = \{h\ |c_1\rangle | h \in \mathbb{C}\}$$

and obviously the elementary operator has unit rank.

Within the class of the elementary operators, a fundamental role is played, especially in Quantum Mechanics, by the operators obtained from the outer product of a ket $|b\rangle$ and the corresponding bra $\langle b|$, namely

$$B = |b\rangle\langle b|. \tag{2.49}$$

For these elementary operators, following the interpretation of Fig. 2.7, we realize that B transforms an arbitrary ket $|x\rangle$ into a ket proportional to $|b\rangle$. As we will see, if $|b\rangle$ is unitary, then $|b\rangle\langle b|$ turns out to be a *projector*.

2.7.1 Properties of an Orthonormal Basis

The elementary operators allow us to reinterpret in a very meaningful way the properties of an orthonormal basis in a Hilbert space $\mathcal{H}$. If $\mathcal{B} = \{|b_i\rangle, i \in I\}$ is an orthonormal basis on $\mathcal{H}$, then $\mathcal{B}$ identifies $k = |I|$ elementary operators $|b_i\rangle\langle b_i|$, and their sum gives the identity

$$\boxed{\sum_{i\in I} |b_i\rangle\langle b_i| = I_{\mathcal{H}} \qquad \text{for every orthonormal} \quad \mathcal{B} = \{|b_i\rangle, i \in I\}.} \tag{2.50}$$

In fact, if $|x\rangle$ is any vector of $\mathcal{H}$, its Fourier expansion (see (2.24)), using the basis $\mathcal{B}$ (see (2.24)), results in

$$\sum_i |b_i\rangle\langle b_i|x\rangle = \sum_i |b_i\rangle \sum_j a_i \,\langle b_i|b_j\rangle$$
$$= \sum_i |b_i\rangle\, a_i = |x\rangle$$

and, recalling that $\langle b_i|b_j\rangle = \delta_{ij}$, we have

$$\sum_i |b_i\rangle\langle b_i|x\rangle = \sum_i a_i|b_i\rangle = |x\rangle.$$

In other words, if we apply to the sum of the elementary operators $|b_i\rangle\langle b_i|$ the ket $|x\rangle$, we obtain again the ket $|x\rangle$ and therefore such sum gives the identity. The property (2.50), illustrated in Fig. 2.8, can be expressed by stating that the elementary operators $|b_i\rangle\langle b_i|$ obtained from an orthonormal basis $\mathcal{B} = \{|b_i\rangle, i \in I\}$ *give a resolution of the identity* $I_{\mathcal{H}}$ on $\mathcal{H}$.

The properties of an orthonormal basis $\mathcal{B} = \{|b_i\rangle, i \in I\}$ on the Hilbert space $\mathcal{H}$ can be so summarized:

(1) $\mathcal{B}$ is composed of *linearly independent* and orthonormal vectors

$$\langle b_i|b_j\rangle = \delta_{ij};$$

(2) the cardinality of $\mathcal{B}$ is, by definition, equal to the dimension of $\mathcal{H}$

$$|\mathcal{B}| = \dim \mathcal{H};$$

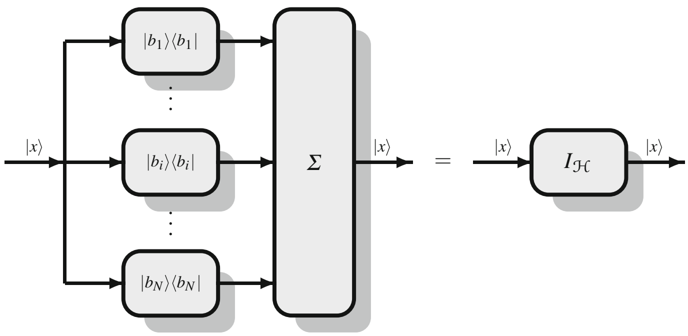

Fig. 2.8 The elementary operators $|b_i\rangle\langle b_i|$ obtained from an orthonormal basis $\mathcal{B} = \{|b_1\rangle, \ldots, |b_N\rangle\}$ provide a *resolution of the identity* $I_{\mathcal{H}}$

(3) the basis $\mathcal{B}$, through its elementary operators $|b_i\rangle\langle b_i|$, gives a *resolution of the identity* on $\mathcal{H}$, as stated by (2.50);
(4) $\mathcal{B}$ makes it possible to develop every vector $|x\rangle$ of $\mathcal{H}$ in the form (Fourier expansion)

$$|x\rangle = \sum_i a_i |b_i\rangle \quad \text{with} \quad a_i = \langle b_i|x\rangle. \tag{2.51}$$

Continuous bases Above we have implicitly assumed that the basis consists of an enumerable set of kets $\mathcal{B}$. In Quantum Mechanics also *continuous bases*, which consist of a continuum of eigenkets, are considered. This will be seen in the final chapters in the context of Quantum Information (see in particular Sect. 11.2).

2.7.2 Useful Identities Through Elementary Operators

Previously, we anticipated two identities requiring the notion of elementary operator.

A first identity, related to the trace, is given by (2.37), namely

$$\langle u|A|u\rangle = \text{Tr}[A|u\rangle\langle u|]$$

where A is an arbitrary operator, and $|u\rangle$ is a vector, also arbitrary. To prove this relation, let us consider an orthonormal basis $\mathcal{B} = \{|b_i\rangle, i \in I\}$ and let us apply the definition of a trace (2.34) to the operator $A|u\rangle\langle u|$. We obtain

$$\begin{aligned}\text{Tr}[A|u\rangle\langle u|] &= \sum_i \langle b_i|A|u\rangle\langle u|b_i\rangle \\ &= \sum_i \langle u|b_i\rangle\langle b_i|A|u\rangle = \langle u| \sum_i |b_i\rangle\langle b_i|A|u\rangle \\ &= \langle u|I_{\mathcal{H}}A|u\rangle = \langle u|A|u\rangle\end{aligned}$$

where we took into account the fact that $\sum_i |b_i\rangle\langle b_i|$ coincides with the identity operator $I_{\mathcal{H}}$ on $\mathcal{H}$ (see (2.50)).

A second identity is (2.33)

$$A = \sum_i \sum_j a_{ij} |b_i\rangle\langle b_j|$$

which makes it possible to reconstruct an operator A from its matrix representation $A_{\mathcal{B}} = [a_{ij}]$ obtained with the basis $\mathcal{B}$. To prove this relation, let us write A in the form $I_{\mathcal{H}} A I_{\mathcal{H}}$, and then let us express the identity $I_{\mathcal{H}}$ in the form (2.50). We obtain

$$A = I_{\mathcal{H}} A I_{\mathcal{H}} = \sum_i |b_i\rangle\langle b_i| A \sum_j |b_j\rangle\langle b_j|$$
$$= \sum_i \sum_j |b_i\rangle\langle b_i|A|b_j\rangle\langle b_j|$$

where (see (2.32)) $\langle b_i|A|b_j\rangle = a_{ij}$.

2.8 Hermitian and Unitary Operators

Basically, in Quantum Mechanics only *unitary* and *Hermitian* operators are used. Preliminary to the introduction of these two classes of operators is the concept of an *adjoint* operator.

The definition of adjoint is given in a very abstract form (see below). If we want to follow a more intuitive way, we can refer to the matrices associated to operators, recalling that if $A = [a_{ij}]$ is a complex square matrix, then:

- A^* indicates the conjugate transpose matrix, that is, the matrix with elements a^*_{ji},
- A is a *Hermitian* matrix, if $A^* = A$,
- A is a *normal* matrix, if $AA^* = A^*A$,
- A is a *unitary* matrix, if $AA^* = I$, where I is the identity matrix.

Note that the class of normal matrices includes as a special cases both Hermitian and unitary matrices (Fig. 2.9). It is also worthwhile to recall the conjugate transpose rule for the product of two square matrices

$$(AB)^* = B^* A^*. \tag{2.52}$$

2.8.1 The Adjoint Operator

The adjoint operator A^* was introduced in Sect. 2.5.1 as the operator for the bras $\langle x|$, $\langle y|$, whereas A is the operator for the kets $|x\rangle$, $|y\rangle$ (see Figs. 2.3 and 2.4). But the standard definition of adjoint is the following.

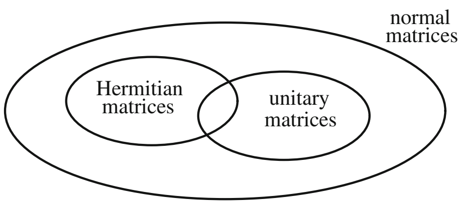

Fig. 2.9 The class of *normal matrices* includes *Hermitian matrices* and unitary matrices

Given an operator $A : \mathcal{H} \to \mathcal{H}$, the *adjoint* (*or Hermitian adjoint*) operator A^* is defined through the inner product from the relation[9]

$$(A|x\rangle, |y\rangle) = \left(|x\rangle, A^*|y\rangle\right), \qquad |x\rangle, |y\rangle \in \mathcal{H}. \tag{2.53}$$

It can be proved that the operator A^* verifying such relation exists and is unique and, also, if $A_{\mathcal{B}} = [a_{ij}]$ is the representative matrix of A, the corresponding matrix of A^* is the *conjugate transpose* of $A_{\mathcal{B}}$, namely the matrix $A^*_{\mathcal{B}} = [a^*_{ji}]$.

In addition, between two operators A and B and their adjoints the following relations hold:

$$\begin{aligned}
&(A^*)^* = A \\
&(A + B)^* = A^* + B^* \\
&(AB)^* = B^* A^* \\
&(aA)^* = a^* A^* \qquad a \in \mathbb{C}
\end{aligned} \tag{2.54}$$

that is, exactly the same relations that are obtained interpreting A and B as complex matrices.

2.8.2 Hermitian Operators

An operator $A : \mathcal{H} \to \mathcal{H}$ is called *Hermitian* (or self-adjoint) if it coincides with its adjoint, that is, if

$$A^* = A.$$

As a consequence, every representative matrix of A is a Hermitian matrix.

A fundamental property is that the spectrum of a Hermitian operator is composed of *real eigenvalues*. To verify this property, we start out by observing that for each vector $|x\rangle$ it results

$$\langle x|A|x\rangle \in \mathbb{R}, \qquad \forall |x\rangle \in \mathcal{H}, \qquad A \text{ Hermitian}. \tag{2.55}$$

In fact, the conjugate of such product gives $(\langle x|A|x\rangle)^* = \langle x|A^*|x\rangle = \langle x|A|x\rangle$. Now, if λ is an eigenvalue of A and $|x_0\rangle$ the corresponding eigenvector, it results that $\langle x_0|A|x_0\rangle = \langle x_0|\lambda x_0\rangle = \lambda\langle x_0|x_0\rangle$. Then $\lambda = \langle x_0|A|x_0\rangle / \langle x_0|x_0\rangle$ is real, being a ratio between real quantities.

Another important property of Hermitian operators is that the *eigenvectors corresponding to distinct eigenvalues are always orthogonal*. In fact, from $A|x_1\rangle = \lambda_1|x_1\rangle$ and $A|x_2\rangle = \lambda_2|x_2\rangle$, remembering that λ_1 and λ_2 are real, it follows that $\lambda_2\langle x_1|x_2\rangle = \langle x_1|A|x_2\rangle = \langle \lambda_1 x_1|x_2\rangle = \lambda_1\langle x_1|x_2\rangle$. Therefore, if $\lambda_1 \neq \lambda_2$, then nec-

[9] In most textbooks the adjoint operator is indicated by the symbol $A^\dagger$ and sometimes by A^+.

essarily $\langle x_1|x_2\rangle = 0$. This property can be expressed in terms of *eigenspaces* (see (2.41)) in the form: $\mathcal{E}_\lambda = \{|x_0\rangle \mid A|x_0\rangle = \lambda|x_0\rangle\}$, $\lambda \in \sigma(A)$, that is, the eigenspaces of a Hermitian operator are orthogonal and this is indicated as follows:

$$\mathcal{E}_\lambda \perp \mathcal{E}_\mu, \qquad \lambda \neq \mu.$$

Example 2.3 The matrix 4×4 defined by (2.43) is Hermitian in $\mathbb{C}^4$. As $\sigma(A) = \{1, 2, 3\}$, we have three eigenspaces $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{E}_3$. For the eigenvalues indicated in (2.44) we have

$$|\tilde{\lambda}_1\rangle, |\tilde{\lambda}_2\rangle \in \mathcal{E}_1, \qquad |\tilde{\lambda}_3\rangle \in \mathcal{E}_2, \qquad |\tilde{\lambda}_4\rangle \in \mathcal{E}_3.$$

It can be verified that orthogonality holds between eigenvectors belonging to different eigenspaces. For example,

$$\langle\lambda_1|\lambda_4\rangle = [1-\mathrm{i}, -\mathrm{i}, 0, 1]\begin{bmatrix}-\mathrm{i}\\-1\\\mathrm{i}\\1\end{bmatrix} = (1-\mathrm{i})(-\mathrm{i}) + (-\mathrm{i})(-1) + 1 = 0.$$

2.8.3 Unitary Operators

An operator $U : \mathcal{H} \to \mathcal{H}$ is called unitary if

$$U\,U^* = I_{\mathcal{H}} \tag{2.56}$$

where $I_{\mathcal{H}}$ is the identity operator. Both unitary and Hermitian operators fall into the more general class of *normal operators*, which are operators defined by the property $AA^* = A^*A$.

From definition (2.56) it follows immediately that U is *invertible*, that is, there exists an operator U^{-1} such that $UU^{-1} = I_{\mathcal{H}}$, given by

$$U^{-1} = U^*. \tag{2.57}$$

Moreover, it can be proved that the spectrum of U is always composed of eigenvalues λ_i *with unit modulus*.

We observe that, if $\mathcal{B} = \{|b_i\rangle, i \in I\}$ is an orthonormal basis of $\mathcal{H}$, all the other bases can be obtained through unitary operators, according to $\{U|b_i\rangle, i \in I\}$.

An important property is that the unitary operators preserve the inner product. In fact, if we apply the same unitary operator U to the vectors $|x\rangle$ and $|y\rangle$, so that $|u\rangle = U|x\rangle$ and $|v\rangle = U|y\rangle$, from $\langle u| = \langle x|U^*$, we obtain

$$\langle u|v\rangle = \langle x|U^*U|y\rangle = \langle x|y\rangle.$$

Example 2.4 A remarkable example of unitary operator in $\mathcal{H} = \mathcal{L}_2(I)$ is the operator/matrix F which gives the discrete Fourier transform (DFT)

$$F = \left(1/\sqrt{N}\right) \; [W_N^{-(r-s)}]_{r,s=0,1,\ldots,N-1} \tag{2.58}$$

where $W_N = \mathrm{e}^{\mathrm{i}2\pi/N}$. Then F is the DFT matrix. The inverse matrix is

$$F^{-1} = F^* = \left(1/\sqrt{N}\right) \; [W_N^{r-s}]_{r,s=0,1,\ldots,N-1}. \tag{2.59}$$

The columns of F, like for any unitary matrix, form an orthonormal basis of $\mathcal{H} = \mathbb{C}^N$.

Problem 2.8 ⋆ Classify the so-called *Pauli matrices*

$$\sigma_0 = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{E2}$$

which have an important role in quantum computation.

2.9 Projectors

Orthogonal projectors (briefly, projectors) are Hermitian operators of absolute importance for Quantum Mechanics, since quantum measurements are formulated with such operators.

2.9.1 Definition and Basic Properties

A projector $P : \mathcal{H} \to \mathcal{H}$ is an *idempotent* Hermitian operator, that is, with the properties

$$\boxed{P^* = P, \qquad P^2 = P} \tag{2.60}$$

and therefore $P^n = P$ for every $n \geq 1$.

Let $\mathcal{P}$ be the image of the projector P

$$\mathcal{P} = \mathrm{im}(P) = P\,\mathcal{H} = \{P|x\rangle \;\big|\; |x\rangle \in \mathcal{H}\}, \tag{2.61}$$

then, if $|s\rangle$ is a vector of $\mathcal{P}$, we get

$$\boxed{P|s\rangle = |s\rangle, \qquad |s\rangle \in \mathcal{P}.} \tag{2.62}$$

In fact, as a consequence of idempotency, if $|s\rangle = P|x\rangle$ we obtain $P|s\rangle = P(P|x\rangle) = P|x\rangle = |s\rangle$. Property (2.62) states that the subspace $\mathcal{P}$ is *invariant* with respect to the operator P.

Expression (2.62) establishes that each $|s\rangle \in \mathcal{P}$ is an eigenvector of P with eigenvalue $\lambda = 1$; the spectrum of P can contain also the eigenvalue $\lambda = 0$

$$\sigma(P) \subset \{0, 1\}. \tag{2.63}$$

In fact, the relation $P|x\rangle = \lambda|x\rangle$, multiplied by P gives

$$P^2|x\rangle = \lambda P|x\rangle = \lambda^2|x\rangle \quad \rightarrow \quad P|x\rangle = \lambda^2|x\rangle = \lambda|x\rangle.$$

Therefore, every eigenvalue satisfies the condition $\lambda^2 = \lambda$, which leads to (2.63).

Finally, (2.63) allows us to state that projectors are *nonnegative* or *positive semi-definite* operators (see Sect. 2.12.1). This property is briefly written as $P \geq 0$.

2.9.2 Why Orthogonal Projectors?

To understand this concept we must introduce the *complementary* projector

$$P_c = I - P \tag{2.64}$$

where $I = I_{\mathcal{H}}$ is the identity on $\mathcal{H}$. P_c is in fact a projector because it is Hermitian, and also $P_c^2 = I^2 + P^2 - IP - PI = I - P = P_c$. Now, in addition to the subspace $\mathcal{P} = P\mathcal{H}$, let us consider the *complementary* subspace $\mathcal{P}_c = P_c\mathcal{H}$. It can be verified that (see Problem 2.9):

(1) all the vectors of $\mathcal{P}_c$ are orthogonal to the vectors of $\mathcal{P}$, that is,

$$\langle s^\perp|s\rangle = 0 \qquad |s\rangle \in \mathcal{P},\ |s^\perp\rangle \in \mathcal{P}_c \tag{2.65}$$

and then we write $\mathcal{P}_c = \mathcal{P}^\perp$.

(2) the following relations hold:

$$P|s^\perp\rangle = 0,\ |s^\perp\rangle \in \mathcal{P}_c \qquad P_c|s\rangle = 0,\ |s\rangle \in \mathcal{P}. \tag{2.66}$$

(3) the decomposition of an arbitrary vector $|x\rangle$ of $\mathcal{H}$

$$|x\rangle = |s\rangle + |s^\perp\rangle \tag{2.67}$$

is uniquely determined by

$$|s\rangle = P|x\rangle, \qquad |s^\perp\rangle = P_c|x\rangle. \tag{2.67a}$$

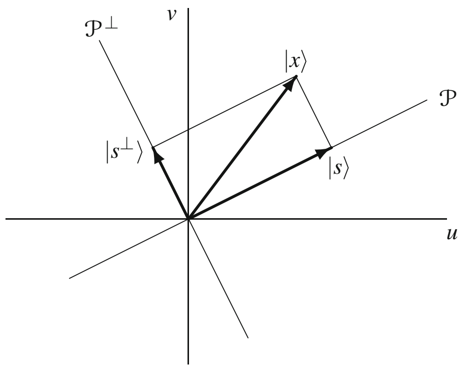

Fig. 2.10 Projection of the vector $|x\rangle$ *on* $\mathcal{P}$ along $\mathcal{P}^\perp$

Property (3) establishes that *the space* $\mathcal{H}$ *is given by the direct sum of the subspaces* $\mathcal{P}$ *and* $\mathcal{P}^\perp$, and we write $\mathcal{H} = \mathcal{P} \oplus \mathcal{P}^\perp$. According to properties (1) and (2), the projector P "projects the space $\mathcal{H}$ on $\mathcal{P}$ along $\mathcal{P}^\perp$".

Example 2.5 Consider the space $\mathcal{H} = \mathbb{R}^2$, which is "slightly narrow" for a Hilbert space, but sufficient to graphically illustrate the above properties. In $\mathbb{R}^2$, let us introduce a system of Cartesian axes u, v (Fig. 2.10), and let us indicate by $|x\rangle = \begin{bmatrix} u \\ v \end{bmatrix}$ the generic point of $\mathbb{R}^2$.

Let P_h be the real matrix

$$P_h = \frac{1}{1+h^2} \begin{bmatrix} 1 & h \\ h & h^2 \end{bmatrix} \tag{2.68}$$

where h is a real parameter. It can be verified that $P_h^2 = P_h$, therefore P_h is a projector. The space generated by P_h is

$$\mathcal{P} = \left\{ P_h \begin{bmatrix} u \\ v \end{bmatrix} | (u, v) \in \mathbb{R}^2 \right\} = \left\{ \begin{bmatrix} u \\ hu \end{bmatrix} | u \in \mathbb{R} \right\}$$

This is a straight line passing through the origin, whose slope is determined by h. We can see that the complementary projector

$$P_h^{(c)} = I - P_h$$

has the same structure as (2.68) with the substitution $h \to -1/h$, and therefore $\mathcal{P}^\perp$ is given by the line through the origin orthogonal to the one above. The conclusion is that the projector P_h projects the space $\mathbb{R}^2$ onto the line $\mathcal{P}$ along the line $\mathcal{P}^\perp$.

It remains to verify that the geometric orthogonality here implicitly invoked, coincides with the orthogonality defined by the inner product. If we denote by $|s\rangle$ the generic point of $\mathcal{P}$, and by $|s^\perp\rangle$ the generic point of $\mathcal{P}^\perp$, we get

$$|s\rangle = \begin{bmatrix} u \\ hu \end{bmatrix} \qquad |s^\perp\rangle = \begin{bmatrix} u_1 \\ (-1/h)u_1 \end{bmatrix}$$

for given u and u_1. Then

$$\langle s^\perp|s\rangle = \begin{bmatrix} u_1, & (-1/h)u_1 \end{bmatrix} \begin{bmatrix} u \\ hu \end{bmatrix} = 0.$$

Finally, the decomposition $\mathbb{R}^2 = \mathcal{P} \oplus \mathcal{P}^\perp$ must be interpreted in the following way (see Fig. 2.10): each vector of $\mathbb{R}^2$ can be uniquely decomposed into a vector of $\mathcal{P}$ and a vector of $\mathcal{P}^\perp$.

Example 2.6 We now present an example less related to the geometric interpretation, a necessary effort if we want to comprehend the generality of Hilbert spaces.

Consider the Hilbert space $\mathcal{L}_2 = \mathcal{L}_2(I)$ of the signals defined in I (see Sect. 2.3, Examples from Signal Theory), and let $\mathcal{E}$ be the subspace constituted by the *even signals*, that is, those verifying the condition $s(t) = s(-t)$. Notice that $\mathcal{E}$ is a subspace because every linear combination of even signals gives an even signal. The orthogonality condition between two signals $x(t)$ and $y(t)$ is

$$\int_I \mathrm{d}t\; x(t)\, y^*(t) = 0.$$

We state that the orthogonal complement of $\mathcal{E}$ is given by the class $\mathcal{O}$ of the *odd* signals, those verifying the condition $s(-t) = -s(t)$. In fact, it can be easily verified that if $x(t)$ is even and $y(t)$ is odd, their inner product is null (this for sufficiency; for necessity, the proof is more complex). Then

$$\mathcal{E}^\perp = \mathcal{O}.$$

We now check that a signal $x(t)$ of $\mathcal{L}_2$, which in general is neither even nor odd, can be uniquely decomposed into an even component $x_p(t)$ and an odd component $x_d(t)$. We have in fact (see Unified Theory [2])

$$x(t) = x_p(t) + x_d(t)$$

where

$$x_p(t) = \frac{1}{2}[x(t) + x(-t)], \qquad x_d(t) = \frac{1}{2}[x(t) - x(-t)].$$

Then (Fig. 2.11)

$$\mathcal{L}_2 = \mathcal{E} \oplus \mathcal{O}.$$

Notice that $\mathcal{E} \cap \mathcal{O} = \{0\}$, where 0 denotes the identically null signal.

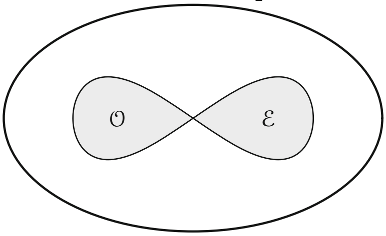

Fig. 2.11 The space $\mathcal{L}_2$ is obtained as a direct sum of the subspaces $\mathcal{E}$ and $\mathcal{O} = \mathcal{E}^\perp$

2.9.3 General Properties of Projectors

We summarize the general properties of a projector P:

(1) the spectrum is always $\sigma(P) = \{0, 1\}$,
(2) the multiplicity of eigenvalue 1 gives the rank of P and the dimension of the subspace $\mathcal{P} = P\mathcal{H}$,
(3) $P \geq 0$: is a positive semidefinite operator (see Sect. 2.12),
(4) $\mathrm{Tr}[P] = \mathrm{rank}(P)$: the trace of P gives the rank of the projector.

(1) has already been seen. (3) is a consequence of (1) and of Theorem 2.6. (4) follows from (2.45).

2.9.4 Sum of Projectors. System of Projectors

The sum of two projectors P_1 and P_2

$$P = P_1 + P_2$$

is not in general a projector. But if two projectors are *orthogonal* in the sense that

$$P_1 P_2 = 0,$$

the sum results again in a projector, as can be easily verified. An example of a pair of orthogonal projectors has already been seen above: P and the complementary projector P_c verify the orthogonality condition $PP_c = 0$, and their sum is

$$P + P_c = I,$$

where $I = I_{\mathcal{H}}$ is the identity on $\mathcal{H}$, which is itself a projector.

The concept can be extended to the sum of several projectors:

$$P = P_1 + P_2 + \cdots + P_k \tag{2.69}$$

where the addenda are pairwise orthogonal

$$P_i P_j = 0, \qquad i \neq j. \tag{2.69a}$$

To (2.69) one can associate $k+1$ subspaces

$$\mathcal{P} = P\mathcal{H} \quad \text{and} \quad \mathcal{P}_i = P_i\mathcal{H}, \ i = 1, \ldots, k$$

and, generalizing what was previously seen, we find that every vector $|s\rangle$ of $\mathcal{P}$ can be uniquely decomposed in the form

$$|s\rangle = |s_1\rangle + |s_2\rangle + \cdots + |s_k\rangle \quad \text{with} \quad |s_i\rangle = P_i|s\rangle. \tag{2.70}$$

Hence $\mathcal{P}$ is given by the *direct sum* of the subspaces $\mathcal{P}_i$

$$\mathcal{P} = \mathcal{P}_1 \oplus \mathcal{P}_2 \oplus \cdots \oplus \mathcal{P}_k.$$

If in (2.69) the sum of the projectors yields the identity I

$$P_1 + P_2 + \cdots + P_k = I \tag{2.71}$$

we say that the projectors $\{P_i\}$ *provide a resolution of the identity on* $\mathcal{H}$ and form a *complete* orthogonal class of projectors. In this case the direct sum gives the Hilbert space $\mathcal{H}$

$$\mathcal{P}_1 \oplus \mathcal{P}_2 \oplus \cdots \oplus \mathcal{P}_k = \mathcal{H}, \tag{2.72}$$

as illustrated in Fig. 2.12 for $k = 4$, where each point $|s\rangle$ of $\mathcal{P}$ is uniquely decomposed into the sum of 4 components $|s_i\rangle$ obtained from the relations $|s_i\rangle = P_i|s\rangle$, also shown in the figure.

For later use we find it convenient to introduce the following definition:

Definition 2.3 A set of operators $\{P_i, i \in I\}$ of the Hilbert space $\mathcal{H}$ constitutes a complete system of orthogonal projectors, briefly **projector system**, if they have the properties:

(1) the P_i are projectors (Hermitian and idempotent),
(2) the P_i are pairwise orthogonal ($P_i P_j = 0$) for $i \neq j$,
(3) the P_i form a resolution of the identity on $\mathcal{H}$ ($\sum_i P_i = \mathcal{I}_{\mathcal{H}}$).

The peculiarities of a projector system have been illustrated in Fig. 2.12.

Rank of projectors A projector has always a reduced rank with respect to the dimension of the space, unless P coincides with the identity I

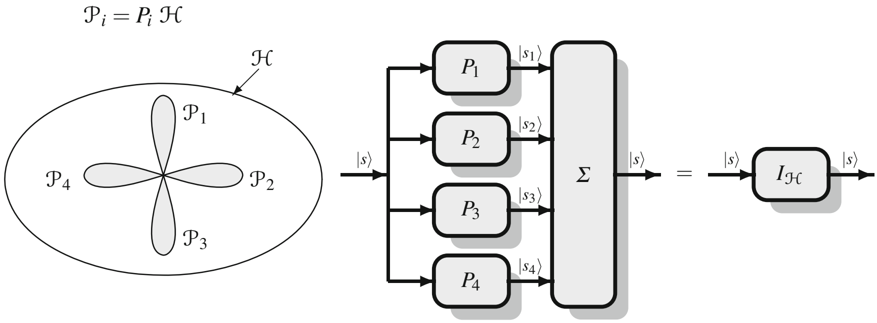

Fig. 2.12 The subspaces $\mathcal{P}_i$ give the space $\mathcal{H}$ as a *direct sum*. The projectors P_i give a *resolution of the identity* $I_{\mathcal{H}}$ and form a **projector system**

$$\mathrm{rank}(P) = \dim\mathcal{P} < \dim\mathcal{H} \qquad (P \neq I).$$

In the decomposition (2.71) it results as

$$\mathrm{rank}(P_1) + \cdots + \mathrm{rank}(P_k) = \mathrm{rank}(I) = \dim\mathcal{H},$$

and therefore in the corresponding direct sum we have that

$$\dim\mathcal{P}_1 + \cdots + \dim\mathcal{P}_k = \dim\mathcal{H}.$$

2.9.5 Elementary Projectors

If $|b\rangle$ is any unitary ket, the elementary operator

$$B = |b\rangle\langle b| \qquad \text{with} \quad ||\mathrm{b}|| = 1 \tag{2.73}$$

is Hermitian and verifies the condition $B^2 = |b\rangle\langle b|b\rangle\langle b| = B$, therefore it is a *unit rank projector*. Applying to B any vector $|x\rangle$ of $\mathcal{H}$ we obtain

$$B|x\rangle = |b\rangle\langle b|x\rangle = k\,|b\rangle \qquad \text{with}\, k = \langle b|x\rangle.$$

Hence B projects the space $\mathcal{H}$ on the straight line through the origin identified by the vector $|b\rangle$ (Fig. 2.13).

If two kets $|b_1\rangle$ and $|b_2\rangle$ are orthonormal, the corresponding elementary projectors $B_1 = |b_1\rangle\langle b_1|$ and $B_2 = |b_2\rangle\langle b_2|$ verify the orthogonality (for operators)

$$B_1 B_2 = |b_1\rangle\langle b_1|b_2\rangle\langle b_2| = 0$$

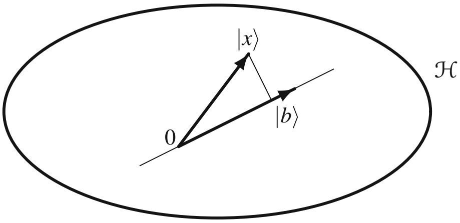

Fig. 2.13 The elementary projector $|b\rangle\langle b|$ *projects* an arbitrary ket $|x\rangle$ of $\mathcal{H}$ on the line of the ket $|b\rangle$

and therefore their sum

$$B_1 + B_2 = |b_1\rangle\langle b_1| + |b_2\rangle\langle b_2|$$

is still a projector (of rank 2).

Proceeding along this way we arrive at the following:

Theorem 2.1 *If $\mathcal{B} = \{|b_i\rangle, i \in I\}$ is an orthonormal basis of $\mathcal{H}$, the elementary projectors $B_i = |b_i\rangle\langle b_i|$ turn out to be orthogonal in pairs, and give the identity resolution*

$$\sum_{i \in I} B_i = \sum_{i \in I} |b_i\rangle\langle b_i| = I_{\mathcal{H}}.$$

In conclusion, through a generic basis of a Hilbert space $\mathcal{H}$ of dimension n it is always possible to "resolve" the space $\mathcal{H}$ through n elementary projectors, which form a projector system.

Problem 2.9 ⋆ Prove properties (2.65), (2.66) and (2.67) for a projector and its complement.

Problem 2.10 ⋆ Prove that projectors are positive semidefinite operators.

2.10 Spectral Decomposition Theorem (EID)

This theorem is perhaps the most important result of Linear Algebra because it sums up several previous results and opens the door to get so many interesting results. It will appear in various forms and will be referred to in different ways, for example, as *diagonalization* of a matrix and also as eigendecomposition or EID.

2.10.1 Statement and First Consequences

Theorem 2.2 *Let A be a Hermitian operator (or unitary) on the Hilbert space $\mathcal{H}$, and let $\{\lambda_i\}, i = 1, 2, \ldots, k$ be the distinct eigenvalues of A. Then A can be uniquely*

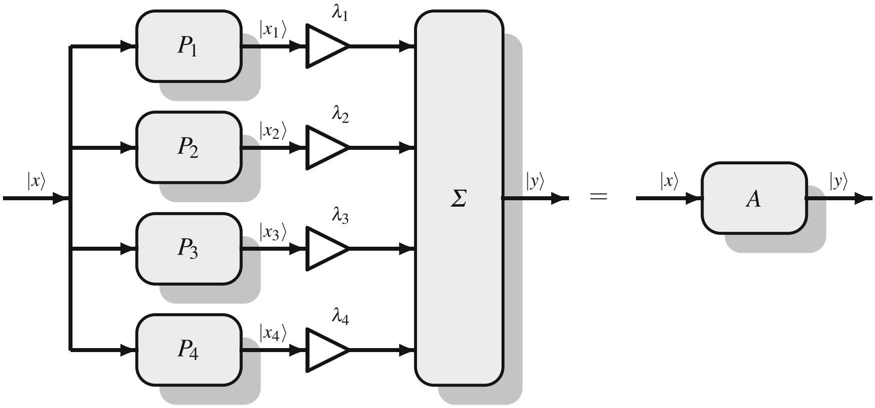

Fig. 2.14 Spectral decomposition of an operator A with four distinct eigenvalues

decomposed in the form

$$A = \sum_{i=1}^{k} \lambda_i \, P_i \tag{2.74}$$

where the $\{P_i\}$ *form a projector system.*

The spectral decomposition is illustrated in Fig. 2.14 for $k = 4$.

The theorem holds both for Hermitian operators, in which case the spectrum of the operator is formed by *real* eigenvalues λ_i, and for unitary operators, the eigenvalues λ_i have *unit modulus*.

We observe that, if p_i is the multiplicity of the eigenvalue λ_i, the rank of the projector P_i is just given by p_i

$$\text{rank}(P_i) = p_i \, .$$

In particular, if the eigenvalues have all unit multiplicity, that is, if A is *nondegenerate*, the projectors P_i have unit rank and assume the form

$$P_i = |\lambda_i\rangle\langle\lambda_i|$$

where $|\lambda_i\rangle$ is the eigenvector corresponding to eigenvalue λ_i. In this case the eigenvectors define an orthonormal basis for $\mathcal{H}$.

Example 2.7 The 4×4 complex matrix considered in Example 2.1

$$A = \frac{1}{4} \begin{bmatrix} 7 & -1+2\,\mathrm{i} & -1 & -1-2\,\mathrm{i} \\ -1-2\,\mathrm{i} & 7 & -1+2\,\mathrm{i} & -1 \\ -1 & -1-2\,\mathrm{i} & 7 & -1+2\,\mathrm{i} \\ -1+2\,\mathrm{i} & -1 & -1-2\,\mathrm{i} & 7 \end{bmatrix} \tag{2.75}$$

is Hermitian. It has been found that the distinct eigenvalues are

$$\lambda_1 = 1, \qquad \lambda_2 = 2, \qquad \lambda_3 = 3$$

with λ_1 of multiplicity 2. Then it is possible to decompose A through 3 projectors, in the form

$$A = \lambda_1 P_1 + \lambda_2 P_2 + \lambda_3 P_3$$

with P_1 of rank 2 and P_2, P_3 of rank 1. Such projectors result in

$$P_1 = \frac{1}{4}\begin{bmatrix} 2 & 1-\mathrm{i} & 0 & 1+\mathrm{i} \\ 1+\mathrm{i} & 2 & 1-\mathrm{i} & 0 \\ 0 & 1+\mathrm{i} & 2 & 1-\mathrm{i} \\ 1-\mathrm{i} & 0 & 1+\mathrm{i} & 2 \end{bmatrix} \qquad P_2 = \frac{1}{4}\begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$$

$$P_3 = \frac{1}{4}\begin{bmatrix} 1 & \mathrm{i} & -1 & -\mathrm{i} \\ -\mathrm{i} & 1 & \mathrm{i} & -1 \\ -1 & -\mathrm{i} & 1 & \mathrm{i} \\ \mathrm{i} & -1 & -\mathrm{i} & 1 \end{bmatrix}.$$

We leave it to the reader to verify the idempotency, the orthogonality, and the completeness of these projectors. In other words, to prove that the set $\{P_1, P_2, P_3\}$ forms a projector system.

2.10.2 Interpretation

The theorem can be interpreted in two ways:

- as *resolution* of a given Hermitian operator A, which makes it possible to identify a projector system $\{P_i\}$, as well as the corresponding eigenvalues $\{\lambda_i\}$,
- as *synthesis*, in which a projector system $\{P_i\}$ is known, and, for so many fixed distinct real numbers λ_i, a Hermitian operator can be built based on (2.74), having the λ_i as eigenvalues.

It is very interesting to see how the spectral decomposition acts on the input $|x\rangle$ and on the output $|y\rangle$ of the operator, following Fig. 2.14. The parallel of projectors decomposes in a unique way the input vector *into orthogonal components* (see (2.70)).

$$|x\rangle = |x_1\rangle + |x_2\rangle + \cdots + |x_k\rangle \tag{2.76}$$

where $|x_i\rangle = P_i|x\rangle$. In fact, as a consequence of the orthogonality of the projectors, we have

$$\langle x_i|x_j\rangle = \langle x|P_iP_j|x\rangle = 0, \qquad i \neq j$$

while (2.76) is a consequence of completeness. In addition, *each component* $|x_i\rangle$ *is an eigenvector of* A with eigenvalue λ_i. In fact, we have

$$A|x_i\rangle = \sum_j \lambda_j P_j |x_i\rangle = \lambda_i |x_i\rangle$$

where we have taken into account that $P_j|x_i\rangle = P_j P_i |x\rangle = 0$ if $i \neq j$.

Yet from Fig. 2.14, it appears that the decomposition of the input (2.76) is followed by the decomposition of the output in the form

$$|y\rangle = \lambda_1|x_1\rangle + \lambda_2|x_2\rangle + \cdots + \lambda_k|x_k\rangle$$

namely, as a sum of eigenvectors multiplied by the corresponding eigenvalues.

Going back to the decomposition of the input, as each component $|x_i\rangle = P_i|x\rangle$ belongs to the eigenspace $\mathcal{E}_{\lambda_i}$, and as $|x\rangle$ is an arbitrary ket of the Hilbert space $\mathcal{H}$, it results that

$$P_i\, \mathcal{H} = \mathcal{E}_{\lambda_i}$$

is the corresponding eigenspace. Furthermore, for completeness we have (see (2.72))

$$\mathcal{E}_{\lambda_1} \oplus \mathcal{E}_{\lambda_2} \oplus \cdots \oplus \mathcal{E}_{\lambda_k} = \mathcal{H}.$$

The reader can realize that the Spectral Decomposition Theorem allows us to refine what we saw in Sect. 2.9.4 on the sum of orthogonal projectors.

2.10.3 Decomposition via Elementary Projectors

In the statement of the theorem the eigenvalues $\{\lambda_i\}$ are assumed distinct, so the spectrum of the operator A is

$$\sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_k\} \quad \text{with} \quad k \leq n$$

where k may be smaller than the space dimension.

As we have seen, if $k = n$ all the eigenvalues have unit multiplicity and the decomposition (2.74) is done with elementary projectors. We can obtain a decomposition with elementary projectors even if $k < n$, that is, not all the eigenvalues have unit multiplicity. In this case we denote with

$$\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_n$$

the eigenvalues with repetition (see (2.42)). Then the spectral decomposition (2.74) takes the form

$$\boxed{A = \sum_{i=1}^{n} \tilde{\lambda}_i \; |b_i\rangle\langle b_i|} \tag{2.77}$$

where now the projectors $B_i = |b_i\rangle\langle b_i|$ are all elementary, and form a complete system of projectors.

To move from (2.74) to (2.77), let us consider an example regarding the space $\mathcal{H} = \mathbb{C}^4$, with A a 4×4 matrix. Suppose now that

$$\sigma(A) = \{\lambda_1, \lambda_2, \lambda_3\} \qquad p_1 = 2, \; p_2 = 1, \; p_3 = 1.$$

Then Theorem 2.2 provides the decomposition

$$A = \lambda_1 P_1 + \lambda_2 P_2 + \lambda_3 P_3$$

with

$$P_1 \quad \text{of rank } 2, \qquad P_2 = |\lambda_2\rangle\langle\lambda_2|, \qquad P_3 = |\lambda_3\rangle\langle\lambda_3|.$$

Consider now the subspace $\mathcal{P}_1 = P_1\mathcal{H}$ having dimension 2, which in turn is a Hilbert space, and therefore with basis composed of two orthonormal vectors, say $|c_1\rangle$ and $|c_2\rangle$. Choosing such a basis, the sum of the corresponding elementary projectors yields (see Theorem 2.1)

$$P_1 = |c_1\rangle\langle c_1| + |c_2\rangle\langle c_2|.$$

In this way we obtain (2.77) with $(\tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\lambda}_3, \tilde{\lambda}_4) = (\lambda_1, \lambda_1, \lambda_2, \lambda_3)$ and $|b_1\rangle = |c_1\rangle$, $|b_2\rangle = |c_2\rangle$, $|b_3\rangle = |\lambda_2\rangle$, $|b_4\rangle = |\lambda_3\rangle$. Notice that the $|c_i\rangle$ and the $|\lambda_i\rangle$ are independent (orthogonal) because they belong to different eigenspaces.

Example 2.8 In Example 2.1 we have seen that the eigenvalues with repetition of the matrix (2.75) are

$$\tilde{\lambda}_1 = 1, \qquad \tilde{\lambda}_2 = 1, \qquad \tilde{\lambda}_3 = 2, \qquad \tilde{\lambda}_4 = 3$$

and the corresponding eigenvectors are

$$|\tilde{\lambda}_1\rangle = \begin{bmatrix} 1+\mathrm{i} \\ \mathrm{i} \\ 0 \\ 1 \end{bmatrix} \quad |\tilde{\lambda}_2\rangle = \begin{bmatrix} -\mathrm{i} \\ 1-\mathrm{i} \\ 1 \\ 0 \end{bmatrix} \quad |\tilde{\lambda}_3\rangle = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \quad |\tilde{\lambda}_4\rangle = \begin{bmatrix} -\mathrm{i} \\ -1 \\ \mathrm{i} \\ 1 \end{bmatrix}. \tag{2.78}$$

These vectors are not normalized (they all have norm 2). The normalization results in

$$|b_i\rangle = \frac{1}{2}|\tilde{\lambda}_i\rangle \tag{2.79}$$

and makes it possible to build the elementary projectors $Q_i = |b_i\rangle\langle b_i|$. Therefore the spectral decomposition of matrix A via elementary projectors becomes

$$A = \tilde{\lambda}_1 Q_1 + \tilde{\lambda}_2 Q_2 + \tilde{\lambda}_3 Q_3 + \tilde{\lambda}_4 Q_4. \tag{2.80}$$

We encourage the reader to verify that with choice (2.79) the elementary operators Q_i are idempotent and form a system of orthogonal (elementary) projectors.

2.10.4 Synthesis of an Operator from a Basis

The Spectral Decomposition Theorem in the form (2.77), revised with elementary projectors, identifies an orthonormal basis $\{|b_i\rangle\}$.

The inverse procedure is also possible: given an orthonormal basis $\{|b_i\rangle\}$, a Hermitian (or unitary) operator can be built choosing an n-tuple of real eigenvalues $\tilde{\lambda}_i$ (or with unit modulus to have a unitary operator). In this way we obtain the synthesis of a Hermitian operator in the form (2.77).

Notice that with synthesis we can also obtain non-elementary projectors, thus arriving at the general form (2.74) established by the Spectral Decomposition Theorem. To this end, it suffices to choose some $\tilde{\lambda}_i$ equal. For example, if we want a rank 3 projector, we let

$$\tilde{\lambda}_1 = \tilde{\lambda}_2 = \tilde{\lambda}_3 = \lambda_1$$

and then

$$\tilde{\lambda}_1 |b_1\rangle\langle b_1| + \tilde{\lambda}_2 |b_2\rangle\langle b_2| + \tilde{\lambda}_3 |b_3\rangle\langle b_3| = \lambda_1 P_1$$

where $P_1 = |b_1\rangle\langle b_1| + |b_2\rangle\langle b_2| + |b_3\rangle\langle b_3|$ is actually a projector, as can be easily verified.

2.10.5 Operators as Generators of Orthonormal Bases

In Quantum Mechanics, orthonormal bases are usually obtained from the EID of operators, mainly Hermitian operators. Then, considering a Hermitian operator B, the starting point is the *eigenvalue relation*

$$B|b\rangle = b\,|b\rangle \tag{2.81}$$

where $|b\rangle$ denotes an eigenket of B and b the corresponding eigenvalue; B is given, while b and $|b\rangle$ are considered unknowns. The solutions to (2.81) provide the spectrum $\sigma(B)$ of B and also an orthonormal basis

$$\mathcal{B} = \{|b\rangle, b \in \sigma(B)\}$$

where $|b\rangle$ are supposed to be normalized, that is, $\langle b|b\rangle = 1$. Note the economic notation (due to Dirac [3]), where a single letter (b or B) is used to denote the operator, the eigenkets, and the eigenvalues.

The bases obtained from operators are used in several ways, in particular to represent kets and bras through the Fourier expansion and operators through the matrix representation. A systematic application of these concepts will be seen in Chap. 11 in the context of Quantum Information (see in particular Sect. 11.2).

2.11 The Eigendecomposition (EID) as Diagonalization

In the previous forms of spectral decomposition (EID) particular emphasis was given to projectors because those are the operators that are used in quantum measurements. Other forms are possible, or better, other interpretations of the EID, evidencing other aspects.

Relation (2.77) can be written in the form

$$\boxed{A = U\tilde{\Lambda}U^*} \tag{2.82}$$

where U is the $n \times n$ matrix having as columns the vectors $|b_i\rangle$, and $\tilde{\Lambda}$ is the diagonal matrix with diagonal elements $\tilde{\lambda}_i$, namely

$$U = [|b_1\rangle, |b_2\rangle, \ldots, |b_n\rangle], \qquad \tilde{\Lambda} = \text{diag}\,[\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_n]\,, \tag{2.82a}$$

and U^* is the conjugate transpose of U having as rows the bras $\langle b_i|$. As the kets $|b_i\rangle$ are orthonormal, the product UU^* gives identity

$$UU^* = I_{\mathcal{H}}. \tag{2.82b}$$

Thus U is a unitary matrix.

The decomposition (2.82) is a consequence of the Spectral Decomposition Theorem and is interpreted as *diagonalization* of the Hermitian matrix A. The result also holds for unitary matrices and more generally for normal matrices. Furthermore, from diagonalization one can obtain the spectral decomposition, therefore the two decompositions are equivalent.

Example 2.9 The diagonalization of the Hermitian matrix A, defined by (2.75), is obtained with

$$U = \frac{1}{2}\begin{bmatrix} 1+\mathrm{i} & -\mathrm{i} & -1 & -\mathrm{i} \\ \mathrm{i} & 1-\mathrm{i} & 1 & -1 \\ 0 & 1 & -1 & \mathrm{i} \\ 1 & 0 & 1 & 1 \end{bmatrix}, \qquad \tilde{\Lambda} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \tag{2.83}$$

2.11.1 On the Nonuniqueness of a Diagonalization

Is the diagonalization $A = U\Lambda U^*$ unique? A first remark is that the eigenvalues and also their multiplicity are unique. Next, consider two diagonalizations of the same matrix

$$A = U\Lambda U^*, \qquad A = U_1\Lambda U_1^*. \tag{2.84}$$

By a left multiplication by U^* and a right multiplication by U, we find

$$U^*U\Lambda U^*U = \Lambda = U^*U_1\Lambda U_1^*U.$$

Hence, a sufficient condition for the equivalence of the two diagonalizations is

$$U^* U_1 = I_{\mathcal{H}} \tag{2.85}$$

which reads: if the unitary matrices U and U_1 verify the condition (2.85), they both diagonalize the same matrix A.

But, we can also permute the order of the eigenvalues in the diagonal matrix Λ, combined with the same permutation of the eigenvectors in the unitary matrix, to get a new diagonalization. These, however, are only formal observations. The true answer to the question is given by [4]:

Theorem 2.3 *A matrix is uniquely diagonalizable, up to a permutation, if and only if its eigenvalues are all distinct.*

2.11.2 Reduced Form of the EID

So far, in the EID we have not considered the rank of matrix A. We observe that the rank of a linear operator is given by the number of nonzero eigenvalues (with multiplicity) $\tilde{\lambda}_i$. Then, if the $n \times n$ matrix A has the eigenvalue 0 with multiplicity p_0, the rank results in $r = n - p_0$. Sorting the eigenvalues with the null ones at the end, (2.77) and (2.82) become

$$\boxed{A = \sum_{i=1}^{r} \tilde{\lambda}_i \; |b_i\rangle\langle b_i| = U_r\,\tilde{\Lambda}_r\,U_r^*} \tag{2.86}$$

where

$$U_r = [|b_1\rangle, |b_2\rangle, \ldots, |b_r\rangle], \qquad \tilde{\Lambda}_r = \text{diag}\,[\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_r]. \tag{2.86a}$$

Therefore, U_r has dimensions $n \times r$ and collects as columns only the eigenvectors corresponding to nonzero eigenvalues, and the $r \times r$ diagonal matrix $\tilde{\Lambda}_r$ collects such eigenvalues. In the form (2.82) the eigenvectors with null eigenvalues are not relevant (because they are multiplied by 0), whereas in (2.86) all the eigenvectors are relevant. These two forms will be often used in quantum decision and, to distinguish them, the first will be called *full* form and the second *reduced* form.

2.11.3 Simultaneous Diagonalization and Commutativity

The diagonalization of a Hermitian operator given by (2.82), that is,

$$A = U \tilde{\Lambda} U^* \tag{2.87}$$

is done *with respect to the orthonormal basis* constituted by the columns of the unitary matrix U. The possibility that another operator B be diagonalizable with respect to the same basis, namely

$$B = U \tilde{\Lambda}_1 U^* \tag{2.88}$$

is bound to the commutativity of the two operators. In fact:

Theorem 2.4 *Two Hermitian operators A and B are commutative, $BA = AB$, if and only if they are simultaneously diagonalizable, that is, if and only if they have a common basis made by eigenvectors.*

The sufficiency of the theorem is immediately verified. In fact, if (2.87) and (2.88) hold simultaneously, we have

$$BA = U \tilde{\Lambda}_1 U^* U \tilde{\Lambda} U^* = U \tilde{\Lambda}_1 \tilde{\Lambda} U^*$$

where the diagonal matrices are always commutable, $\tilde{\Lambda}_1 \tilde{\Lambda} = \tilde{\Lambda} \tilde{\Lambda}_1$, thus $BA = AB$. Less immediate is the proof of necessity (see [5, p. 229]).

Commutativity and non-commutativity of Hermitian operators play a role in Heisenberg's Uncertainty Principle (see Sect. 3.9).

2.12 Functional Calculus

One of the most interesting applications of the Spectral Theorem is Functional Calculus, which allows for the introduction of arbitrary functions of an operator A, such as

$$A^m, \qquad e^A, \qquad \cos A, \qquad \sqrt{A}.$$

We begin by observing that from the idempotency and the orthogonality, in decomposition (2.74) we obtain

$$A^m = \sum_k \lambda_k^m P_k.$$

Thus a polynomial $p(\lambda)$, $\lambda \in \mathbb{C}$ over complex numbers is extended to the operators in the form

$$p(A) = \sum_k p(\lambda_k) P_k.$$

This idea can be extended to an arbitrary complex function $f : \mathbb{C} \to \mathbb{C}$ through

$$f(A) = \sum_k f(\lambda_k) P_k. \tag{2.89}$$

An alternative form of (2.89), based on the compact form (2.82), is given by

$$\boxed{f(A) = U f(\tilde{\Lambda}) U^*} \tag{2.89a}$$

where

$$f(\tilde{\Lambda}) = \text{diag}\,[f(\tilde{\lambda}_1), \ldots, f(\tilde{\lambda}_n)]. \tag{2.89b}$$

The following theorem links the Hermitian operators to the unitary operators [4]

Theorem 2.5 *An operator U is unitary if and only if it can be written in the form*

$$U = e^{iA} \qquad \text{with } A \text{ Hermitian operator},$$

where, from (2.89),

$$e^{iA} = \sum_k e^{i\lambda_k} P_k. \tag{2.90}$$

For a proof, see [4]. As a check, we observe that, if A is Hermitian, its eigenvalues λ_k are real. Then, according to (2.90), the eigenvalues of U are $e^{i\lambda_k}$, which, as it must be, have unit modulus.

Importance of the exponential of an operator In Quantum Mechanics a fundamental role is played by the exponential of an operator, in particular in the form e^{A+B}, where A and B do not commute. This topic will be developed in Sect. 11.6.

2.12.1 Positive Semidefinite Operators

We first observe that if A is a Hermitian operator, the quantity $\langle x|A|x\rangle$ is always a real number. Then a Hermitian operator is:

- *nonnegative* or *positive semidefinite* and is written as $A \geq 0$ if

$$\langle x|A|x\rangle \geq 0 \qquad \forall\ |x\rangle \in \mathcal{H} \tag{2.91}$$

- *positive definite* and is written as $A > 0$ if

$$\langle x|A|x\rangle > 0 \qquad \forall\ |x\rangle \neq 0. \tag{2.92}$$

From the Spectral Theorem it can be proved that [1, Theorem 10.23]:

Theorem 2.6 *In a finite dimensional space, a Hermitian operator A is positive semidefinite (positive) if and only if its eigenvalues are nonnegative (positive).*

Remembering that the spectrum of a projector P is $\sigma(P) = \{0, 1\}$, we find, as anticipated in Sect. 2.9.3:

Corollary 2.1 *The projectors are always positive semidefinite operators.*

2.12.2 Square Root of an Operator

The square root of a positive semidefinite Hermitian operator $A \geq 0$ is introduced starting from its spectral resolution

$$A = \lambda_1 P_1 + \cdots + \lambda_k P_k \qquad \lambda_j \geq 0$$

in the following way:

$$\sqrt{A} = \sqrt{\lambda_1}\, P_1 + \cdots + \sqrt{\lambda_k}\, P_k \tag{2.93a}$$

or in equivalent form (see (2.89a))

$$\sqrt{A} = U\, \sqrt{\tilde{\Lambda}}\, U^* \tag{2.93b}$$

where $\sqrt{\tilde{\Lambda}} = \text{diag}\,[\sqrt{\tilde{\lambda}_1}, \ldots, \sqrt{\tilde{\lambda}_n}]$. The definition of $\sqrt{A}$ is unique and it can be soon verified that from (2.93) it follows that

$$\left(\sqrt{A}\right)^2 = A.$$

The square root of a Hermitian operator will find interesting applications in Quantum Communications starting from Chap. 6.

2.12.3 Polar Decomposition

These decompositions regard arbitrary operators and therefore not necessarily Hermitian or unitary [4]

Theorem 2.7 (Polar decomposition) *Let A be an arbitrary operator. Then there always exists a unitary operator U and two positive definite Hermitian operators J and K such that*

$$A = UJ = KU$$

where J and K are unique with

$$J = \sqrt{A^*A} \quad \text{and} \quad K = \sqrt{AA^*}. \tag{2.94}$$

The theorem can be considered as an extension to square matrices of the polar decomposition of complex numbers: $z = |z| \exp(\mathrm{i} \arg z)$.

2.12.4 Singular Value Decomposition

So far we have considered operators of the Hilbert space, which in particular become complex square matrices. The singular value decomposition (SVD) considers more generally rectangular matrices.

Theorem 2.8 *Let A be an $m \times n$ complex matrix. Then the singular value decomposition of A results in*

$$A = UDV^*, \tag{2.95}$$

where

- *U is an $m \times m$ unitary matrix,*
- *V is an $n \times n$ unitary matrix,*
- *D is an $m \times n$ diagonal matrix with real nonnegative values on the diagonal.*

The positive values d_i of the diagonal matrix D are called the *singular values* of A. It can be proved that the SVD of a matrix A is strictly connected to the EIDs of the Hermitian matrices AA^* and A^*A (see [4] and Chap. 5).

If the matrix has rank r, the positive values d_i are r and a more explicit form can be given for the decomposition

$$A = U_r D_r V_r^* = \sum_{i=1}^{r} d_i \, |u_i\rangle\langle v_i| \tag{2.96}$$

where

- $U_r = [|u_1\rangle \cdots |u_r\rangle]$ is an $m \times r$ matrix,
- $V_r = [|v_1\rangle \cdots |v_r\rangle]$ is an $n \times r$ matrix,
- D is an $r \times r$ diagonal matrix collecting on the diagonal the singular values d_i.

The form (2.96) will be called the *reduced* form of the SVD. Both forms play a fundamental role in the theory of quantum decision.

Example 2.10 Consider the 4×2 matrix

$$A = \frac{1}{12\sqrt{2}} \begin{bmatrix} 5 & 1 \\ 3-2\mathrm{i} & 3+2\mathrm{i} \\ 1 & 5 \\ 3+2\mathrm{i} & 3-2\mathrm{i} \end{bmatrix}.$$

The SVD UDV^* of A results in

$$U = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{-\mathrm{i}}{2} & -\frac{1}{2} & \frac{\mathrm{i}}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\mathrm{i}}{2} & -\frac{1}{2} & \frac{-\mathrm{i}}{2} \end{bmatrix}, \quad V = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad D = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Therefore the singular values are $d_1 = 1/2$ and $d_2 = 1/3$. The reduced form $U_r\, D_r\, V_r^*$ becomes

$$U_r = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{-\mathrm{i}}{2} \\ \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\mathrm{i}}{2} \end{bmatrix}, \quad V_r = V = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad D_r = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{9} \end{bmatrix}.$$

We leave it to the reader to verify that, carrying out the products in these decompositions, one obtains the original matrix A.

2.12.5 Cholesky's Decomposition

Another interesting decomposition for Hermitian matrices is given by Cholesky's decomposition [4].

Theorem 2.9 (Cholesky's decomposition) *Let A be an $n \times n$ positive semidefinite Hermitian matrix, then there exists an upper triangular matrix U, of dimensions $n \times n$, with nonnegative elements on the main diagonal, such that*

$$A = U^* U.$$

If A is positive semidefinite then the matrix U is unique, and the elements of its main diagonal are all positive.

For a given matrix A it can turn out to be useful as an alternative to the EID for the factor decomposition of the density operators (see Chap. 5).

Example 2.11 Consider again the Hermitian matrix of Example 2.1

$$A = \frac{1}{4}\begin{bmatrix} 7 & -1+2\mathrm{i} & -1 & -1-2\mathrm{i} \\ -1-2\mathrm{i} & 7 & -1+2\mathrm{i} & -1 \\ -1 & -1-2\mathrm{i} & 7 & -1+2\mathrm{i} \\ -1+2\mathrm{i} & -1 & -1-2\mathrm{i} & 7 \end{bmatrix}.$$

Cholesky's decomposition, obtained with `Mathematica`, is specified by the triangular factor

$$U = \begin{bmatrix} \frac{\sqrt{7}}{2} & -\frac{\frac{1}{2}-\mathrm{i}}{\sqrt{7}} & -\frac{1}{2\sqrt{7}} & -\frac{\frac{1}{2}+\mathrm{i}}{\sqrt{7}} \\ 0 & \sqrt{\frac{11}{7}} & -\frac{2-3\mathrm{i}}{\sqrt{77}} & -\frac{1+\mathrm{i}}{\sqrt{77}} \\ 0 & 0 & \sqrt{\frac{17}{11}} & -\frac{3-4\mathrm{i}}{\sqrt{187}} \\ 0 & 0 & 0 & 2\sqrt{\frac{6}{17}} \end{bmatrix} = \begin{bmatrix} 1.32 & -0.19+\mathrm{i}\,0.38 & -0.19 & -0.19-\mathrm{i}\,0.38 \\ 0 & 1.25 & -0.23+\mathrm{i}\,0.34 & -0.11-\mathrm{i}\,0.11 \\ 0 & 0 & 1.24 & -0.22+\mathrm{i}\,0.29 \\ 0 & 0 & 0 & 1.19 \end{bmatrix}.$$

The decomposition is unique and has positive elements on the main diagonal, as predicted by Theorem 2.9.

Problem 2.11 ⋆⋆⋆ Let A be an arbitrary operator of the Hilbert space $\mathcal{H}$. Show that the operator AA^* is always positive semidefinite.

Hint: use diagonalization of A.

2.13 Tensor Product

The *tensor product* makes it possible to combine two or more vector spaces to obtain a larger vector space. In Quantum Mechanics it is used in Postulate 4 to combine quantum systems.

Before giving the definition, we introduce the symbolism that will be used. If $\mathcal{H}_1$ and $\mathcal{H}_2$ are two Hilbert spaces (on complex numbers), their tensor product is indicated in the form

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$$

and, as we will see, the new Hilbert space $\mathcal{H}$ has dimension

$$\dim(\mathcal{H}_1 \otimes \mathcal{H}_2) = \dim(\mathcal{H}_1)\dim(\mathcal{H}_2).$$

If $|x\rangle \in \mathcal{H}_1$ and $|y\rangle \in \mathcal{H}_2$ the kets and the bras of $\mathcal{H}$ are indicated, respectively, in the form

$$|x\rangle \otimes |y\rangle \qquad \langle x| \otimes \langle y| \;, \tag{2.97}$$

which is sometimes simplified as

$$|x\rangle|y\rangle \qquad \langle x|\langle y|. \tag{2.97a}$$

If A is an operator of $\mathcal{H}_1$ and B is an operator of $\mathcal{H}_2$, the operator of $\mathcal{H}$ is indicated in the form

$$A \otimes B.$$

We now list the abstract definitions of the vectors and of the operators that are obtained through the tensor product. However, as these definitions are very abstract, or better, scarcely operational, they can be skipped and the reader may move to the next section, where the tensor product is developed for matrices and is more easily understood.

2.13.1 Abstract Definition ⇓

We want to combine two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ using the tensor product $\otimes$, and let us denote by $|x\rangle$, $|x_1\rangle$, $|x_2\rangle$ arbitrary kets of $\mathcal{H}_1$, and by $|y\rangle$, $|y_1\rangle$, $|y_2\rangle$ arbitrary kets of $\mathcal{H}_2$. Then, by definition, the tensor product of kets must have the following properties:

(1) homogeneity

$$a(|x\rangle \otimes |y\rangle) = (a|x\rangle) \otimes |y\rangle = |x\rangle \otimes a(|y\rangle) \tag{2.98a}$$

(2) linearity with respect to the first factor

$$(|x_1\rangle + |x_2\rangle) \otimes |y\rangle = |x_1\rangle \otimes |y\rangle + |x_2\rangle \otimes |y\rangle \tag{2.98b}$$

(3) linearity with respect to the second factor

$$|x\rangle \otimes (|y_1\rangle + |y_2\rangle) = |x\rangle \otimes |y_1\rangle + |x\rangle \otimes |y_2\rangle. \tag{2.98c}$$

Once defined the tensor products between the kets, imposing the above conditions, the tensor product between bras can be obtained as follows:

$$\langle x| \otimes \langle y| = (|x\rangle)^* \otimes (|y\rangle)^*.$$

Then we can move on to define the tensor product of two operators in this way:

$$(A \otimes B)(|x\rangle \otimes |y\rangle) = (A|x\rangle) \otimes (B|y\rangle) \tag{2.99}$$

where on the right-hand side we find the tensor product of two kets, which has already been defined.

Finally, the inner product on $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ is defined by the inner products on $\mathcal{H}_1$ and $\mathcal{H}_2$, through the relation

$$(\langle x_1| \otimes \langle y_1|)(|x_2\rangle \otimes |y_2\rangle) = \langle x_1|x_2\rangle\ \langle y_1|y_2\rangle. \tag{2.100}$$

2.13.2 Kronecker's Product of Vector and Matrices

Consider two row vectors written in standard notation

$$x = [x_0, \ldots, x_{m-1}], \qquad y = [y_0, \ldots, y_{n-1}]$$

and suppose we want to form a "product" containing all possible products between the elements of two vectors

$$x_i\, y_j, \qquad i = 0, 1, \ldots, m-1,\ \ j = 0, 1, \ldots, n-1$$

The natural procedure would be to build the $m \times n$ matrix

$$[x_i\, y_j]$$

but with the Kronecker product we want to build *a vector* containing all the mn products. The problem is that of passing from a bidimensional configuration (2D), like the matrix $[x_i\, y_j]$, to a 1D configuration, as a vector is. But, while in 1D there is a natural order of the indexes, namely 0, 1, 2, ..., in 2D such order does not exist for the indexes (i, j). We must then introduce a conventional ordering to establish, for example, whether (1, 2) comes before or after (2, 1). A solution to the problem is given by the *lexicographical order*[10] obtained as follows: in the pair of indexes (i, j) we fix the first index starting from $i = 0$ and let run the second index j along

[10] This name comes from the order given to words in the dictionary: a word of k letters, $a = (a_1, \ldots, a_k)$ appears in the dictionary before the word $b = (b_1, \ldots, b_k)$, symbolized $a < b$, *if and only if the first a_i which is different from b_i comes before b_i in the alphabet*. In our context the

its range, obtaining $(0, 0), (0, 1), \ldots, (1, n-1)$, we then move to the value $i = 1$, until $i = m - 1$. In this way we associate the pair of integer indexes to a single index given by

$$h = j + (i-1)\,n, \qquad j = 0, 1, \ldots, n-1 \quad i = 0, 1, \ldots, m-1$$

which gives the required 1D ordering. The resulting vector can be written in the compact form:

$$x \otimes y = [x_1 y, x_2 y, \ldots, x_m y] \tag{2.101}$$

where the form $x_i y$ indicates the n-tuple $(x_i y_1, \ldots, x_i y_n)$. Relation (2.101) defines Kronecker's product of two vectors x and y.

Next we consider two matrices

$$\begin{aligned} A &= [a_{ir}], \qquad i = 1, \ldots, m, \;\; r = 1, \ldots, p \\ B &= [b_{is}], \qquad j = 1, \ldots, n, \;\; s = 1, \ldots, q \end{aligned}$$

where the dimensions are respectively $m \times p$ and $n \times q$. To collect all the products of the entries of the two matrices we would have to form a 4D matrix

$$[a_{ir}\, b_{js}]$$

but, if we want a standard 2D matrix, we apply the lexicographical order to the pairs of indexes (i, j) and (r, s), given by the integers

$$h = j + i(p-1), \qquad k = s + r(q-1)$$

where h goes from 1 to mp and k from 1 to nq. In this way we build an $A \otimes B$ matrix of dimension $mn \times pq$.

The compact form for $A \otimes B$ is

$$A \otimes B = \begin{bmatrix} a_{11}B & \ldots & a_{1p}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \ldots & a_{mp}B \end{bmatrix} \tag{2.102}$$

where on the left-hand side we notice the "blocks" $a_{ij}B$, which are $n \times q$ matrices. Once we expand these blocks, we can see that on the left-hand side the resulting matrix has dimensions $mn \times pq$.

(Footnote 10 continued)
alphabet is given by the set of integers. Then we find, e.g., that $(1, 3) < (2, 1)$, $(0, 3, 2) < (1, 0, 1)$ and $(1, 1, 3) < (1, 2, 0)$.

It can be verified that (2.102) falls into the abstract definition of tensor product, based on the previous "abstract" conditions (1), (2), and (3).

Relation (2.102) extends to matrices in the compact form (2.101) seen for vectors and represents the definition of the *Kronecker product* for matrices. It includes the case of vectors, provided that vectors are regarded as matrices.

Example 2.12 If

$$|a\rangle = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \qquad |b\rangle = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

the tensor product gives

$$|a\rangle \otimes |b\rangle = \begin{bmatrix} a_1 b \\ a_2 b \end{bmatrix} = \begin{bmatrix} a_1 b_1 \\ a_1 b_2 \\ a_1 b_3 \\ a_2 b_1 \\ a_2 b_2 \\ a_2 b_3 \end{bmatrix},$$

which is a 6×1 vector. In particular,

$$|a\rangle = \begin{bmatrix} 1+\mathrm{i} \\ 2+\mathrm{i} \end{bmatrix}, \qquad |b\rangle = \begin{bmatrix} 1+\mathrm{i} \\ 2+2\mathrm{i} \\ 3+2\mathrm{i} \end{bmatrix} \qquad \to \qquad |a\rangle \otimes |b\rangle = \begin{bmatrix} 2\mathrm{i} \\ 4\mathrm{i} \\ 5+\mathrm{i} \\ 1+3\mathrm{i} \\ 2+6\mathrm{i} \\ 4+7\mathrm{i} \end{bmatrix}.$$

If

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$$

are two matrices of dimensions, respectively, 2×2 and 3×2, the tensor product yields

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{bmatrix}$$

which is a 6×4 matrix. In particular, if

$$A = \begin{bmatrix} \mathrm{i} & 3 \\ 2+\mathrm{i} & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2\mathrm{i} & 1 \\ 2 & \mathrm{i} \\ \mathrm{i} & 3 \end{bmatrix} \to A \otimes B = \begin{bmatrix} -2 & \mathrm{i} & 6\mathrm{i} & 3 \\ 2\mathrm{i} & -1 & 6 & 3\mathrm{i} \\ -1 & 3\mathrm{i} & 3\mathrm{i} & 9 \\ -2+4\mathrm{i} & 2+\mathrm{i} & 2\mathrm{i} & 1 \\ 4+2\mathrm{i} & -1+2\mathrm{i} & 2 & \mathrm{i} \\ -1+2\mathrm{i} & 6+3\mathrm{i} & \mathrm{i} & 3 \end{bmatrix}.$$

For Kronecker's product (2.102) the following rules can be established. The transpose and the conjugate transpose simply result in

$$(A \otimes B)^{\mathrm{T}} = A^{\mathrm{T}} \otimes B^{\mathrm{T}}, \qquad (A \otimes B)^* = A^* \otimes B^* \tag{2.103}$$

and also the important rule holds (valid if the dimensions are compatible with ordinary products)

$$\boxed{(A \otimes B)(C \otimes D) = (AC) \otimes (BD)} \tag{2.104}$$

which contains both the Kronecker product and the ordinary matrix product and will be called **mixed-product law**. In addition, for two *invertible square* matrices we have

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}. \tag{2.105}$$

2.13.3 Properties of the Tensor Product

Kronecker's product of matrices allows us now to interpret and verify the definition and the properties of the tensor product on Hilbert spaces. This is done directly when $\mathcal{H}_1 = \mathbb{C}^m$ an $\mathcal{H}_2 = \mathbb{C}^n$ and it turns out that $\mathcal{H}_1 \otimes \mathcal{H}_2 = \mathbb{C}^{mn}$, but using the matrix representation it can be done for arbitrary Hilbert spaces of finite dimension (and with some effort even of infinite dimension).

Then, if $\mathcal{H}_1$ and $\mathcal{H}_2$ have dimensions, respectively, m and n, and if $|x\rangle \in \mathcal{H}_1$, $|y\rangle \in \mathcal{H}_2$, it results that:

- $|x\rangle \otimes |y\rangle$ is a ket of dimension mn (column vector)
- $\langle x| \otimes \langle y|$ is a bra of dimension mn (row vector).

For example, given the two kets $|x\rangle \in \mathcal{H}_1 = \mathbb{C}^2$ and $|y\rangle \in \mathcal{H}_2 = \mathbb{C}^3$

$$|x\rangle = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad |y\rangle = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

the tensor product gives

$$|x\rangle \otimes |y\rangle = \begin{bmatrix} x_1 y_1 \\ x_1 y_2 \\ x_1 y_3 \\ x_2 y_1 \\ x_2 y_2 \\ x_2 y_3 \end{bmatrix}, \qquad \langle x| \otimes \langle y| = \left[x_1^* y_1^*, x_1^* y_2^*, x_1^* y_3^*, x_2^* y_1^*, x_2^* y_2^*, x_2^* y_3^*\right].$$

If A is an operator of $\mathcal{H}_1$ and B is an operator of $\mathcal{H}_2$, then

- $A \otimes B$ is an operator to which an $mn \times mn$ square matrix must be associated.

The following general properties can also be established:

(1) If $\{|b_i\rangle, i \in I\}$ is a basis for $\mathcal{H}_1$ and $\{|c_j\rangle, j \in J\}$ is a basis for $\mathcal{H}_2$, then

$$\{|b_i\rangle \otimes |c_j\rangle, i \in I, j \in J\} \tag{2.106}$$

is a basis for $\mathcal{H}_1 \otimes \mathcal{H}_2$.

(2) $\dim(\mathcal{H}_1 \otimes \dim \mathcal{H}_2) = \dim \mathcal{H}_1 \ \dim \mathcal{H}_2$.

(3) If $\{\lambda_i, i \in I\}$ is the spectrum of A and $\{\mu_j, j \in J\}$ is the spectrum of B, the spectrum of $A \otimes B$ results in

$$\sigma(A \otimes B) = \{\lambda_i \, \mu_j, \ i \in I, j \in J\}. \tag{2.107}$$

Analogously, the eigenvalues of $A \otimes B$ are given by $\{|\lambda_i\rangle \otimes |\mu_i\rangle\}$.

(4) If A and B are (unitary) Hermitian operators, also $A \otimes B$ is a (unitary) Hermitian operator.

(5) If A and B are positive definite Hermitian operators, also $A \otimes B$ is a positive definite Hermitian operator.

(6) For the trace, the simple rule holds that

$$\text{Tr}[A \otimes B] = \text{Tr}[A] \ \text{Tr}[B]. \tag{2.108}$$

Final Comment on Tensor Product

As mentioned, the tensor product appears in Postulate 4 of Quantum Mechanics. The properties of this product, albeit with a rather heavy symbolism, seem natural enough at first glance. However, just these apparently "intuitive" properties lead to paradoxical consequences, which are at the foundations of very interesting applications, as we will see at the end of the following chapter.

Problem 2.12 ⋆ Prove that if A and B are Hermitian operators, also $A \otimes B$ is a Hermitian operator.

Problem 2.13 ⋆⋆ Establish the compatibility conditions for the dimensions of the matrices in the mixed-product law (2.104).

Problem 2.14 ⋆⋆ Prove property (2.107) of the Kronecker product and, more specifically, prove that, if λ is an eigenvalue of A with eigenvector $|\lambda\rangle$ and μ is an eigenvalue of B with eigenvector $|\mu\rangle$, then $\lambda\mu$ is an eigenvalue of $A \otimes B$ with eigenvector $|\lambda\rangle \otimes |\mu\rangle$.

Problem 2.15 ⋆ ⋆ ⋆ The mixed-product law can be extended in several ways. In particular,

$$(A_1 \otimes A_2)(B_1 \otimes B_2)(C_1 \otimes C_2) = (A_1 B_1 C_1) \otimes (A_2 B_2 C_2). \tag{E5}$$

Prove this relation using (2.104).

Problem 2.16 ⋆⋆ Prove that, if the matrices A_1 and A_2 have, respectively, the diagonalizations (see (2.87))

$$A_1 = U_1 \Lambda_1 U_1^*, \qquad A_2 = U_2 \Lambda_2 U_2^*$$

then

$$A_1 \otimes A_2 = (U_1 \otimes U_2)(\Lambda_1 \otimes \Lambda_2)(U_1^* \otimes U_2^*) \tag{E6}$$

is a diagonalization of $A_1 \otimes A_2$.

2.14 Other Fundamentals Developed Throughout the Book

This chapter developed the essential fundamentals necessary for the comprehension of the elements of Quantum Mechanics that will be used in the next chapter, which in turn are indeed required in the study of Quantum Communications systems developed in Part II.

On the other hand, the mathematics encountered in the field of Quantum Mechanics is very extensive and a further development of fundamentals is out of the scope of this book. Considering our philosophy of introducing the needed preliminaries in a gradual form, a few fundamentals, which will be needed in Part III on Quantum Information, will be introduced just before describing the applications. We mention in particular the EID with a continuous spectrum and the partial trace.

References

1. S. Roman, *Advanced Linear Algebra* (Springer, New York, 1995)
2. G. Cariolaro, *Unified Signal Theory* (Springer, London, 2011)
3. P.A.M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, Oxford, 1958)
4. R.A. Horn, C.R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1998)
5. M.A. Nielsen, I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000)