Working notes for an introductory course on quantum information theory

Contents

This course will be about the **theoretical foundations of information processing**. Along this course, the word "information" will be used to denote the ability to distinguish between alternatives. Such an ability can be *encoded* in a physical object, which can be *communicated* from one place to another and *transformed* into another physical object. It is important to introduce a "fundamental unit", with respect to which one quantifies information. One *bit* of information is an abstract quantity defined as the amount of information contained in the answer to the question "which one between two alternatives?" M bits can (and are necessary to) indicate one among 2^M alternatives.

One **abstract bit** of information is always encoded in the state of a **physical object**, which admits *two distinguishable states*. Information is acquired by performing an **observation** upon the physical object carrying it. To understand how information can be acquired and processed at the quantum level, we need to **understand which states, which transformations, and which observations are allowed by Quantum Theory** (QT). We will find that the theory of information constructed upon QT is substantially different from its classical analogue: tasks like copying and deleting are generally impossible within QT, while new protocols like quantum teleportation and super-dense coding will arise naturally.

☞ Paragraphs denoted by the symbol "☞" usually contains some remark, whose explanation is only sketched, leaving the complete proof of the claim to the reader.

1 Basic mathematical tools

1.1 The space $\mathbb{C}^n$ with the dot-product

The mathematical framework of Quantum Theory is described by linear algebra. Therefore, it is necessary to recall some basic notions in linear algebra.

Definition 1.1 (Matrices)**.** Let m and n be two positive integer numbers. An $m \times n$ complex matrix is an array of mn complex numbers c_{ij}, $1 \leq i \leq m$, $1 \leq j \leq n$, arranged as

$$\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix}.$$

We will usually denote matrices by capital letters $A, B, C, \ldots$. The set of all $m \times n$ matrices forms a linear space, in the sense that, given two $m \times n$ matrices A and B, their linear composition $\alpha A + \beta B$, with $\alpha, \beta \in \mathbb{C}$, given by

$$\begin{aligned} \alpha A + \beta B &= \alpha \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} + \beta \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} \\ &\equiv \begin{pmatrix} \alpha a_{11} + \beta b_{11} & \alpha a_{12} + \beta b_{12} & \cdots & \alpha a_{1n} + \beta b_{1n} \\ \alpha a_{21} + \beta b_{21} & \alpha a_{22} + \beta b_{22} & \cdots & \alpha a_{2n} + \beta b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m1} + \beta b_{m1} & \alpha a_{m2} + \beta b_{m2} & \cdots & \alpha a_{mn} + \beta b_{mn} \end{pmatrix}, \end{aligned}$$

is also a complex $m \times n$ matrix. Matrices can also be *multiplied* according to the matrix-multiplication rule, i.e. given an $m \times n$ matrix A and an $n \times p$ matrix B, the matrix $C := AB$

is defined as the $m \times p$ matrix with matrix elements c_{ij} given by $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$, for every $1 \leq i \leq m$, $1 \leq j \leq p$.

☞ A complex number, then, can be seen as a 1×1 complex matrix.

Definition 1.2 (Complex Vector Spaces)**.** Let n be a positive integer number. The set of all $n \times 1$ complex matrices

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

forms an n-dimensional complex vector space, denoted by $\mathbb{C}^n$, where addition and scalar multiplication are the same as for matrices. Elements of $\mathbb{C}^n$ are called (column) vectors and will be usually denoted by lower-case Greek letters $\psi, \phi, \chi, \ldots$. Any $m \times n$ matrix A induces a linear operator $A : \mathbb{C}^n \to \mathbb{C}^m$, whose action on vectors ψ in $\mathbb{C}^n$ is given by $A\psi$, understood as the multiplication of an $m \times n$ matrix with an $n \times 1$ matrix, resulting in an $m \times 1$ matrix.

☞ The so-called *identity matrix* on $\mathbb{C}^n$ is the $n \times n$ matrix with 1's on its diagonal, and 0's everywhere else. The symbol used to denote such matrix is $\mathbb{1}_n$.

☞ The set of $m \times n$ complex matrices will be denoted by $\mathbb{M}(\mathbb{C}^n, \mathbb{C}^m)$. The set of square $n \times n$ complex matrices will be denoted by $\mathbb{M}(\mathbb{C}^n)$. The reason to adopt such notation will be made clear later.

Definition 1.3 (Dot product)**.** For a positive integer number n, let us consider the complex vector space $\mathbb{C}^n$. Given two vectors $\psi, \phi \in \mathbb{C}^n$,

$$\psi = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \quad \text{and} \quad \phi = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix},$$

the *dot product* between ψ and ϕ, denoted as $\langle \psi, \phi \rangle$, is defined as

$$\langle \psi, \phi \rangle := \sum_{i=1}^{n} c_i^* d_i \in \mathbb{C}.$$

According to the matrix multiplication rule, we have that

$$\langle \psi, \phi \rangle = \begin{pmatrix} c_1^* & c_2^* & \cdots & c_n^* \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix}.$$

Two vectors are called *orthogonal* if their dot product is equal to zero. The *norm* of a vector $\psi \in \mathbb{C}^n$ is defined as $\|\psi\| := \sqrt{\langle \psi, \psi \rangle}$. □

Definition 1.4 (Standard basis)**.** Let n be a positive integer number. The set of vectors $\mathbf{e} = \{e_i : 1 \leq i \leq n\}$, where

$$e_i := \begin{pmatrix} \delta_{1i} \\ \delta_{2i} \\ \vdots \\ \delta_{ni} \end{pmatrix}, \quad \text{with} \quad \delta_{kl} = \begin{cases} 1, & k = l, \\ 0, & k \neq l, \end{cases}$$

constitutes the so-called *standard basis* of $\mathbb{C}^n$. Every vector $\psi \in \mathbb{C}^n$ can be expanded as $\psi = \sum_{i=1}^n \langle e_i, \psi \rangle e_i$. □

☞ It is easy to check that the set **e** defined above is a set of orthogonal vectors, all with norm equal to 1. Vectors which are pairwise orthogonal and all norm-one are called, in short, *orthonormal*.

1.2 Complex inner product spaces: Hilbert spaces

From a more formal level, it is useful to introduce the following notion of an *abstract* linear space equipped with an inner product, generalizing the notion of the vector space $\mathbb{C}^n$ with the dot product:

Definition 1.5 (Hilbert spaces)**.** A (complex) *Hilbert space* is a vector space $\mathcal{H}$ over the field of complex numbers $\mathbb{C}$, equipped with a function, called *inner (or scalar) product*, $\langle \phi, \psi \rangle \in \mathbb{C}$, such that
1. $\langle \phi, \psi \rangle = \langle \psi, \phi \rangle^*$ (hermitian symmetric);
2. $\langle \phi, c_1\psi_1 + c_2\psi_2 \rangle = c_1\langle \phi, \psi_1 \rangle + c_2\langle \phi, \psi_2 \rangle$ (right linear);
3. $0 \leq \langle \phi, \phi \rangle < \infty$, $\forall \phi \in \mathcal{H}$ (positive semi-definite);
4. $\langle \phi, \phi \rangle = 0$ if and only if $\phi = 0$ (non-degenerate). □

☞ For a mathematically rigorous definition, the following condition must be added: the *norm* $\|\phi\| := \sqrt{\langle \phi, \phi \rangle}$ turns $\mathcal{H}$ into a complete metric space. Without going here into an explanation of what this exactly means, we just notice that, in all the examples treated in this course, such condition is *always* and *automatically* satisfied, so that we can effectively "forget" about it.

☞ Note that, from properties 1 and 2, it follows that $\langle c_1\phi_1 + c_2\phi_2, \psi \rangle = c_1^*\langle \phi_1, \psi \rangle + c_2^*\langle \phi_2, \psi \rangle$, i.e. the inner product is left anti-linear.

Definition 1.6. Two vectors $\phi, \psi \in \mathcal{H}$ are called *orthogonal* if $\langle \phi, \psi \rangle = 0$. A vector $\psi \in \mathcal{H}$ is called *normalized* if $\|\psi\| := \sqrt{\langle \psi, \psi \rangle} = 1$.

Proposition 1.1 (Complete orthonormal system)**.** *A family $\{\psi_i\}_i$ of elements in a Hilbert space $\mathcal{H}$ contitutes an* orthonormal system *(abbrev.* ONS*) if $\langle \psi_i, \psi_j \rangle = \delta_{ij}$. For any orthonormal system $\{\psi_i\}_i$ and any $\phi \in \mathcal{H}$,*

$$\|\phi\|^2 \leq \sum_i |\langle \psi_i, \phi \rangle|^2.$$

If the equation above holds with equality for all $\phi \in \mathcal{H}$, then the orthonormal system $\{\psi_i\}_i$ is called complete *(abbrev.* CONS*). For any complete orthonormal system $\{\psi_i\}_i$ and any $\phi \in \mathcal{H}$, the following* expansion formula *holds:*

$$\phi = \sum_i \langle \psi_i, \phi \rangle \psi_i. \qquad \square \tag{1.1}$$

☞ Everywhere in this course, we assume that for any Hilbert space $\mathcal{H}$ there exists a complete orthonormal system with a *finite* number of elements. It can be proved that such a number is uniquely defined. It is called the *dimension* of $\mathcal{H}$, and it is denoted by $\dim \mathcal{H}$.

☞ **Canonical isomorphism**. Let $\mathcal{H}$ a d-dimensional Hilbert space. Let us fix in $\mathcal{H}$ a complete orthonormal system $\{\psi_i : 1 \leq i \leq d\}$. We can then construct an isomorphism $\mathcal{H} \leftrightarrow \mathbb{C}^d$ as follows:

$$\psi_i \longleftrightarrow e_i \equiv \begin{pmatrix} \delta_{1i} \\ \delta_{2i} \\ \vdots \\ \delta_{di} \end{pmatrix}. \tag{1.2}$$

By means of the expansion formula (1.1), the correspondence constructed above induces a one-to-one correspondence between elements in $\mathcal{H}$ with elements in $\mathbb{C}^d$ as follows:

$$\phi \longleftrightarrow \begin{pmatrix} \langle \psi_1, \phi \rangle \\ \langle \psi_2, \phi \rangle \\ \vdots \\ \langle \psi_d, \phi \rangle \end{pmatrix}. \tag{1.3}$$

With this correspondence, the inner product $\langle , \rangle$ in $\mathcal{H}$ becomes the dot product in $\mathbb{C}^d$. For these reasons, in the following, every d-dimensional Hilbert space where a complete orthonormal system has been fixed, will be considered as being, essentially, $\mathbb{C}^d$ with the dot product. We will refer to such isomorphism as the *canonical isomorphism*.

Theorem 1.1 (Cauchy-Schwarz inequality)**.** *For any Hilbert space $\mathcal{H}$ and any $\psi, \phi \in \mathcal{H}$,*

$$|\langle \psi, \phi \rangle| \leq \|\psi\| \, \|\phi\| . \tag{1.4}$$

Proof. If $\psi = 0$ then the inequality holds. We will then assume $\psi \neq 0$. Let us define $\omega := \phi - \frac{\langle \psi, \phi \rangle}{\langle \psi, \psi \rangle} \psi$. By construction, then, $\langle \psi, \omega \rangle = 0$. But then, $\|\phi\|^2 = \left\| \omega + \frac{\langle \psi, \phi \rangle}{\langle \psi, \psi \rangle} \psi \right\|^2 = \|\omega\|^2 + \left\| \frac{\langle \psi, \phi \rangle}{\langle \psi, \psi \rangle} \psi \right\|^2 \geq \frac{|\langle \psi, \phi \rangle|^2}{\|\psi\|^4} \|\psi\|^2$, i.e. $\|\phi\|^2 \|\psi\|^2 \geq |\langle \psi, \phi \rangle|^2$. ∎

In proving the Cauchy-Schwarz inequality, we decomposed a given vector ϕ as a sum of two vectors, $\omega + \frac{\langle \psi, \phi \rangle}{\langle \psi, \psi \rangle} \psi$, where the first one (i.e. ω) is orthogonal to ψ, while the second component is parallel to ψ. It is then easy to recognize that the component $\frac{\langle \psi, \phi \rangle}{\langle \psi, \psi \rangle} \psi$ represents the *orthogonal projection* of ϕ onto ψ:

Definition 1.7 (Orthogonal Projection)**.** Let $\psi \in \mathcal{H}$ be a non-zero vector. Then, for any $\phi \in \mathcal{H}$, the *orthogonal projection* of ϕ onto ψ is given by the action of the following operator:

$$\Pi_\psi(\phi) := \frac{\langle \psi, \phi \rangle}{\langle \psi, \psi \rangle} \psi = \left\langle \frac{\psi}{\|\psi\|}, \phi \right\rangle \frac{\psi}{\|\psi\|}.$$

As a consequence of the right-linearity of the inner product, the operator Π_ψ is linear, i.e. for any $c_1, c_2 \in \mathbb{C}$ and $\phi_1, \phi_2 \in \mathcal{H}$, $\Pi_\psi(c_1\phi_1 + c_2\phi_2) = c_1\Pi_\psi(\phi_1) + c_2\Pi_\psi(\phi_2)$.

1.3 Linear mappings and Dirac's *bra-ket* notation

As we noticed in Definition 1.7, the orthogonal projection is linear in its input ϕ. More generally, we have the following definition:

Definition 1.8. Let $\mathcal{H}$ and $\mathcal{K}$ be two Hilbert spaces. A mapping $F : \mathcal{H} \to \mathcal{K}$ is called *linear* if, for any $\psi_1, \psi_2 \in \mathcal{H}$ and any $c_1, c_2 \in \mathbb{C}$, $F(a_1\psi_1 + a_2\psi_2) = a_1F(\psi_1) + a_2F(\psi_2)$.

Theorem 1.2 (Riesz representation theorem)**.** *For any linear mapping $F : \mathcal{H} \to \mathbb{C}$ (usually called linear functional), there exists a unique $\phi \in \mathcal{H}$ such that $F(\psi) = \langle \phi, \psi \rangle$, for all $\psi \in \mathcal{H}$.*

Theorem 1.3. *A linear mapping F is completely and uniquely specified by its action on a complete orthonormal system $\{\psi_i\}_i$ of $\mathcal{H}$. Conversely, given a complete orthonormal system $\{\psi_i\}_i$ in $\mathcal{H}$ and any vectors $\{\phi_i\}$ in $\mathcal{K}$, there exists a unique linear mapping $F : \mathcal{H} \to \mathcal{K}$ such that $F(\psi_i) = \phi_i$, for all i.*

Indeed, once the vectors $F(\psi_i) \in \mathcal{K}$ are specified, the action of F on any vector $\phi \in \mathcal{H}$ is uniquely given by means of the expansion formula (1.1):

$$F(\phi) = F\left(\sum_i \langle \psi_i, \phi \rangle \psi_i\right) = \sum_i \langle \psi_i, \phi \rangle F(\psi_i) = \sum_i \langle \psi_i, \phi \rangle \phi_i. \tag{1.5}$$

The above equation shows that any linear mapping $F : \mathcal{H} \to \mathcal{K}$ can be written as a sum of elementary linear mappings $f : \mathcal{H} \to \mathcal{K}$ of the form

$$f(\bullet) := \langle \psi, \bullet \rangle \omega, \tag{1.6}$$

where $\psi \in \mathcal{H}$ and $\omega \in \mathcal{K}$, and the symbol "$\bullet$" is used to denote any input vector in $\mathcal{H}$. The "bra-ket notation" introduced by Dirac is very convenient to denote such elementary linear mappings:

Definition 1.9 (Dirac *bra-ket* notation)**.** Let $\mathcal{H}$ be a d-dimensional Hilbert space. For any $\psi \in \mathcal{H}$, the Dirac's *ket* symbol $|\psi\rangle$ is used to denote the vector ψ, while the Dirac's *bra* symbol $\langle \psi|$ is used to denote the linear functional $\langle \psi, \bullet \rangle : \mathcal{H} \to \mathbb{C}$. In an expression, when a bra and a ket appear consecutively, i.e. $\langle \phi || \psi \rangle$, one bar is dropped, and the resulting symbol $\langle \phi | \psi \rangle$ (that is, a *bra(c)ket*!) represents the inner product $\langle \phi, \psi \rangle$. Correspondingly, an elementary linear mapping as that in Eq. (1.6) can be conveniently represented as *outer product* as follows:

$$f \longleftrightarrow |\omega\rangle\langle\psi|,$$

so that the action of f can be written as a multiplication of objects from left to right, i.e.

$$|\omega\rangle\langle\psi| \ |\phi\rangle = |\omega\rangle\langle\psi|\phi\rangle \longleftrightarrow \omega\langle\psi, \phi\rangle.$$

☞ In terms of the correspondence $\mathcal{H} \leftrightarrow \mathbb{C}^d$, Dirac's ket symbol $|\psi\rangle$ is represented by a column vector as follows:

$$|\psi\rangle \equiv \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_d \end{pmatrix},$$

while Dirac's bra symbol $\langle \psi|$ is represented by a row vector as follows:

$$\langle \psi| \equiv \begin{pmatrix} c_1^* & c_2^* & \cdots & c_d^* \end{pmatrix}.$$

Therefore, while the correspondence $\psi \mapsto |\psi\rangle$ is linear, i.e. $|a_1\psi_1 + a_2\psi_2\rangle = a_1|\psi_1\rangle + a_2|\psi_2\rangle$, the correspondence $\phi \mapsto \langle \phi|$ is anti-linear, i.e. $\langle b_1\phi_1 + b_2\phi_2| = b_1^*\langle\phi_1| + b_2^*\langle\phi_2|$.

☞ From Eq. (1.5), any linear mapping $F : \mathcal{H} \to \mathcal{K}$ can be written as a sum $F = \sum_j |F(\psi_j)\rangle\langle\psi_j|$. Let now $\{\omega_i\}_i$ be an orthonormal system in $\mathcal{K}$. Let us expand $F(\psi_j) = \sum_i \langle \omega_i | F(\psi_j) \rangle \omega_i$. Correspondingly, we can write $F = \sum_{ij} \langle \omega_i | F(\psi_j) \rangle |\omega_i\rangle\langle\psi_j|$. According to the canonical isomorphism $\mathcal{H} \leftrightarrow \mathbb{C}^n$, $\mathcal{K} \leftrightarrow \mathbb{C}^m$, the mapping F is represented by the $m \times n$ matrix $[[f_{ij}]_{i=1}^m]_{j=1}^n \in \mathbb{M}(\mathbb{C}^n, \mathbb{C}^m)$, where $f_{ij} = \langle \omega_i | F(\psi_j) \rangle$.

Example 1.1. How to write the identity mapping $I : \mathcal{H} \to \mathcal{H}$ in Dirac's notation? Let $\{\psi_i\}_{i=1}^n$ be a complete orthonormal system for $\mathcal{H}$. Then the expansion formula (1.1) holds, i.e.

$$|\phi\rangle = \sum_{i=1}^n |\psi_i\rangle\langle\psi_i|\phi\rangle,$$

for any $\phi \in \mathcal{H}$. Therefore, the identity mapping $I : \mathcal{H} \to \mathcal{H}$ can be written in bra-ket notation as $I = \sum_i |\psi_i\rangle\langle\psi_i|$. Notice that the choice of the orthonormal system $\{\psi_i\}_i$ does not matter: let $\{\omega_i\}_i$ be another complete orthonormal system for $\mathcal{H}$; then, again, $I = \sum_i |\omega_i\rangle\langle\omega_i|$. As expected, the matrix in $\mathbb{M}(\mathbb{C}^n)$ corresponding to I is the identity matrix $\mathbb{1}_n$.

Example 1.2. How to write the orthogonal projection on a vector $\psi \neq 0$? Remember (see Definition 1.7) that $\Pi_\psi(\phi) = \|\psi\|^{-2} \langle\psi, \phi\rangle\psi$. Therefore, in bra-ket notation, $\Pi_\psi(|\phi\rangle) = \|\psi\|^{-2} |\psi\rangle\langle\psi|\phi\rangle$, i.e.

$$\Pi_\psi = \frac{1}{\|\psi\|^2} |\psi\rangle\langle\psi|.$$

Example 1.3. How to write the orthogonal projection on a subspace $V \subset \mathcal{H}$? Let $\{v_k\}_k$ an orthonormal system spanning the subspace V. Then, the orthogonal projection onto $\Pi_V : \mathcal{H} \to V$ can be written as

$$\Pi_V = \sum_k |v_k\rangle\langle v_k|.$$

As in the case of the identity (Example 1.1), also here the choice of the orthonormal set spanning V does not matter in the definition of Π_V.

1.4 Kernel, Image, and rank of a linear map

kernel, image, rank
isomorphisms

Exercise 1.1. What is the rank of an elementary linear mapping as that in Eq. (1.6)?

1.5 The algebra $\mathscr{L}(\mathcal{H})$

The set $\mathscr{L}(\mathcal{H})$ is defined as the set of all linear mappings $F : \mathcal{H} \to \mathcal{H}$. Such a set is in fact an algebra, since it is a vector space with the natural composition rule $G \circ F$.

☞ Let d be the dimension of $\mathcal{H}$. We saw before that, by fixing a complete orthonormal system in $\mathcal{H}$, $\mathcal{H}$ becomes equivalent to $\mathbb{C}^d$, and the set $\mathscr{L}(\mathcal{H})$ becomes equivalent to the set $\mathbb{M}(\mathbb{C}^d)$ of $d \times d$ matrices of complex numbers. We will denote the complete orthonormal system fixed in $\mathcal{H}$ as $\mathbf{e} = \{e_i : 1 \leq i \leq d\}$.

For any $A \in \mathscr{L}(\mathcal{H})$, the following definitions are given:

- the action of A on a ket $|\psi\rangle$ is defined as $A|\psi\rangle := |A\psi\rangle$;
- A can be written as $\sum_{i,j} a(i,j)|e_i\rangle\langle e_j|$, with $a(i,j) \in \mathbb{C}$; the square $d \times d$ matrix of numbers $[[a(i,j)]]_{ij}$ is the *matrix representation* (with respect to $\mathbf{e}$) of A (the matrix representation depends on the choice of basis);
- the *kernel* of A is the linear subspace $\mathsf{Ker}A := \{\psi \in \mathcal{H} : A\psi = 0\}$; the *nullity* of A is defined as the dimension of $\mathsf{Ker}A$; the *support* of A is the linear subspace $\mathsf{Supp}A := (\mathsf{Ker}A)^\perp$; the *range* or *image* of A is the linear subspace $\mathsf{Rng}A := \{A\psi : \psi \in \mathcal{H}\}$; the *rank* of A is defined as $\mathsf{r}(A) := \dim \mathsf{Supp}A = \dim \mathsf{Rng}A$; the *rank-nullity theorem* states that $\mathsf{r}(A) + \dim \mathsf{Ker}A = d$;
- the *Hermite conjugate* (or *adjoint* or *dagger*) operator $A^\dagger$ is defined by the relation $\langle\phi, A^\dagger\psi\rangle := \langle A\phi, \psi\rangle = \langle\psi, A\phi\rangle^*$, for all $\phi, \psi \in \mathcal{H}$; the adjoint of a product AB is equal to $B^\dagger A^\dagger$;

- A is *normal* if $[A, A^\dagger] := AA^\dagger - A^\dagger A = 0$, i.e. if A and $A^\dagger$ *commute*;
- A is *self-adjoint* (or *Hermitian*) if $A^\dagger = A$; A is *skew-Hermitian* if $A^\dagger = -A$; any Hermitian and skew-Hermitian operators are, in particular, normal;
- a self-adjoint operator A is an *orthogonal projector* if $A^2 = A$ and $\|A\| = 1$;
- A is *unitary* if $AA^\dagger = A^\dagger A = \mathbb{1}$; any unitary operator is, in particular, normal;
- the *norm of a linear operator* A is defined as $\|A\| := \max_{\|\psi\|=1} \|A\psi\|$;
- the *complex conjugate* (with respect to $\mathbf{e}$) of A is defined as $A^* := \sum_{i,j} a(i,j)^* |e_i\rangle\langle e_j|$ (complex conjugation depends on the choice of basis); the complex conjugation of a product AB is equal to A^*B^*;
- the *transpose* (with respect to $\mathbf{e}$) of A is defined as $A^T := \sum_{i,j} a(i,j)|e_j\rangle\langle e_i|$ (transposition depends on the choice of basis); the transposition of a product AB is equal to $B^T A^T$;
- the hermitian conjugate of a linear operator A can also be expressed as the linear operator corresponding to the matrix $(A^*)^T = (A^T)^*$; notice that, even though both complex conjugation and transposition are basis dependent, the hermitian conjugate is basis independent, as noticed above;
- Optional: the *pseudoinverse* A^{-1} is uniquely defined by the four conditions (i) $AA^{-1}A = A$; (ii) $A^{-1}AA^{-1} = A^{-1}$; (iii) $(A^{-1}A)^\dagger = A^{-1}A$; (iv) $(AA^{-1})^\dagger = AA^{-1}$; A is *invertible* if $A^{-1}A = AA^{-1} = I$, in which case A^{-1} is called the *inverse* of A;
- the *trace* of A is defined as $\operatorname{Tr} A := \sum_{i=1}^d \langle e_i|A|e_i\rangle = \sum_i a(i,i)$ (Theorem 1.4 below shows that the trace is a basis-independent quantity); the trace is linear, i.e. $\operatorname{Tr}[a_1A_1 + a_2A_2] = a_1 \operatorname{Tr} A_1 + a_2 \operatorname{Tr} A_2$, for any $A_1, A_2 \in \mathscr{L}(\mathcal{H})$ and $a_1, a_2 \in \mathbb{C}$; finally, one can directly verify that $\operatorname{Tr}[A\,|\psi\rangle\langle\phi|] = \langle\phi|A|\psi\rangle = \langle\phi, A\psi\rangle$, for any A, ψ, ϕ.

☞ If $\mathbf{e} = \{e_i\}$ and $\mathbf{f} = \{f_i\}$ are two complete orthonormal systems for $\mathcal{H}$, then there exists a unitary operator $U \in \mathscr{L}(\mathcal{H})$ such that $f_i = Ue_i$, for all i.

Theorem 1.4 (Cyclicity and invariance of trace)**.** *The trace operation* Tr *satisfies* $\operatorname{Tr} AB = \operatorname{Tr} BA$*, for any* $A, B \in \mathscr{L}(\mathcal{H})$*. This implies that* $\operatorname{Tr} A$ *does not depend on the choice of the basis* $\mathbf{e}$ *used to compute it.*

Proof. Let us expand A and B as $A = \sum_{i,j} a(i,j)|e_i\rangle\langle e_j|$ and $B = \sum_{k,l} b(k,l)|e_k\rangle\langle e_l|$. Then, $AB = \sum_{i,j,l} a(i,j)b(j,l)|e_i\rangle\langle e_l|$ and $BA = \sum_{k,l,j} b(k,l)a(l,j)|e_k\rangle\langle e_j|$, where we used the fact that $\langle e_j|e_k\rangle = \delta_{j,k}$. This implies that $\operatorname{Tr} AB = \sum_{i,j} a(i,j)b(j,i) = \sum_{k,l} b(k,l)a(l,k) = \operatorname{Tr} BA$.

Now, suppose that we are given another orthonormal basis $\mathbf{f} = \{f_1, \cdots, f_d\}$ for $\mathcal{H}$. We know that there exists a unitary operator U such that $|f_i\rangle = U|e_i\rangle$, for all i. The trace of A computed with respect to the basis $\mathbf{f}$ is equal to $\operatorname{Tr}_{\mathbf{f}} A := \sum_i \langle f_i|A|f_i\rangle = \sum_i \langle e_i|U^\dagger AU|e_i\rangle = \operatorname{Tr} U^\dagger AU = \operatorname{Tr} AUU^\dagger = \operatorname{Tr} A$. ∎

Remark 1.1. Another way to introduce the trace operation, such that the independence of the particular choice of basis is made apparent from the beginning, is to define

$$\operatorname{Tr}[|u\rangle\langle v|] = \langle v|u\rangle,$$

for any $u, v \in \mathcal{H}$, and then extend this definition by linearity to any linear operator, via decomposition (1.5).

Theorem 1.5 (Conditions for equality)**.** *For any $A, B \in \mathscr{L}(\mathcal{H})$, the following are equivalent:*

- $A = B$*;*
- $A\psi = B\psi$*, for all* $\psi \in \mathcal{H}$*;*
- $\mathrm{Tr}[A\,|\psi\rangle\langle\phi|] = \mathrm{Tr}[B\,|\psi\rangle\langle\phi|]$*, for all* $\psi, \phi \in \mathcal{H}$*;*
- $\mathrm{Tr}[A\,|\psi\rangle\langle\psi|] = \mathrm{Tr}[B\,|\psi\rangle\langle\psi|]$*, for all* $\psi \in \mathcal{H}$*.*

Proof. The first three conditions essentially represent the definition of the identity $A = B$. The last condition is a consequence of the generalized polarization identity, which can be expressed as

$$\langle\phi, A\psi\rangle = \frac{1}{4}\sum_a a\langle a\phi + \psi, A(a\phi + \psi)\rangle, \qquad a \in \{1, -1, i, -i\}. \tag{1.7}$$

∎

Remark 1.2. Up to here, and until the end of the notes, we take the underlying field to be the set of complex numbers $\mathbb{C}$. It seems important to notice, therefore, that some of the results given so far holds only in such a case. For example, the last condition of Theorem 1.5 is not equivalent to the other three if the underlying field is the set of real numbers $\mathbb{R}$. This can be easily seen by considering, e.g., the operator $T(x, y) = (-y, x)$ acting on $\mathbb{R}^2$. Then, for any $\psi \in \mathbb{R}^2$, $\langle\psi|T(\psi)\rangle = 0$, even though $T \neq 0$.

1.6 The Spectral Theorem and other canonical forms

Theorem 1.6 (Spectral theorem, matrix form)**.** *$A \in \mathscr{L}(\mathcal{H})$ is normal, if and only if there exists a unitary operator U and a diagonal (w.r.t.* **e***) matrix $\Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \cdots, \lambda_d]$, $\lambda_i \in \mathbb{C}$, such that $A = U\Lambda U^\dagger$. The diagonal entries of Λ are the* eigenvalues *of A, and the columns of U are the corresponding* eigenvectors *and they are orthonormal. If the same eigenvalue appears in Λ more than once, that eigenvalue is said to be* degenerate*. For each eigenvalue λ_i, the* degeneracy parameter *ν_i is the number of times the eigenvalue λ_i appears in Λ.*

☞ A is self-adjoint if and only if $\lambda_i \in \mathbb{R}$, for all i. A is unitary if and only if $|\lambda_i| = 1$, for all i.

☞ One can restate the above theorem as follows: $A \in \mathscr{L}(\mathcal{H})$ is normal, if and only if there exists a complete orthonormal system $\{\psi_k : 1 \leq k \leq d\}$ in $\mathcal{H}$ such that $A = \sum_{k=1}^d \lambda_k |\psi_k\rangle\langle\psi_k|$.

☞ Any orthogonal projector A is self-adjoint, and hence it is, in particular, normal. Therefore, it can be written as $A = \sum_{j=1}^d \lambda_j |\alpha_j\rangle\langle\alpha_j|$, where the vectors $\alpha_j \in \mathcal{H}$ form a complete orthonormal system. By the condition $A^2 = A$, every λ_j must be either one or zero. Recalling that the rank of A, denoted by $\mathsf{r}(A)$ is the dimension of $\mathsf{Supp}A$, one can conclude that for any orthogonal projector A, there exists a complete orthonormal system $\{\alpha_j : 1 \leq j \leq d\}$ in $\mathcal{H}$, such that A can be written as $A = \sum_{j=1}^{\mathsf{r}(A)} |\alpha_j\rangle\langle\alpha_j|$. Conversely, for any given (not necessarily complete) orthonormal system $\mathbf{b} = \{\beta_k : 1 \leq k \leq s\}$, the operator $\Pi_{\mathbf{b}} := \sum_{k=1}^s |\beta_k\rangle\langle\beta_k|$ is the orthogonal projector onto the subspace $\mathcal{S} = \mathrm{span}\{\beta_k; 1 \leq k \leq s\} \equiv \mathsf{Supp}\Pi_{\mathbf{b}} = \mathsf{Rng}\Pi_{\mathbf{b}}$.

Remark 1.3. As a consequence of the Spectral Theorem, any normal operator A is unitarily equivalent to its transpose A^T. This can be easily proved by noticing that the condition $A = U\Lambda U^\dagger$ implies that $A^T = U^*\Lambda U^T$, or, equivalently, $\Lambda = U^T A^T U^*$, since $\Lambda^T = \Lambda$. Then, $A = UU^T A^T U^* U^\dagger$. In general, however, this is not true. A counter-example is given in [A. George and K. D. Ikramov, Lin. Alg. Appl. **349**, 11-16 (2002)] as

$$\begin{pmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 0 & 2 & 5 \end{pmatrix}. \tag{1.8}$$

The proof relies on advanced techniques and intensive numerical search. For a complete characterization of matrices which are unitarily equivalent to their transpose, see [S. R. Garcia and J. E. Tener, arXiv:0908.2107v3]. On the other hand, it is important to stress that A and A^T always have the same eigenvalues. (This is because $\det(A - \lambda\mathbb{1}) = 0$ if and only if $\det(A^T - \lambda\mathbb{1}) = 0$, since the determinant is invariant under transposition.) □

According to the Spectral Theorem above, a matrix $A \in \mathscr{L}(\mathcal{H})$ is normal, i.e. $A^\dagger A = AA^\dagger$, if and only if there exists a complete orthonormal system $\{\psi_k : 1 \leq k \leq d\} \subset \mathcal{H}$ and a family of complex numbers $\{\lambda_k : 1 \leq k \leq d\} \subset \mathbb{C}$, such that $A = \sum_{k=1}^d \lambda_k |\psi_k\rangle\langle\psi_k|$. We know that some of the λ_k's can be equal: we can always reorder the terms of the sum in such a way that, in general, the list of the λ_k's will look like the following:

$$\underbrace{\lambda_1 = \lambda_2 = \cdots = \lambda_{\nu_1}}_{\mu_1 \text{ rep. } \nu_1 \text{ times}} \neq \underbrace{\lambda_{\nu_1+1} = \cdots = \lambda_{\nu_1+\nu_2}}_{\mu_2 \text{ rep. } \nu_2 \text{ times}} \neq \underbrace{\lambda_{\nu_1+\nu_2+1} = \cdots}_{\cdots} \neq \underbrace{\lambda_{\sum_{i=1}^{\ell-1}\nu_i+1} = \cdots = \lambda_d}_{\mu_\ell \text{ rep. } \nu_\ell \text{ times}}. \tag{1.9}$$

According to this reordering, each eigenvalue λ_k of A is addressed by specifying two numbers, i.e. a complex number μ_i and an integer j between 1 and ν_i. The eigevector ψ_k is correspondingly denoted as $\psi_{(\mu_i,j)}$. The general form of a normal operator A can then be written as:

$$A = \sum_{i=1}^{\ell} \sum_{j=1}^{\nu_i} \mu_i |\psi_{(\mu_i,j)}\rangle\langle\psi_{(\mu_i,j)}|, \tag{1.10}$$

where $\ell \equiv \ell(A)$ is the number of *distinct* eigenvalues of A.

The notation is further simplified by dropping the letter ψ in the bra's and ket's, and by taking the first sum to run over all distinct eigenvalues of A, so that

$$A = \sum_{\mu} \sum_{j=1}^{\nu(\mu)} \mu |\mu, j\rangle\langle\mu, j|. \tag{1.11}$$

For each distinct eigenvalue μ of A, we introduce the following orthogonal projectors

$$\Pi^A(\mu) := \sum_{j=1}^{\nu(\mu)} |\mu, j\rangle\langle\mu, j|, \tag{1.12}$$

so that

$$A = \sum_{\mu} \mu \Pi^A(\mu). \tag{1.13}$$

Theorem 1.7 (Spectral theorem, operator form)**.** *$A \in \mathscr{L}(\mathcal{H})$ is normal, if and only if there exists a set of complex numbers $\{\mu_i\}_i$, with $\mu_i \neq \mu_j$, and a set of orthogonal projectors $\{\Pi_i^A\}_i$, with $\Pi_i^A \Pi_j^A = \delta_{ij}\Pi_i^A$ and $\sum_i \Pi_i^A = I$, such that*

$$A = \sum_i \mu_i \Pi_i^A.$$

Theorem 1.8 (Simultaneous diagonalisation)**.** *A family of operators* $(A_k; k = 1, \cdots, K)$ *is* simultaneous diagonalisable*, i.e. there exists a unitary operator* U *such that* $A_k = U\Lambda_k U^\dagger$ *for all* k*, if and only if the family is* commutative*, i.e.* $[A_k, A_{k'}] = 0$*, for all* k *and* k'*.*

Definition 1.10 (Positive operators)**.** The operator A is said to be *positive* if $\langle\psi|A|\psi\rangle \geq 0$, for all $\psi \in \mathcal{H}$; A is said to be *strictly positive* if $\langle\psi|A|\psi\rangle > 0$, for all $\psi \in \mathcal{H}$, $\psi \neq 0$.

Theorem 1.9 (Conditions for positivity)**.** *For any* $A \in \mathscr{L}(\mathcal{H})$*, the following are equivalent:*

- A *is positive;*
- A *is self-adjoint and all its eigenvalues are non-negative (strictly positive if and only if* A *is strictly positive);*
- $A = B^\dagger B$ *for some operator* B*;*
- $A = B^2$ *for some positive operator* B*; such a* B *is unique and it is equivalently denoted as* $A^{1/2}$ *or* $\sqrt{A}$*.*

☞ For any operator $A \in \mathscr{L}(\mathcal{H})$, the operator $A^\dagger A$ is positive. We can hence consider its square root $\sqrt{A^\dagger A}$. The resulting positive operator is called the *absolute value* of A and it is denoted by $|A|$.

☞ A 2×2 self-adjoint matrix is positive, if and only if both its trace and determinant are non-negative.

Theorem 1.10 (Singular-value decomposition (SVD))**.** *For any* $A \in \mathscr{L}(\mathcal{H})$*, there exist unitary operators* $U, W \in \mathscr{L}(\mathcal{H})$ *such that*

$$A = U\Sigma W, \tag{1.14}$$

where $\Sigma = \mathrm{diag}[s_1, s_2, \cdots, s_d]$*,* $s_i \in \mathbb{R}$*, and* $s_1 \geq s_2 \geq \cdots \geq s_d \geq 0$*. The positive numbers* s_i *are called the* singular values *of* A*.*

☞ For any $A \in \mathscr{L}(\mathcal{H})$, the singular values of A are the eigenvalues of $|A|$, repeated according to their degeneracy and listed in decreasing order.

Theorem 1.11 (Polar decomposition)**.** *For any* $A \in \mathscr{L}(\mathcal{H})$*, there exists a unitary operators* $U, V \in \mathscr{L}(\mathcal{H})$ *such that*

$$A = U(A^\dagger A)^{1/2}, \qquad A = (AA^\dagger)^{1/2}V. \tag{1.15}$$

Corollary 1.1. *For any* $A \in \mathscr{L}(\mathcal{H})$*, there exists a unitary operator* $W \in \mathscr{L}(\mathcal{H})$ *such that* $A^\dagger A = W(AA^\dagger)W^\dagger$*.*

Proof. From eq. (1.15), for any $A \in \mathscr{L}(\mathcal{H})$, there exists a unitary $U \in \mathscr{L}(\mathcal{H})$ such that $A = U(A^\dagger A)^{1/2}$. This implies that $AA^\dagger = U(A^\dagger A)^{1/2}(A^\dagger A)^{1/2}U^\dagger = U(A^\dagger A)U^\dagger$. By choosing $W = U^\dagger$ the statement is proved. ∎

Definition 1.11 (Schatten p-norms)**.** Given an operator $A \in \mathscr{L}(\mathcal{H})$, the *Schatten* p*-norm* of A is defined as

$$\|A\|_p := \left[\sum_{j=1}^{d} (s_j(A))^p\right]^{1/p}, \qquad 1 \leq p \leq \infty. \tag{1.16}$$

From p-norms, we recover the usual operator norm $\|A\|$ of A as $\|A\|_\infty = s_1(A)$. The 1-norm $\|A\|_1 = \sum_{j=1}^{d} s_j(A) = \mathrm{Tr}\,|A|$ is also called *trace-norm* of A. The 2-norm $\|A\|_2 = \sqrt{\mathrm{Tr}[A^\dagger A]}$ is also called the *Hilbert-Schmidt* or *Frobenius* norm. □

Theorem 1.12. *For any $A, B \in \mathscr{L}(\mathcal{H})$ and any $0 \leq p \leq \infty$, $\|A+B\|_p \leq \|A\|_p + \|B\|_p$ and $\|AB\|_p \leq \|A\|_p \|B\|_p$. The analogue of the Cauchy-Schwarz inequality for operators is $|\operatorname{Tr}[A^\dagger B]| \leq \|A\|_2 \|B\|_2$.*

1.7 Some facts about $\mathscr{L}(\mathcal{H}, \mathcal{K})$

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be finite dimensional Hilbert spaces, possibly with $d_1 := \dim \mathcal{H}_1 \neq \dim \mathcal{H}_2 := d_2$. We denote the set of all linear operators from $\mathcal{H}_1$ to $\mathcal{H}_2$ by $\mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$. Let $\mathbf{e} = \{e_1, \cdots, e_{d_1}\}$ and $\mathbf{f} = \{f_1, \cdots, f_{d_2}\}$ be the standard bases chosen for $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. Then, the set $\mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$ becomes equivalent to the set $\mathbb{M}(\mathbb{C}^{d_1}, \mathbb{C}^{d_2})$ of $d_2 \times d_1$ complex matrices.

For any $A \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$, the following definitions are given:

- the *kernel* of A is the linear subspace $\mathsf{Ker}A := \{\psi \in \mathcal{H}_1 : A\psi = 0\} \subseteq \mathcal{H}_1$; the *support* of A is the linear subspace $\mathsf{Supp}A := (\mathsf{Ker}A)^\perp \subseteq \mathcal{H}_1$; the *range* of A is the linear subspace $\mathsf{Rng}A := \{A\psi : \psi \in \mathcal{H}_1\} \subseteq \mathcal{H}_2$; the *rank* of A is defined as $\mathsf{r}(A) := \dim \mathsf{Supp}A = \dim \mathsf{Rng}A$;

- the operator norm of A is defined as $\|A\| := \max_{\|\psi\|=1} \|A\psi\|$;

- A can be written as $\sum_{i=1}^{d_2} \sum_{j=1}^{d_1} a(i,j) |f_i\rangle\langle e_j|$, with $a(i,j) \in \mathbb{C}$; the rectangular $d_2 \times d_1$ matrix of numbers $[[a(i,j)]]_{ij}$ is the *matrix representation* (w.r.t. $\mathbf{e}$ and $\mathbf{f}$) of A (the matrix representation depends on the choice of basis);

- for $A \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$, the *complex conjugate* (w.r.t. $\mathbf{e}$ and $\mathbf{f}$) $A^* \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$ of A is defined as $A^* := \sum_{i,j} a(i,j)^* |f_i\rangle\langle e_j|$ (complex conjugation depends on the choice of basis); if $A \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$ and $B \in \mathscr{L}(\mathcal{H}_2, \mathcal{H}_3)$, $(BA)^* = B^* A^* \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_3)$;

- for $A \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$, the *transpose* (w.r.t. $\mathbf{e}$ and $\mathbf{f}$) $A^T \in \mathscr{L}(\mathcal{H}_2, \mathcal{H}_1)$ of A is defined as $A^T := \sum_{i,j} a(i,j) |e_j\rangle\langle f_i|$ (transposition depends on the choice of basis); if $A \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$ and $B \in \mathscr{L}(\mathcal{H}_2, \mathcal{H}_3)$, $(BA)^T = A^T B^T \in \mathscr{L}(\mathcal{H}_3, \mathcal{H}_1)$;

- for $A \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$, the *Hermite conjugate* (or *adjoint*) $A^\dagger \in \mathscr{L}(\mathcal{H}_2, \mathcal{H}_1)$ of A is defined by the relation $\langle \phi, A^\dagger \psi \rangle := \langle A\phi, \psi \rangle = \langle \psi, A\phi \rangle^*$, for all $\phi \in \mathcal{H}_1$ and all $\psi \in \mathcal{H}_2$; the definition of the adjoint is hence basis *independent*; equivalently, using the notations introduced above, $A^\dagger = (A^*)^T$; if $A \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$ and $B \in \mathscr{L}(\mathcal{H}_2, \mathcal{H}_3)$, $(BA)^\dagger = A^\dagger B^\dagger \in \mathscr{L}(\mathcal{H}_3, \mathcal{H}_1)$;

- the *pseudoinverse* $A^{-1} \in \mathscr{L}(\mathcal{H}_2, \mathcal{H}_1)$ is uniquely defined by the four conditions (i) $AA^{-1}A = A$; (ii) $A^{-1}AA^{-1} = A^{-1}$; (iii) $(A^{-1}A)^\dagger = A^{-1}A$; (iv) $(AA^{-1})^\dagger = AA^{-1}$; if $\dim \mathcal{H}_1 \neq \dim \mathcal{H}_2$ there is no notion of a full inverse;

- the trace operation is not defined for rectangular operators;

- conditions like normality and self-adjointness are defined only for square operators; hence, there are no rectangular projectors, no rectangular positive operators, no rectangular unitary operators;

- $A \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$ is a *partial isometry* if both $A^\dagger A$ and $AA^\dagger$ are orthogonal projectors in $\mathscr{L}(\mathcal{H}_1)$ and $\mathscr{L}(\mathcal{H}_2)$, respectively; $A \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$ is an *isometry* (or an *isometric embedding*) if $A^\dagger A = I$, where I denotes the identity map in $\mathscr{L}(\mathcal{H}_1)$.

While the Spectral Theorem is about square operators only, the Singular-value Decomposition (SVD) and the Polar Decomposition can be given also for rectangular operators.

Theorem 1.13 (Polar decomposition for rectangular operators)**.** *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces, and let d_1 and d_2 be their dimensions. For any $A \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$, the following statements hold:*

1. *if $d_1 \leq d_2$, there exists an isometry $W \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$ such that*

$$A = WQ, \tag{1.17}$$

   where $Q = (A^\dagger A)^{1/2}$ is a positive semidefinite operator in $\mathscr{L}(\mathcal{H}_1)$;

2. *if $d_1 \geq d_2$, there exists an isometry $V \in \mathscr{L}(\mathcal{H}_2, \mathcal{H}_1)$ such that*

$$A = PV^\dagger, \tag{1.18}$$

   where $P = (AA^\dagger)^{1/2}$ is a positive semidefinite operator in $\mathscr{L}(\mathcal{H}_2)$.

Corollary 1.2. *Let $A \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$. If $d_1 \leq d_2$, there exists an isometry $W \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$ such that $AA^\dagger = WA^\dagger AW^\dagger$. If $d_1 \geq d_2$, there exists an isometry $V \in \mathscr{L}(\mathcal{H}_2, \mathcal{H}_1)$ such that $A^\dagger A = VAA^\dagger V^\dagger$. In other words, $AA^\dagger$ and $A^\dagger A$ have the same positive eigenvalues.*

Theorem 1.14 (SVD for rectangular operators)**.** *For any $A \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$, there exist unitary operators $V \in \mathscr{L}(\mathcal{H}_2)$ and $W \in \mathscr{L}(\mathcal{H}_1)$ such that*

$$A = V\Sigma W, \tag{1.19}$$

where $\Sigma \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$ is the $d_2 \times d_1$ rectangular matrix defined as $\Sigma = \sum_{i=1}^{d_2} \sum_{j=1}^{d_1} s_{ij} |f_i\rangle\langle e_j|$, with $s_{ij} = 0$ for $i \neq j$, and $s_{11} \geq s_{22} \geq \cdots \geq s_{qq} \geq 0$, where $q = \min\{d_1, d_2\}$. The positive numbers s_{ii} are called the singular values *of A.*

☞ For any $A \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_2)$, the singular values s_{ii} are equal to the eigenvalues of $(AA^\dagger)^{1/2}$, repeated according to their degeneracy and listed in decreasing order.

2 Mathematical description of quantum systems

2.1 Physical states and physical properties of quantum systems

The starting point of any empirical theory (i.e. a theory based on experiments and observations) is to provide rules to describe the *states* of a physical system. As it is usually found on textbooks, the first postulate is as follows:

> **Postulate 1** (Representations of quantum states)**.** *Any quantum system Q is associated with a unique Hilbert space $\mathcal{H}_Q$. Any state of Q is represented by a normalized vector in $\mathcal{H}_Q$, i.e. $\psi \in \mathcal{H}_Q$ such that $\langle\psi|\psi\rangle = 1$.*

Postulate 1 does not specify anything about the correspondence between states and vectors: we are not told, for example, whether such a correspondence is one-to-one or not. In order to clarify this important point, we first need to understand what does it mean that *two states are different*. Intuitively, two states are (defined to be) different if there exists at least one physical property that distinguish between the two. The problem is, however, that we do not know yet what a *physical property* is in quantum theory.

The next postulate tells us how physical properties are represented in quantum theory. We recall here the spectral decomposition (1.13) of self-adjoint operators $A = \sum_\mu \mu \Pi^A(\mu)$. We further define, for any interval of the real line $\Delta \subseteq \mathbb{R}$, the projector $\Pi^A(\Delta) := \sum_{\mu\in\Delta} \Pi^A(\mu)$. Since the eigenvalues of every observable $A \in \mathscr{L}(\mathcal{H})$ are finite,

$$\Pi^A(\mathbb{R}) = \sum_\mu \Pi^A(\mu) = \mathbb{1}, \tag{2.1}$$

for every observable A.

> **Postulate 2** (Representation of physically measurable quantities)**.** *Any physical property (or dynamical variable) of Q is represented in one-to-one correspondence by a self-adjoint operator $A \in \mathscr{L}(\mathcal{H}_Q)$. Such operators are called the* observables *of Q. For any physical property A of Q, the eigenvalues of A are the only possible values that a measurement of the physical property A on Q can give.*

> **Postulate 3** (Born statistical formula)**.** *Any observable A can be measured in any state. In the case in which the state is represented by the vector $\psi \in \mathcal{H}_Q$, the measurement of A returns a value in an interval $\Delta \subseteq \mathbb{R}$ with probability $\langle\psi|\Pi^A(\Delta)|\psi\rangle$. We denote this probability by $\Pr\{A \in \Delta \| \psi\}$.*

We can now refine Postulate 1 as follows: consider two vectors of $\mathcal{H}_Q$, ψ and ϕ, such that $\phi = z\psi$, where z is a complex phase (i.e. $z^*z = 1$). The Born statistical formula tells us that

$$\begin{aligned}\Pr\{A \in \Delta\|\phi\} =& \langle\phi|\Pi^A(\Delta)|\phi\rangle = \langle z\psi|\Pi^A(\Delta)|z\psi\rangle = z^*z\langle\psi|\Pi^A(\Delta)|\psi\rangle = \langle\psi|\Pi^A(\Delta)|\psi\rangle \\ =& \Pr\{A \in \Delta\|\psi\},\end{aligned}$$

for all intervals Δ and all observables A. In other words, the two vectors ψ and ϕ will produce exactly the same outcome statistics in any possible measurement, that is to say, both ψ and ϕ represent the *same physical state*. We have therefore the following refinement of Postulate 1:

> **Postulate 1'** (Representations of quantum states)**.** *Any quantum system Q is associated with a unique Hilbert space $\mathcal{H}_Q$. Any state of Q is represented, in one-to-one correspondence, by a projector $|\psi\rangle\langle\psi|$, $\psi \in \mathcal{H}_Q$ with $\langle\psi|\psi\rangle = 1$.*

By representing states by projectors, rather than vectors, we automatically get rid of the unphysical overall phase, i.e. $|\psi\rangle\langle\psi| = |z\psi\rangle\langle z\psi|$, for any $z \in \mathbb{C}$ such that $z^*z = 1$. The Born statistical formula too gets updated as follows:

> **Postulate 3'** (Born statistical formula)**.** *Any observable A can be measured in any state. In the case in which the state is represented by the projector $|\psi\rangle\langle\psi|$, the measurement of A returns a value in an interval $\Delta \subseteq \mathbb{R}$ with probability* $\mathrm{Tr}[\Pi^A(\Delta)\ |\psi\rangle\langle\psi|]$. *We denote this probability by* $\Pr\{A \in \Delta \| \psi\}$.

Example 2.1. Let us consider a two-dimensional quantum system Q, i.e. a quantum system with Hilbert space $\mathcal{H} \cong \mathbb{C}^2$. Such a system is the simplest quantum system and it is called a *quantum bit*, or, in short, *qubit*. Let the quantum system Q be in the state corresponding to the vector $\psi = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, i.e.

$$|\psi\rangle\langle\psi| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Suppose that we want to measure the observable

$$A = \begin{pmatrix} a_0 & 0 \\ 0 & a_1 \end{pmatrix},$$

with $a_0, a_1 \in \mathbb{R}$, $a_0 \neq a_1$. Postulate 3 tells us that $\Pr\{A = a_0 \| \psi\} = 1$, while $\Pr\{A = a_1 \| \psi\} = 0$. This means that, in this case, a measurement of the physical quantity A will return the value a_0 with certainty. In this case, hence, Quantum Theory can *predict* the result of the measurement.

Example 2.2. Let us consider the same observable A as in the example above, but the state of the quantum system is now described by $\psi = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, i.e.

$$|\psi\rangle\langle\psi| = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

In this case, Postulate 3' tells us that $\Pr\{A = a_0 \| \psi\} = \Pr\{A = a_1 \| \psi\} = 1/2$. This means that, in this case, a measurement of the physical quantity A will return either the value a_0 or the value a_1, each with 50% probability. In this case, hence, the possible value of the measurement is *completely random*, even if we know the state of the quantum system in advance. In other words, quantum theory *cannot predict the result of the measurement*, in this case.

☞ From the previous examples, we learn that if we know that the state of a quantum system Q, immediately before the measurement of an observable A, is an eigenvector $|\mu\rangle\langle\mu|$ of A, the measurement of A on Q will return outcome μ with probability one. However, in general, it is impossible to predict the outcome of a measurement, even if we know the observable A and the state of the system ρ, exactly. *In quantum theory, the outcome of a measurement is intrinsically random.*

Remark 2.1. In this course we consider only finite dimensional Hilbert spaces. This means that we consider only finite dimensional quantum systems. What does it mean that a quantum system is finite dimensional? From Postulates 2 and 3 above, a quantum system Q is finite dimensional if and only if any physically measurable quantity of Q can assume only a finite number of possible values. □

2.2 Distinguishing states of quantum systems

Example 2.3 (Distinguishing quantum states)**.** Suppose that a quantum system Q can randomly be in a quantum state chosen between two possible ones, $|\psi_0\rangle\langle\psi_0|$ and $|\psi_1\rangle\langle\psi_1|$. Is there an observable on Q that we can measure so to determine in which state the system is? The problem amounts to finding a measurable physical quantity (i.e. an observable) that, with probability one, assumes values in disjoint intervals for $|\psi_0\rangle\langle\psi_0|$ and $|\psi_1\rangle\langle\psi_1|$.

First case: the two states are orthogonal, i.e. $\langle\psi_0|\psi_1\rangle = 0$. Then, it is easy to see that the observable

$$A := \alpha|\psi_0\rangle\langle\psi_0| + \beta|\psi_1\rangle\langle\psi_1| \tag{2.2}$$

satisfies our requirement, whenever $\alpha \neq \beta$. Hence, *if the states are orthogonal, there always exists a measurement distinguishing them.*

Second case: $\langle\psi_0|\psi_1\rangle = c_0 \neq 0$. Then, we can linearly decompose $|\psi_1\rangle$ as $|\psi_1\rangle = c_0|\psi_0\rangle +$ (other terms). The state $|\psi_1\rangle\langle\psi_1|$ is then equal to

$$|\psi_1\rangle\langle\psi_1| = |c_0|^2|\psi_0\rangle\langle\psi_0| + (\text{other terms}). \tag{2.3}$$

Now, an observable A assumes a value in $\Delta \subseteq \mathbb{R}$ on $|\psi_0\rangle\langle\psi_0|$ with probability one if and only if $\mathrm{Tr}[\Pi^A(\Delta)\,|\psi_0\rangle\langle\psi_0|] = 1$. Due to the decomposition (2.3) and the linearity of the trace, we have that

$$\begin{aligned}\mathrm{Tr}\left[\Pi^A(\Delta)\,|\psi_1\rangle\langle\psi_1|\right] =& |c_0|^2\,\mathrm{Tr}\left[\Pi^A(\Delta)\,|\psi_0\rangle\langle\psi_0|\right] + (\text{other terms})\\ =& |c_0|^2 + (\text{other terms})\\ \geq& |c_0|^2\\ >& 0.\end{aligned} \tag{2.4}$$

The above calculation shows that there is a non-zero probability that A assumes a value in the same interval Δ also on $|\psi_1\rangle\langle\psi_1|$. This means that there is no observable that assumes values in disjoint intervals for $|\psi_0\rangle\langle\psi_0|$ and $|\psi_1\rangle\langle\psi_1|$, with probability one. Hence, *if the states are not orthogonal, there always exists a non-zero probability of mis-identification.* □

It is not difficult to extend the same arguments to an arbitrary (finite) number of mixed states, so that it is possible to prove the following:

Theorem 2.1 (Distinguishable states)**.** *A family of states $(\psi_i)_i$ is* perfectly distinguishable *if and only if $\langle\psi_i|\psi_j\rangle = 0$, for any $i \neq j$, i.e. they are all pairwise orthogonal.*

2.3 Random samples of quantum systems

Let us now imagine a source of quantum particles, such that each particle can be in a state $|\psi_j\rangle\langle\psi_j|$, chosen among a family of possible vectors $\{\psi_1, \cdots, \psi_n\} \subset \mathcal{H}$, with probability p_j. We say that each particle is a *random sample* from the ensemble $(\{p_j\}, \{\psi_j\})$.

Theorem 2.2 (Random samples)**.** *Any random sample from the ensemble $(\{p_j\}, \{\psi_j\})$ is associated with the operator $\rho = \sum_j p_j|\psi_j\rangle\langle\psi_j|$. In other words, there is no way to distinguish among different ensembles having the same average state.*

Proof. Any observable A, measured on a random sample of $(\{p_j\}, \{\psi_j\})$, takes a value in the interval Δ with probability

$$\Pr\{A \in \Delta \| (\{p_j\}, \{\psi_j\})\} = \sum_j p_j\,\mathrm{Tr}[\Pi^A(\Delta)\,|\psi_j\rangle\langle\psi_j|]. \tag{2.5}$$

By linearity of the trace, $\Pr\{A \in \Delta \| (\{p_j\}, \{\psi_j\})\} = \mathrm{Tr}[\Pi^A(\Delta)\,\rho]$, with $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$. Since this holds for *any* observable, the state of a random sample is correctly described by ρ. ∎

Since any operator ρ of the form $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$ is positive (i.e. $\langle v|\rho|v\rangle \geq 0$ for all $v \in \mathcal{H}$) and has trace equal to one (i.e. $\mathrm{Tr}[\rho] = 1$), we update Postulate 1 one final time as follows:

> **Postulate 1"** (Representations of quantum states)**.** *Any quantum system Q is associated with a unique Hilbert space $\mathcal{H}_Q$. Any state of Q is represented, in one-to-one correspondence, by a positive operator $\rho \in \mathscr{L}(\mathcal{H}_Q)$, with $\mathrm{Tr}\,\rho = 1$. Such operators are equivalently called* states, density operators, *or* density matrices *of Q.*

Definition 2.1. A state $\rho \in \mathscr{L}(\mathcal{H})$ is called *pure* if and only if the rank of ρ is equal to one. This is equivalent to say that there exists a vector $\psi \in \mathcal{H}$ with $\|\psi\| = 1$ such that $\rho = |\psi\rangle\langle\psi|$. If ρ is not a pure state, then it is called *mixed*.

If the state we assign to a quantum system is pure, it means that *we have perfect knowledge* about the system. Assigning a mixed state always implies that *our knowledge of the system is incomplete*.

Proposition 2.1. *A given density matrix ρ corresponds to a pure state if and only if $\rho^2 = \rho$, or equivalently,* $\mathrm{Tr}[\rho^2] = 1$.

We can now formulate also Postulate 3 in its more general form:

> **Postulate 3"** (Born statistical formula)**.** *Any observable A can be measured in any state. In the case in which the state is represented by the density operator ρ, the measurement of A returns a value in an interval $\Delta \subseteq \mathbb{R}$ with probability* $\mathrm{Tr}[\Pi^A(\Delta)\ \rho]$. *We denote this probability by* $\Pr\{A \in \Delta \| \rho\}$.

☞ Given an observable A and a state ρ of a quantum system Q, while it is (in general) impossible to compute the value that A assumes on Q, it is instead possible to compute the *expected* value of A on Q. Such an "average" value, usually denoted by $\langle A\rangle_\rho$ and called the *expectation value* of A on ρ, is defined as

$$\langle A\rangle_\rho := \mathrm{Tr}[A\,\rho]. \tag{2.6}$$

By expanding the observable as $A = \sum_\mu \mu \Pi^A(\mu)$, we have that $\langle A\rangle_\rho = \sum_\mu \mu \Pr\{A = \mu \| \rho\}$. This justifies the name "expectation value" given to $\langle A\rangle_\rho$.

Example 2.4. Suppose, for example, that a source emits quantum particles randomly. The only thing we know is that, half of the times (i.e. with probability 1/2), the particle's state is

$$\rho_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

half of the times is

$$\rho_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

According to Theorem 2.2, Quantum Theory tells us that the state of *every* particle coming out from the source is correctly described by the state

$$\bar\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}.$$

Example 2.5. Again, as in the previous example, we have a source emitting quantum particles randomly. This time, however, the source emits with probability 1/2 a particle in the state

$$\sigma_1 = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix},$$

with probability 1/2 a particle in the state

$$\sigma_2 = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}.$$

Even if the source is different from the source described in Example 2.4, the state of every particle coming out from this source, according to Quantum Theory, is still given by

$$\bar{\sigma} = \frac{1}{2}\sigma_1 + \frac{1}{2}\sigma_2 = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} = \bar{\rho}.$$

Exercise 2.1. Find probabilities $p_1, p_2, p_3 = 1 - p_1 - p_2$ and three quantum states τ_1, τ_2, τ_3 such that $p_1\tau_1 + p_2\tau_2 + p_3\tau_3 = \bar{\rho}$, where $\bar{\rho}$ is the same average state of the preceding examples.

☞ According to Postulate 1, the state of a quantum system is described by a density matrix ρ, i.e. a positive matrix with $\operatorname{Tr}\rho = 1$. Since ρ is positive, it is also self-adjoint, so that we can apply the Spectral Theorem 1.7 and write ρ in diagonal form $\rho = \sum_j \lambda_j |\psi_j\rangle\langle\psi_j|$, where λ_j are the eigenvalues and $|\psi_j\rangle$ are the corresponding orthonormal eigenvectors. Since ρ is positive, $\lambda_j \geq 0$, for all j. Since $\operatorname{Tr}\rho = 1$, $\sum_j \lambda_j = 1$. These two conditions together mean that the numbers λ_j form a *probability distribution*. Moreover, the matrices $|\psi_j\rangle\langle\psi_j|$ are themselves density matrices, for all j. This means that Quantum Theory allows us to interpret *any* mixed quantum state ρ as a random sample taken from a source emitting quantum particles in the state $|\psi_j\rangle\langle\psi_j|$ with probability λ_j. A very important, very subtle point to stress now is that, even if we can interpret any mixed state as a random sample, this does not mean that every quantum system in a mixed state was *actually* produced as a random sampling. This feature of Quantum Theory is a consequence of the phenomenon known as "quantum entanglement" (see below).

Question 2.1 (Very difficult)**.** Given a mixed state $\rho \in \mathscr{L}(\mathcal{H})$, how to characterize all the ensembles of pure states $(\{p_j\}, \{\psi_j\})$ such that $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$?

2.4 Dynamics, composite quantum systems, quantum entanglement

Up to now, we learnt only how Quantum Theory describes the states and the physical properties of quantum systems. However, we still know nothing about how a quantum system evolve in time. This is the topic of the next postulate:

> **Postulate 4** (Schrödinger Equation)**.** *If a quantum system Q is isolated during the time interval $[t, t']$, $t \leq t'$, there exists a unitary operator $U \in \mathscr{L}(\mathcal{H}_Q)$, called the* time evolution operator*, such that, if Q was in state ρ at time t, it will be in state $\rho' = U\rho U^\dagger$ at time t'.*

Remark 2.2. The name "Schrödinger Equation" given to Postulate 4 is not completely correct. In fact, Postulate 4 describes how the state of a quantum system changes from an initial time t to a final time t': it deals hence with *discrete* time evolution. The equation originally proposed by Schrödinger, however, deals with *continuous* time evolutions, and gives a rule to describe how the state of a quantum system changes from an initial time t to a final time $t + \mathrm{d}t$, which is infinitesimally close to t.

Remark 2.3. Postulate 4 tells us that any closed evolution is described by a unitary operator. However, we usually assume that also the opposite is true: that to any unitary operator there exists an arrangement that is able to implement it as a closed evolution. In other words, in quantum computation one assumes that *any unitary operator is a legitimate gate* that can be used in a logical circuit. See the following Example.

Example 2.6 (NOT and $\sqrt{\text{NOT}}$ gates)**.** As in classical computation, the NOT gate, i.e. $x \mapsto x \oplus 1$ is important also in quantum computation. It is realized by the unitary matrix $U_{\text{NOT}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. But in quantum computation we also have the square root of NOT, i.e. $U_{\sqrt{\text{NOT}}} = \frac{1}{2}\begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix}$, which has no analogue in classical computation.

The last postulate of Quantum Theory that tells us how to "put together" (or "combine") quantum systems into *composite* ones. This is very important: it is very often the case, in fact, that a quantum system is not *elementary*, but composed of "smaller" building blocks (like a molecule, for example, which is composed by atoms, which are in turn made of protons, neutrons, and electrons, which are in turn.....). The following postulate, hence, provides us the tools to "reverse engineer" such complex quantum systems, as compositions of interacting parts.

☞ However, before stating the postulate, we need to introduce the notion of "tensor product" between vector spaces. Given the complex vector spaces $\mathbb{C}^m$ and $\mathbb{C}^n$, let

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix} \in \mathbb{C}^m \quad \text{and} \quad w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{C}^n.$$

The tensor product $v \otimes w$ (read "v tensor w") is defined as the $1 \times mn$ matrix

$$v \otimes w := \begin{pmatrix} v_1 \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \\ v_2 \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \\ \vdots \\ v_m \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \end{pmatrix} \equiv \begin{pmatrix} v_1 w_1 \\ v_1 w_2 \\ \vdots \\ v_1 w_n \\ v_2 w_1 \\ v_2 w_2 \\ \vdots \\ v_m w_n \end{pmatrix}.$$

☞ The following relations hold for any $v, v_1, v_2 \in \mathbb{C}^m$, any $w, w_1, w_2 \in \mathbb{C}^n$, and any $c, c_1, c_2 \in \mathbb{C}$:

$$\begin{aligned} c(v \otimes w) &= cv \otimes w = v \otimes cw \\ (c_1 v_1 + c_2 v_2) \otimes w &= c_1 v_1 \otimes w + c_2 v_2 \otimes w \\ v \otimes (c_1 w_1 + c_2 w_2) &= c_1 v \otimes w_1 + c_2 v \otimes w_2. \end{aligned}$$

The tensor product space $\mathbb{C}^m \otimes \mathbb{C}^n$ is defined as the set containing all linear combinations of tensor product vectors of the form $v \otimes w$, for any $v \in \mathbb{C}^m$ and any $w \in \mathbb{C}^n$. It is easy to see that $\mathbb{C}^m \otimes \mathbb{C}^n \equiv \mathbb{C}^{mn}$.

☞ What form have the linear operators acting on a tensor product vector space? Let $A \in \mathscr{L}(\mathbb{C}^m)$ and $B \in \mathscr{L}(\mathbb{C}^n)$ be written as

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}.$$

Then, the tensor product operator $A \otimes B$ is defined as

$$\begin{aligned} A \otimes B := & \begin{pmatrix} a_{11} \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} & \cdots & a_{1m} \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} \\ \vdots & & \ddots & \vdots \\ a_{m1} \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} & \cdots & a_{mm} \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} \end{pmatrix} \\ \equiv & \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & \cdots & a_{11}b_{1n} & a_{12}b_{11} & a_{12}b_{12} & \cdots & a_{1m}b_{1n} \\ a_{11}b_{21} & a_{11}b_{22} & \cdots & a_{11}b_{2n} & a_{12}b_{21} & a_{12}b_{22} & \cdots & a_{1m}b_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{n1} & a_{m1}b_{n2} & \cdots & a_{m1}b_{nn} & a_{m2}b_{n1} & a_{m2}b_{n2} & \cdots & a_{mm}b_{nn} \end{pmatrix}. \end{aligned}$$

☞ The following relations hold for any $A, A_1, A_2 \in \mathscr{L}(\mathbb{C}^m)$, any $B, B_1, B_2 \in \mathscr{L}(\mathbb{C}^n)$, and any $c, c_1, c_2 \in \mathbb{C}$:

$$\begin{aligned} &c(A \otimes B) = cA \otimes B = A \otimes cB \\ &(c_1 A_1 + c_2 A_2) \otimes B = c_1 A_1 \otimes B + c_2 A_2 \otimes B \\ &A \otimes (c_1 B_1 + c_2 B_2) = c_1 A \otimes B_1 + c_2 A \otimes B_2. \end{aligned} \tag{2.7}$$

The set of linear operators $\mathscr{L}(\mathbb{C}^m \otimes \mathbb{C}^n)$ is defined as the set containing all linear combinations of tensor product operators of the form $A \otimes B$, for any $A \in \mathscr{L}(\mathbb{C}^m)$ and any $B \in \mathscr{L}(\mathbb{C}^n)$. Since, as we noticed before, $\mathbb{C}^m \otimes \mathbb{C}^n \equiv \mathbb{C}^{mn}$, $\mathscr{L}(\mathbb{C}^m \otimes \mathbb{C}^n) \equiv \mathscr{L}(\mathbb{C}^{mn})$. Moreover, by treating $\mathscr{L}(\mathbb{C}^m)$ and $\mathscr{L}(\mathbb{C}^n)$ as complex vector spaces by themselves, it is easy to see that $\mathscr{L}(\mathbb{C}^m \otimes \mathbb{C}^n) = \mathscr{L}(\mathbb{C}^{mn}) = \mathscr{L}(\mathbb{C}^m) \otimes \mathscr{L}(\mathbb{C}^n)$.

We are now ready to state the last postulate:

Postulate 5 (Composition of quantum systems)**.** *Given two quantum systems Q and R associated with Hilbert spaces $\mathcal{H}_Q$ and $\mathcal{H}_R$, the composite bipartite system QR is associated with the* tensor product *$\mathcal{H}_{QR} = \mathcal{H}_Q \otimes \mathcal{H}_R$. The states of the composite quantum system QR are in one-to-one correspondence with density matrices in $\mathscr{L}(\mathcal{H}_Q \otimes \mathcal{H}_R)$. The physical properties of the composite quantum system QR are in one-to-one correspondence with self-adjoint operators in $\mathscr{L}(\mathcal{H}_Q \otimes \mathcal{H}_R)$. Any physical property of system Q, represented by the self-adjoint operator $A \in \mathscr{L}(\mathcal{H}_Q)$, is identified with the observable $A \otimes \mathbb{1}_R \in \mathscr{L}(\mathcal{H}_Q \otimes \mathcal{H}_R)$ of the composite system QR. Analogously, any physical property of system R, represented by the self-adjoint operator $B \in \mathscr{L}(\mathcal{H}_R)$, is identified with the observable $\mathbb{1}_Q \otimes B \in \mathscr{L}(\mathcal{H}_Q \otimes \mathcal{H}_R)$ of the composite system QR.*

☞ Let us suppose that $\mathcal{H}_Q$ is an m-dimensional Hilbert space, and that $\mathcal{H}_R$ is an n-dimensional Hilbert space. By using the correspondence described in Eq. (1.2), we have that $\mathcal{H}_Q \cong \mathbb{C}^m$ and $\mathcal{H}_R \cong \mathbb{C}^n$. Then, $\mathcal{H}_Q \otimes \mathcal{H}_R \cong \mathbb{C}^m \otimes \mathbb{C}^n$.

☞ By convention, the Dirac ket's obtained from tensor product elements of $\mathcal{H}_Q \otimes \mathcal{H}_R$, for example $\psi \otimes \phi$, are equivalently written as $|\psi \otimes \phi\rangle$, or $|\psi\rangle \otimes |\phi\rangle$, or even as $|\psi\rangle|\phi\rangle$. The corresponding bra's can be written as $\langle\psi \otimes \phi|$, $\langle\psi| \otimes \langle\phi|$, or $\langle\psi|\langle\phi|$.

Exercise 2.2. Consider, for example, the case of $\mathbb{C}^2 \otimes \mathbb{C}^2 \cong \mathbb{C}^4$. Now, consider the normalized vector

$$y = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \in \mathbb{C}^2 \otimes \mathbb{C}^2.$$

Find vectors $v, w \in \mathbb{C}^2$ such that $y = v \otimes w$. □

Exercise 2.3. Are all vectors in a tensor product space, for example $\mathbb{C}^m \otimes \mathbb{C}^n$, in tensor product form, for example $v \otimes w$? Let's consider again $\mathbb{C}^2 \otimes \mathbb{C}^2 \cong \mathbb{C}^4$. Given the normalized vector

$$z = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \in \mathbb{C}^2 \otimes \mathbb{C}^2,$$

find vectors $v, w \in \mathbb{C}^2$ such that $z = v \otimes w$. $\square$

Postulate 5 tells not only how to combine quantum systems, but also how to "split" them. It gives us a rule that answer the following question: "If I know the density matrix describing the joint state of a composite quantum system, how to derive the density matrix correctly describing the state of each component?"

☞ Suppose, for example, that the state of a composite system QR is described by the density matrix $\rho_{QR} \in \mathbb{M}(\mathbb{C}^m \otimes \mathbb{C}^n)$. From Postulate 5, we know that any physical property (i.e. self-adjoint operator) $A \in \mathscr{L}(\mathbb{C}^m)$ of Q is associated with the physical property (i.e. self-adjoint operator) $A' := A \otimes \mathbb{1}_n \in \mathscr{L}(\mathbb{C}^m \otimes \mathbb{C}^n)$ of QR. What is the relation between A of Q and A' fo QR? Suppose that the spectral decomposition of A is $A = \sum_\mu \mu \Pi^A(\mu)$. By equation (2.7), $A' := A \otimes \mathbb{1}_n = \left[\sum_\mu \mu \Pi^A(\mu)\right] \otimes \mathbb{1}_n = \sum_\mu \mu \left[\Pi^A(\mu) \otimes \mathbb{1}_n\right]$. This implies that A' *has the same set of eigenvalues of* A, and that *the corresponding spectral projectors are* $E^{A'}(\mu) = \Pi^A(\mu) \otimes \mathbb{1}_n$.

Postulate 5 states the following: measuring the observable A' on the composite system QR is *equivalent* to measuring the observable A on Q alone. In other words, the probability that the observable A takes a value μ in Q, is equal to the probability that the observable $A' = A \otimes \mathbb{1}_n$ takes the same value μ in QR. Therefore, given that the composite system QR is in state ρ_{QR}, Postulate 5 *implicitly determines* the correct state ρ_Q of the system Q by the relation:

$$\mathrm{Tr}\left[\Pi^A(\mu)\,\rho_Q\right] \equiv \Pr\{A = \mu \| \rho_Q\} \overset{\text{"Post. 5"}}{=} \Pr\{A' = \mu \| \rho_{QR}\} \equiv \mathrm{Tr}\left[E^{A'}(\mu)\,\rho_{QR}\right]. \tag{2.8}$$

The above equation, that must hold for all self-adjoint operators $A \in \mathscr{L}(\mathbb{C}^m)$ and for all eigenvalues μ of each A, determines a set of linear equations that uniquely identify the density matrix ρ_Q, representing the state of Q.

We now introduce the following definition:

Definition 2.2 (Partial trace)**.** Given a tensor product space $\mathbb{C}^m \otimes \mathbb{C}^n$, the operation *partial trace* over $\mathbb{C}^n$

$$\mathrm{Tr}_{\mathbb{C}^n} : \mathbb{M}(\mathbb{C}^m \otimes \mathbb{C}^n) \to \mathbb{M}(\mathbb{C}^m), \tag{2.9}$$

is defined on tensor product operators of the form $A \otimes B$ by the relation

$$\mathrm{Tr}_{\mathbb{C}^n}[A \otimes B] = A\,\mathrm{Tr}[B]. \tag{2.10}$$

In terms of outer products, the above definition becomes $\mathrm{Tr}_A[|u_A\rangle\langle v_A| \otimes |w_B\rangle\langle z_B|] = \langle v_A|u_A\rangle |w_B\rangle\langle z_B|$. Definition (2.10) is then extended, by linearity, to all operators in $\mathscr{L}(\mathbb{C}^m \otimes \mathbb{C}^n)$. Of course, the following holds:

$$\mathrm{Tr}[Z_{AB}] = \mathrm{Tr}_A\left\{\mathrm{Tr}_B[Z_{AB}]\right\} = \mathrm{Tr}_B\left\{\mathrm{Tr}_A[Z_{AB}]\right\}. \qquad \square$$

Example 2.7 (Partial trace of a generic matrix)**.** Let us consider a linear operator $C \in \mathscr{L}(\mathbb{C}^m \otimes \mathbb{C}^n)$, which we write as an $(m \times n) \times (m \times n)$ matrix

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1,mn} \\ c_{21} & c_{22} & \cdots & c_{2,mn} \\ \vdots & \vdots & \ddots & \vdots \\ c_{mn,1} & c_{mn,2} & \cdots & c_{mn,mn} \end{pmatrix}. \tag{2.11}$$

We can divide the matrix C into an $m \times m$ matrix of $n \times n$ matrices, i.e.

$$C = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1,m} \\ C_{21} & C_{22} & \cdots & C_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ C_{m,1} & C_{m,2} & \cdots & C_{m,m} \end{pmatrix}, \tag{2.12}$$

where, for example, $C_{11} = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix}$. Then, according to our definition of partial trace,

$$\mathrm{Tr}_{\mathbb{C}^m}[C] = \sum_{i=1}^{m} C_{ii} \in \mathscr{L}(\mathbb{C}^n), \tag{2.13}$$

and

$$\mathrm{Tr}_{\mathbb{C}^n}[C] = \begin{pmatrix} \mathrm{Tr}[C_{11}] & \mathrm{Tr}[C_{12}] & \cdots & \mathrm{Tr}[C_{1,m}] \\ \mathrm{Tr}[C_{21}] & \mathrm{Tr}[C_{22}] & \cdots & \mathrm{Tr}[C_{2,m}] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Tr}[C_{m,1}] & \mathrm{Tr}[C_{m,2}] & \cdots & \mathrm{Tr}[C_{m,m}] \end{pmatrix} \in \mathscr{L}(\mathbb{C}^m). \ \square \tag{2.14}$$

☞ One can prove that Eq. (2.8) holds for any observable $A \in \mathscr{L}(\mathbb{C}^m)$ if and only if $\rho_Q = \mathrm{Tr}_R[\rho_{QR}] \equiv \mathrm{Tr}_{\mathbb{C}^n}[\rho_{QR}]$.

The following proposition summarizes a property of partial trace, which turns out to be very useful when performing calculations:

Proposition 2.2. *The operation of partial trace satisfies the following property:*

$$\mathrm{Tr}_B[(X_A \otimes \mathbb{1}_B)\ Z_{AB}\ (Y_A \otimes \mathbb{1}_B)] = X_A\, \mathrm{Tr}_B[Z_{AB}] Y_A. \tag{2.15}$$

Going back to Postulate 5: we are now able to reformulate it in the following way:

> **Postulate 5'.** *Given two quantum systems Q and R associated with Hilbert spaces $\mathcal{H}_Q$ and $\mathcal{H}_R$, the composite bipartite system QR is associated with the* tensor product *$\mathcal{H}_{QR} = \mathcal{H}_Q \otimes \mathcal{H}_R$. The states of the composite quantum system QR are in one-to-one correspondence with density matrices in $\mathscr{L}(\mathcal{H}_Q \otimes \mathcal{H}_R)$. The physical properties of the composite quantum system QR are in one-to-one correspondence with self-adjoint operators in $\mathscr{L}(\mathcal{H}_Q \otimes \mathcal{H}_R)$. Moreover, given that the composite system is in state ρ_{QR}, the state of system Q is given by $\rho_Q := \mathrm{Tr}_R[\rho_{QR}]$. Analogously, the state of system R is given by $\rho_R := \mathrm{Tr}_Q[\rho_{QR}]$.*

Remark 2.4 (Discarding subsystems). The operation partial trace describe that we are *discarding* part of a composite system. We hence say "*Perform the partial trace over subsystem R*" and mean "*Discard subsystem R*".

☞ We can now easily prove why the vector considered in Exercise 2.3 *cannot* be written as a tensor product $v \otimes w$. Let us first introduce the vectors $|e_0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|e_1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Then,

$$|z\rangle = \frac{1}{\sqrt{2}}|e_0\rangle \otimes |e_0\rangle + \frac{1}{\sqrt{2}}|e_1\rangle \otimes |e_1\rangle. \tag{2.16}$$

Let us compute the matrix $|z\rangle\langle z|$ as follows:

$$\begin{aligned}
&|z\rangle\langle z| \\
=&\frac{1}{2}\left(|e_0\otimes e_0\rangle + |e_1\otimes e_1\rangle\right)\left(\langle e_0\otimes e_0| + \langle e_1\otimes e_1|\right) \\
=&\frac{1}{2}\left(|e_0\rangle\langle e_0|\otimes|e_0\rangle\langle e_0| + |e_0\rangle\langle e_1|\otimes|e_0\rangle\langle e_1| + |e_1\rangle\langle e_0|\otimes|e_1\rangle\langle e_0| + |e_1\rangle\langle e_1|\otimes|e_1\rangle\langle e_1|\right) \\
=&\frac{1}{2}\begin{pmatrix}1&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&1\end{pmatrix}.
\end{aligned} \tag{2.17}$$

The matrix $|z\rangle\langle z| \in \mathscr{L}(\mathbb{C}^2\otimes\mathbb{C}^2)$ is, in fact, a density matrix, i.e. a positive operator and with unit trace. This means that $|z\rangle\langle z|$ corresponds to a possible state a composite quantum system, say QR, obtained as a combination of Q (with $\mathcal{H}_Q \cong \mathbb{C}^2$) and R (with $\mathcal{H}_R \cong \mathbb{C}^2$). We notice, in particular, that the matrix $|z\rangle\langle z|$ describes a *pure* state, since $|z\rangle\langle z|^2 = |z\rangle\langle z|$ (see Proposition 2.1).
What are the reduced states obtained from $\rho_{QR} = |z\rangle\langle z|$? By the definition of partial trace, the reduced state of Q is given by

$$\mathrm{Tr}_R[|z\rangle\langle z|] = \frac{1}{2}\left(|e_0\rangle\langle e_0| + |e_1\rangle\langle e_1|\right) = \frac{1}{2}\mathbb{1}_2. \tag{2.18}$$

The reduced state of $|z\rangle\langle z|_{QR}$ on Q is described by a *mixed* state! (Just check that $(\mathbb{1}/2)^2 = \mathbb{1}/4$.) This gives a proof that the vector $|z\rangle$ cannot be written as a tensor product of two vectors. In fact, the reduced state of a tensor product state, like $|\psi\otimes\phi\rangle\langle\psi\otimes\phi|$, would be pure as well. However, for $|z\rangle\langle z|$, the reduced state is mixed.

States like that of Exercise 2.3 are called *entangled*:

Definition 2.3. Let $|\Psi_{AB}\rangle\langle\Psi_{AB}|$ be a pure state in $\mathcal{H}_A\otimes\mathcal{H}_B$. Then, $|\Psi_{AB}\rangle\langle\Psi_{AB}|$ is called *entangled* if and only if $|\Psi_{AB}\rangle = |v_A\rangle\otimes|w_B\rangle$, for some normalized vectors $v_A \in \mathcal{H}_A$ and $w_B \in \mathcal{H}_B$. Otherwise, $|\Psi_{AB}\rangle\langle\Psi_{AB}|$ is called *separable*.

Proposition 2.3. *A pure bipartite state $|\Psi_{AB}\rangle\langle\Psi_{AB}|$ is separable if and only if, for $\rho_A = \mathrm{Tr}_B[|\Psi_{AB}\rangle\langle\Psi_{AB}|]$, $\rho_A^2 = \rho_A$.*

There exist mixed entangled states as well (see Definition 2.4 below), but the theory in this case is much more complicated. In particular, there are no easy ways, in general, to decide whether a given mixed bipartite state is entangled or separable. The following few paragraphs provide a sketch of some basic ideas.

Theorem 2.3 (Purification of mixed states)**.** *For any density matrix $\rho \in \mathbb{M}(\mathbb{C}^m)$, there exists a normalized vector $|\Psi_\rho\rangle \in \mathbb{C}^m\otimes\mathbb{C}^m$ such that*

$$\rho = \mathrm{Tr}_{\mathbb{C}^m}\left[|\Psi_\rho\rangle\langle\Psi_\rho|\right]. \tag{2.19}$$

Proof. Let $\rho = \sum_{i=1}^m \lambda_i|\psi_i\rangle\langle\psi_i|$, $\lambda_i \geq 0$, $\sum_{i=1}^m \lambda_i = 1$, be a diagonalization of ρ. Let $\{|e_j\rangle : 1 \leq j \leq m\}$ be a complete orthonormal system in $\mathbb{C}^m$. Then, the vector in $\mathbb{C}^m\otimes\mathbb{C}^m$ defined by

$$|\Psi_\rho\rangle := \sum_{i=1}^m \sqrt{\lambda_i}(|\psi_i\rangle\otimes|e_i\rangle), \tag{2.20}$$

is normalized, i.e. $\langle\Psi_\rho|\Psi_\rho\rangle = 1$, and such that Eq. (2.19) holds. ∎

A very useful result is the following:

Theorem 2.4 (Schmidt decomposition)**.** *For any vector z in the tensor space $\mathbb{C}^m \otimes \mathbb{C}^n$, there always exists a complete orthonormal system $\{e_i\}_i \in \mathbb{C}^m$ and a complete orthonormal system $\{f_j\}_j \in \mathbb{C}^n$ such that*

$$z = \sum_k r_k (e_k \otimes f_k), \tag{2.21}$$

where $0 \leq r_k \in \mathbb{R}$. The number of non-zero coefficients r_k appearing in (2.21) is called the Schmidt number *of z and it is uniquely defined for any vector $z \in \mathbb{C}^m \otimes \mathbb{C}^n$.*

☞ A direct consequence of the Schmidt decomposition is that, given a pure state $|\Psi\rangle\langle\Psi| \in \mathbb{M}(\mathbb{C}^m \otimes \mathbb{C}^n)$, the reduced states $\mathrm{Tr}_{\mathbb{C}^m}[|\Psi\rangle\langle\Psi|]$ and $\mathrm{Tr}_{\mathbb{C}^n}[|\Psi\rangle\langle\Psi|]$ have the same eigenvalues and the same degeneracy indices.

The example studied in Exercise 2.3 explicitly shows that there exist vectors in tensor product spaces that cannot be written in tensor product form. Such vectors are called *entangled*. From entangled vectors, one defines *entangled states* as follows:

Definition 2.4 (Separable and Entangled Mixed States)**.** Given a composite bipartite system QR, a state $\rho_{QR} \in \mathscr{L}(\mathcal{H}_Q \otimes \mathcal{H}_R)$ is called *separable* if and only if there exist pure states $|\psi_j\rangle\langle\psi_j| \in \mathscr{L}(\mathcal{H}_Q)$ and pure states $|\phi_j\rangle\langle\phi_j| \in \mathscr{L}(\mathcal{H}_R)$, such that

$$\rho_{QR} = \sum_{j=1}^{J} p_j \Big(|\psi_j\rangle\langle\psi_j|_Q \otimes |\phi_j\rangle\langle\phi_j|_R \Big), \tag{2.22}$$

where J is finite and p_j's are probabilities, i.e. $p_j \geq 0$, $\sum_{j=1}^{J} p_j = 1$. Any state, which is not separable, is called *entangled*.

Remark 2.5. The phenomenon of entanglement is a purely quantum feature: two quantum particles, when in an entangled state, should be considered as a *single* quantum system, in the sense that a composite system in an entangled state is not *simply* the sum of its constituents. The example given in Exercise 2.3 taught us that having perfect knowledge about the state of the composite system does not imply (in general) any knowledge about the states of the constituents, in sharp contradiction with common sense.

3 Processing of quantum systems: quantum processors and quantum instruments

We are now ready to learn how quantum systems can be manipulated. This is the starting point to the understanding of how information can be encoded, transmitted, and decoded in a quantum information processing device.

Postulates 3, 4, and 5 tell that the following operations are allowed by quantum theory:

1. **Preparation**: one can prepare any quantum system in any chosen state
2. **Composition**: two (or more) quantum systems can always be composed together to form a composite quantum system
3. **Erasure**: one can always discard quantum systems, in any preferred order
4. **Closure**: any quantum system (composite or not) can always be perfectly isolated during a chosen time interval and made evolve according to any chosen unitary operator
5. **Measurement**: one can measure any observable of any quantum system

The above basic operations are at the basis of the so-called *quantum circuit* model of quantum information processing.

Example 3.1 (Evolution of open quantum systems)**.** How do open quantum systems evolve? Postulate 4, by itself, tells us how *closed*, i.e. perfectly isolated, quantum systems evolve. However, such an assumption is rarely satisfied in the real world, even classically (think of, e.g., thermalization or friction). The idea is that, in principle, we can consider a quantum system together with all the systems with which it is interacting, so that the *composite* system is actually an isolated system.

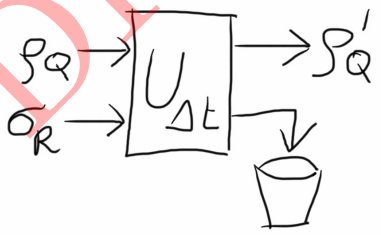

Figure 1: Simplest model of an open evolution.

The simplest way to model the time evolution of an open (i.e. non isolated) quantum system Q, initially in state ρ_Q, is as follows (see Figure 1): we first prepare another quantum system R in some initial state σ_R. We then compose Q with R and isolate them for a time interval Δt. The composite system QR, initially in state $\rho_Q \otimes \sigma_R$, after time Δt has evolved to $U_{\Delta t}(\rho_Q \otimes \sigma_R)U_{\Delta t}^\dagger$, where $U_{\Delta t}$ is a unitary operator in $\mathscr{L}(\mathcal{H}_Q \otimes \mathcal{H}_R)$, according to Postulate 4. Finally, the quantum system R is discarded. The overall operation on Q is mathematically written as follows:

$$\rho_Q \mapsto \rho'_Q = \mathrm{Tr}_R \left[U_{\Delta t}(\rho_Q \otimes \sigma_R)U_{\Delta t}^\dagger \right] . \square \tag{3.1}$$

☞ We now extend the correspondence in equation (3.1) to a mapping $\mathcal{E}$ as follows:

$$\mathcal{E}(X) := \mathrm{Tr}_R \left[U_{\Delta t}(X \otimes \sigma_R) U_{\Delta t}^\dagger \right], \tag{3.2}$$

where $X \in \mathscr{L}(\mathcal{H}_Q)$ is any linear operator. The situation is like the one depicted in Figure 2: we are now considering the apparatus consisting of σ_R, the unitary operator, and the partial trace as a *black box*, or a *quantum processor*, performing some kind of operation on the quantum system Q. Which are the properties of such a mapping?

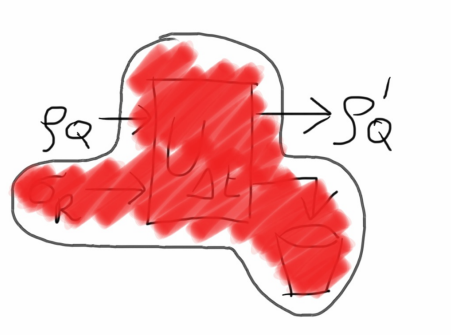

Figure 2: A quantum "processor", with one input wire (or the left) and one output wire (on the right).

1. The map $\mathcal{E} : \mathscr{L}(\mathcal{H}_Q) \to \mathscr{L}(\mathcal{H}_Q)$ is *linear*: the tensor product, the unitary operator and the partial trace are all linear, so that their composition is linear as well.

2. The map $\mathcal{E}$ is *trace-preserving*, i.e. $\mathrm{Tr}[\mathcal{E}(X)] = \mathrm{Tr}[X]$, for all $X \in \mathscr{L}(\mathcal{H}_Q)$. This property can be proved by using the cyclicity property of the trace (Theorem 1.4) as follows: $\mathrm{Tr}[\mathcal{E}(X)] = \mathrm{Tr}[U_{\Delta t}(X \otimes \sigma_R)U_{\Delta t}^\dagger] = \mathrm{Tr}[U_{\Delta t}^\dagger U_{\Delta t}(X \otimes \sigma_R)]$. Since the operator $U_{\Delta t}$ is unitary, $U_{\Delta t}^\dagger U_{\Delta t} = \mathbb{1}_{QR}$, so that $\mathrm{Tr}[\mathcal{E}(X)] = \mathrm{Tr}[X \otimes \sigma_R] = \mathrm{Tr}[X]\,\mathrm{Tr}[\sigma_R]$. Finally, since σ_R is a state, $\mathrm{Tr}[\sigma_R] = 1$, which implies that $\mathrm{Tr}[\mathcal{E}(X)] = \mathrm{Tr}[X]$, for all $X \in \mathscr{L}(\mathcal{H}_Q)$.

3. It is *positive*: if ρ_Q is a density matrix, $\mathcal{E}(\rho_Q)$ *must* be a density matrix, because we constructed the map $\mathcal{E}$ using only operations which are allowed by the postulates of quantum theory. By linearity then, for any positive operator $P \in \mathscr{L}(\mathcal{H}_Q)$, the operator $\mathcal{E}(P)$ is also a positive operator, i.e. the map $\mathcal{E}$ is linear, trace-preserving, and preserves positivity.

☞ Actually, there is more than positivity! In fact, instead of considering the initial quantum system Q on its own, we could imagine to input into our black box a subsystem of a composite quantum system Q_1Q_2, as shown in Figure 3. In this case, we obtain the following transformation:

$$\rho_{Q_1Q_2} \mapsto \rho'_{Q_1Q_2} = \mathrm{Tr}_R \left[(\mathbb{1}_{Q_1} \otimes U_{\Delta t})(\rho_{Q_1Q_1} \otimes \sigma_R)(\mathbb{1}_{Q_1} \otimes U_{\Delta t}^\dagger) \right]. \tag{3.3}$$

The above transformation can also be written as

$$\rho'_{Q_1Q_2} = (\mathrm{id}_{Q_1} \otimes \mathcal{E}_{Q_2})(\rho_{Q_1Q_2}), \tag{3.4}$$

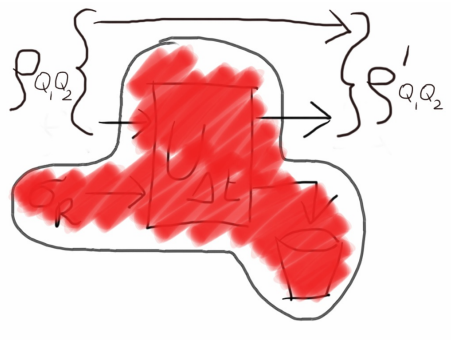

Figure 3: Why complete positivity?

where $\mathrm{id}_{Q_1} : \mathscr{L}(\mathcal{H}_{Q_1}) \to \mathscr{L}(\mathcal{H}_{Q_1})$ is the *identity map* (different from the identity matrix!!), i.e. $\mathrm{id}(X) = X$, for all $X \in \mathscr{L}(\mathcal{H}_{Q_1})$. Then, following the same arguments, the map $\mathrm{id}_{Q_1} \otimes \mathcal{E}_{Q_2}$ is again linear , trace-preserving, and positive. In fact, it is positive *for any choice* of Q_1! This property, which is a property of the map $\mathcal{E}$, is called *complete positivity.*

Theorem 3.1 (Unitary Representation of Quantum Processors)**.** *Given a quantum processor acting on quantum system Q, i.e. given a linear, trace-preserving, completely positive map $\mathcal{E} : \mathscr{L}(\mathcal{H}_Q) \to \mathscr{L}(\mathcal{H}_Q)$, there always exist an auxiliary quantum system R with Hilbert space $\mathcal{H}_R$, a density matrix $\sigma_R \in \mathscr{L}(\mathcal{H}_R)$, and a unitary operator $U \in \mathscr{L}(\mathcal{H}_Q \otimes \mathcal{H}_R)$, such that*

$$\mathcal{E}(\rho_Q) = \mathrm{Tr}_R \left[U(\rho_Q \otimes \sigma_R) U^\dagger \right], \tag{3.5}$$

for all density matrices $\rho_Q \in \mathscr{L}(\mathcal{H}_Q)$. Moreover, the density matrix σ_R can always be chosen to be a pure state, i.e. $\sigma_R^2 = \sigma_R$.

Remark 3.1. The evolution of an open quantum system Q during a time interval Δt, can always be represented by a quantum processor acting on the initial state. A closed system is a particular case of an open system. In that case, $\mathcal{E}_{\Delta t}(\rho) = U_{\Delta t} \rho_Q U_{\Delta t}^\dagger$, for any state ρ_Q of Q. □

Example 3.2 (No-Cloning Theorem)**.** Is the linearity condition important? (to be continued)

Example 3.3 (Partial Transposition)**.** As an example of a linear, trace-preserving map which is positive but not completely positive, consider, for any $X \in \mathbb{M}(\mathbb{C}^m)$, the transposition map $T : X \mapsto X^T$. First, let us check that the three properties of linearity, trace-preservation, and positivity are satisfied:

1. linearity: $(c_1 X + c_2 Y)^T = c_1 X^T + c_2 Y^T$, for all $X, Y \in \mathbb{M}(\mathbb{C}^m)$ and $c_1, c_2 \in \mathbb{C}$;
2. trace-preserving: the trace of an operator X equals the sum of the diagonal elements of a matrix representation of X; since the transposition operation does not modify the diagonal elements, it also preserves the trace;
3. positivity: let P be a positive operator; this implies that P is normal, i.e. $P^\dagger = P$; due to Remark 1.3, P^T is unitarily equivalent to P, that is, there exists a unitary operator $U \in \mathbb{M}(\mathbb{C}^m)$ such that $P^T = U^\dagger P U$; hence, P^T is positive if and only if P is positive.

However, the transposition map is not completely positive. As a counter-example, let us consider again the bipartite pure entangled state $|z\rangle\langle z|$ written in Eq. (2.17). In that example, we had a composite Hilbert space, made up of two two-dimensional parts $\mathbb{C}^2 \otimes \mathbb{C}^2$. We apply the transposition only on the first component, i.e. we apply the map $T_1 \otimes \mathrm{id}_2$,

$$(T_1 \otimes \mathrm{id}_2)(|z\rangle\langle z|) = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.6}$$

Now, if we find that the above matrix is not positive, i.e. has at least one negative eigenvalue, then we know that the map $T_1 \otimes \mathrm{id}_2$ does not preserve positivity, since the initial operator $|z\rangle\langle z|$ is positive (it is a density matrix!). If true, this would imply that the map T is not *completely* positive.

First of all, we notice that the matrix $(T_1 \otimes \mathrm{id}_2)(|z\rangle\langle z|)$ is self-adjoint, so it has a complete set of eigenvectors. Let us now consider the action of $(T_1 \otimes \mathrm{id}_2)(|z\rangle\langle z|)$ on the vector $|\chi\rangle = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}$.

We find that

$$\frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}. \tag{3.7}$$

We hence discovered that the vector $|\chi\rangle$ is an eigenvector of $(T_1 \otimes \mathrm{id}_2)(|z\rangle\langle z|)$ corresponding to the negative eigenvalue $-1/2$. Hence, the matrix $(T_1 \otimes \mathrm{id}_2)(|z\rangle\langle z|)$ is not positive, and the map T is not *completely positive*, even though linear, trace-preserving, and *positive*. □

Question 3.1 (reduced dynamics of a C-NOT gate)**.** In classical information, given two bits $a \in \{0,1\}$ and $b \in \{0,1\}$, the *two-bit gate* C-NOT (controlled-NOT) acts as follows

$$(a,b) \overset{\text{C-NOT}}{\longmapsto} (a, b \oplus a). \tag{3.8}$$

The first bit is called the *control bit*, the second bit is the *target bit*. If the value of the control bit is 0, then the target bit is left unchanged. If the value of the control bit is 1, then the target bit is *flipped*, i.e. if it was 0 it is mapped in 1, and viceversa. Noteworthy is the fact that *the control bit is left unchanged*.

In quantum computation, the operation generalizing the C-NOT gate is defined as follows. The two bits are replaced by two *qubits* (see Example 2.1), i.e. two quantum systems C (control) and T (target), with Hilbert spaces $\mathcal{H}_C \cong \mathbb{C}^2$ and $\mathcal{H}_T \cong \mathbb{C}^2$. The *quantum C-NOT gate* is described by the following unitary operator:

$$U_{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{3.9}$$

In the conventional C-NOT gate, we noticed that the control bit is left unchanged. Is this the case also for the quantum C-NOT gate? Let us consider the situation in which the target qubit is initialized in the state $\rho_T = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. In this case, the transformation that the control qubit

undergoes is described by the following linear, trace-preserving, completely positive map:

$$\rho_C \longmapsto \mathrm{Tr}_T \left\{ U_{CNOT} \left[\rho_C \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}_T \right] U_{CNOT}^\dagger \right\}. \tag{3.10}$$

How to write the above transformation in a more explicit form? □

Theorem 3.2 (Kraus Representation of Quantum Processors)**.** *A map $\mathcal{E} : \mathbb{M}(\mathbb{C}^m) \to \mathbb{M}(\mathbb{C}^n)$ is linear, trace-preserving and completely positive (in short, CPTP) if and only if there exists a finite family of $(m \times n)$ matrices $\{E_k; 1 \leq k \leq K\}$, satisfying the normalization condition $\sum_{k=1}^K E_k^\dagger E_k = \mathbb{1}_m$, such that, for all states $\rho \in \mathbb{M}(\mathbb{C}^m)$,*

$$\mathcal{E}(\rho) = \sum_{k=1}^K E_k \rho E_k^\dagger. \tag{3.11}$$

Remark 3.2 (Very useful for advanced applications). Any completely positive, trace preserving linear map can always be written in the Kraus form (3.11). Any linear map $\mathcal{L} : \mathbb{M}(\mathbb{C}^m) \to \mathbb{M}(\mathbb{C}^n)$ can always be written as $\mathcal{L}(\rho) = \sum_k A_k \rho B_k$, where $A_k \in \mathscr{L}(\mathbb{C}^m, \mathbb{C}^n)$ $B_k \in \mathbb{M}(\mathbb{C}^n, \mathbb{C}^m)$: such a linear map is trace-preserving if and only if $\sum_k B_k A_k = \mathbb{1}_m$. A linear map $\mathcal{L} : \mathbb{M}(\mathbb{C}^m) \to \mathbb{M}(\mathbb{C}^n)$ maps self-adjoint matrices into self-adjoint matrices if and only if $\mathcal{L}(\rho) = \sum_k c_k A_k \rho A_k^\dagger$, with $A_k \in \mathscr{L}(\mathbb{C}^m, \mathbb{C}^n)$ and $c_k \in \mathbb{R}$.
A natural question is the following: is there a "generalized Kraus form" for positive (possibly not completely positive) linear maps? No! Only when $\dim \mathcal{H}_Q = 2$, then any positive map $\mathcal{P} : \mathbb{M}(\mathbb{C}^2) \to \mathbb{M}(\mathbb{C}^2)$ can be decomposed as $\mathcal{P}(\rho) = \sum_k E_k \rho E_k^\dagger + \sum_{k'} F_{k'} \rho^T F_{k'}^\dagger$. □

Exercise 3.1 (C-NOT gate, continued)**.** With a little effort, we should now be able to answer Question 3.1. Our aim is to write the Kraus representation of the map

$$\rho_C \longmapsto \mathrm{Tr}_T \left\{ U_{CNOT} \left[\rho_C \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}_T \right] U_{CNOT}^\dagger \right\}. \tag{3.12}$$

A simple computation shows that the map acts as follows:

$$\rho_C = \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \xmapsto{\mathcal{E}} \mathcal{E}(\rho_C) = \begin{pmatrix} r_{11} & 0 \\ 0 & r_{22} \end{pmatrix}, \tag{3.13}$$

for all $\rho_C \in \mathscr{L}(\mathcal{H}_C)$. In Kraus representation, this can be written as

$$\mathcal{E}(\rho_C) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \rho_C \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \rho_C \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.14}$$

Exercise 3.2 (Operator-sum representation)**.** Here we derive the Kraus form, Theorem 3.2, for quantum processors. The ingredients we need to perform the calculation are the following:

1. a preferred basis $\{\psi_i\}$ in $\mathcal{H}_Q$
2. a preferred basis $\{\phi_j\}$ in $\mathcal{H}_R$
3. the diagonal form $\sum_\lambda \lambda |\lambda\rangle\langle\lambda|_R$ of σ_R

4. the expansion of U as $U = \sum_{i,i'} \sum_{j,j'} u_{ij,i'j'} |\psi_i \otimes \phi_j\rangle\langle\psi_{i'} \otimes \phi_{j'}|$

We then have

$$
\begin{aligned}
&\mathrm{Tr}_R[U(\rho_Q \otimes \sigma_R)U^\dagger] \\
&= \sum_\lambda \lambda \,\mathrm{Tr}_R \left[U(\rho_Q \otimes |\lambda\rangle\langle\lambda|_R)U^\dagger \right] \\
&= \sum_\lambda \lambda \sum_{i,i'} \sum_{j,j'} u_{ij,i'j'} \,\mathrm{Tr}_R \left[(|\psi_i\rangle\langle\psi_{i'}|\rho_Q) \otimes \left(|\phi_j\rangle \underbrace{\langle\phi_{j'}|\lambda\rangle}_{c_{j'\lambda}\in\mathbb{C}} \langle\lambda|_R \right) U^\dagger \right] \\
&= \sum_\lambda \lambda \sum_{i,i',j,j'} u_{ij,i'j'} c_{j'\lambda} \,\mathrm{Tr}_R \left[(|\psi_i\rangle\langle\psi_{i'}|\rho_Q) \otimes |\phi_j\rangle\langle\lambda|_R \, U^\dagger \right] \\
&= \sum_\lambda \lambda \sum_{i,i',j,j'} u_{ij,i'j'} c_{j'\lambda} \sum_{\iota,\iota',\upsilon,\upsilon'} u^*_{\iota'\upsilon',\iota\upsilon} \,\mathrm{Tr}_R \left[(|\psi_i\rangle\langle\psi_{i'}|\rho_Q|\psi_\iota\rangle\langle\psi_{\iota'}|) \otimes \left(|\phi_j\rangle \underbrace{\langle\lambda|\phi_\upsilon\rangle}_{d_{\lambda\upsilon}=c^*_{\upsilon\lambda}} \langle\phi_{\upsilon'}|_R \right) \right] \\
&= \sum_\lambda \lambda \sum_{i,i',j,j'} u_{ij,i'j'} c_{j'\lambda} \sum_{\iota,\iota',\upsilon,\upsilon'} u^*_{\iota'\upsilon',\iota\upsilon} c^*_{\upsilon\lambda} \,\mathrm{Tr}_R \left[(|\psi_i\rangle\langle\psi_{i'}|\rho_Q|\psi_\iota\rangle\langle\psi_{\iota'}|) \otimes |\phi_j\rangle\langle\phi_{\upsilon'}| \right] \\
&= \sum_\lambda \lambda \sum_{i,i',j,j'} u_{ij,i'j'} c_{j'\lambda} \sum_{\iota,\iota',\upsilon,\upsilon'} u^*_{\iota'\upsilon',\iota\upsilon} c^*_{\upsilon\lambda} \delta_{j,\upsilon'} |\psi_i\rangle\langle\psi_{i'}|\rho_Q|\psi_\iota\rangle\langle\psi_{\iota'}| \\
&= \sum_\lambda \lambda \sum_{i,i',j,j'} u_{ij,i'j'} c_{j'\lambda} \sum_{\iota,\iota',\upsilon} u^*_{\iota'j,\iota\upsilon} c^*_{\upsilon\lambda} |\psi_i\rangle\langle\psi_{i'}|\rho_Q|\psi_\iota\rangle\langle\psi_{\iota'}| \\
&= \sum_{i,i'} \sum_{\iota,\iota'} \sum_\lambda \sum_{j,j'} \sum_\upsilon \lambda u_{ij,i'j'} c_{j'\lambda} u^*_{\iota'j,\iota\upsilon} c^*_{\upsilon\lambda} |\psi_i\rangle\langle\psi_{i'}|\rho_Q|\psi_\iota\rangle\langle\psi_{\iota'}| \\
&= \sum_{i,i'} \sum_{\iota,\iota'} \sum_\lambda \sum_{j,j'} \lambda u_{ij,i'j'} c_{j'\lambda} \underbrace{\sum_\upsilon u^*_{\iota'j,\iota\upsilon} c^*_{\upsilon\lambda}}_{a_{\iota'\iota j\lambda}} |\psi_i\rangle\langle\psi_{i'}|\rho_Q|\psi_\iota\rangle\langle\psi_{\iota'}| \\
&= \sum_{i,i'} \sum_{\iota,\iota'} \sum_\lambda \sum_j \lambda \underbrace{\sum_{j'} u_{ij,i'j'} c_{j'\lambda} a_{\iota'\iota j\lambda}}_{a^*_{ii'j\lambda}} |\psi_i\rangle\langle\psi_{i'}|\rho_Q|\psi_\iota\rangle\langle\psi_{\iota'}| \\
&= \sum_\lambda \sum_j \underbrace{\left(\sqrt{\lambda} \sum_{i,i'} a^*_{ii'j\lambda} |\psi_i\rangle\langle\psi_{i'}| \right)}_{A_{j\lambda}} \rho_Q \underbrace{\left(\sqrt{\lambda} \sum_{\iota,\iota'} a_{\iota'\iota j\lambda} |\psi_\iota\rangle\langle\psi_{\iota'}| \right)}_{A^\dagger_{j\lambda}} .
\end{aligned}
\tag{3.15}
$$

After this lengthy, though straightforward calculation, we finally arrive at the form

$$
\mathrm{Tr}_R[U(\rho_Q \otimes \sigma_R)U^\dagger] = \sum_\lambda \sum_j A_{j\lambda} \rho_Q A^\dagger_{j\lambda}, \tag{3.16}
$$

for a suitable family of operators $A_{j\lambda} \in \mathscr{L}(\mathcal{H}_Q)$. □

3.1 Read-out stage: quantum measurement processes

In building a quantum processor, up to now we only used the operations of preparation, composition, erasure, and closure. We still didn't use the fact that measurements can be also performed: any non-trivial computation needs to terminate with the read-out of the computation results! (Indeed, what would be the point of performing a computation, if the results are not observed?)

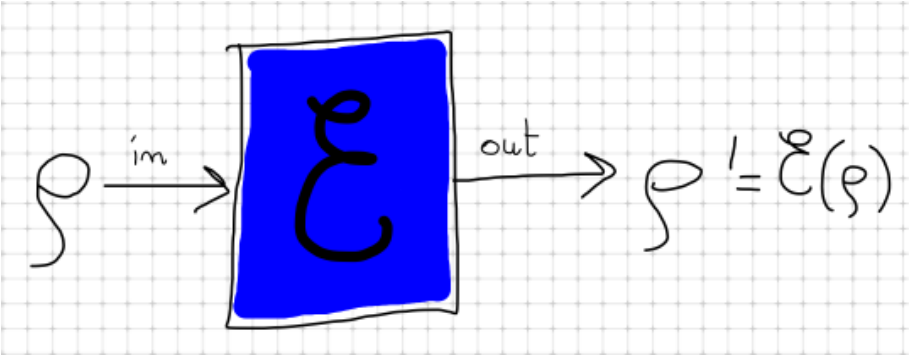

Figure 4: Any quantum processor is represented by a completely positive, trace-preserving linear map $\mathcal{E} : \mathscr{L}(\mathcal{H}_{\text{in}}) \to \mathscr{L}(\mathcal{H}_{\text{out}})$. Still, in this picture, there is no measurement taking place.

The read-out stage always corresponds to the measurement of some physical quantity: if the measurement gives value μ, for example, then we say that the computation terminated and returned the result $f(\mu)$, where f is some appropriate function. For example, when we say that we are reading a bit recorded on a hard disk drive, we are actually measuring the direction of a magnetic dipole, and saying that the bit is zero or one, depending on (i.e. as a function of) the measured value.

The same happens in quantum information theory. Figure 5 depicts a situation often encountered: an input state ρ is fed into a circuit composed by various quantum processors, representing time-evolutions, computations, or arbitrarily complicated combinations of both. After the processors has performed their action, in order to obtain some information about the result of the computation, a measurement (let's say, the measurement of some observable A) is performed on the output produced.

Now, let us imagine a quantum processor whose output is a composite quantum system, for example QR, as shown in Figure 6. Let us imagine that only the output branch labeled by R is measured. The question is the following: since the measurement is performed only on the R subsystem, is subsystem Q affected by it? If so, what happens to subsystem Q *after* the measurement on R is performed?

The postulates of quantum theory imply the following:

Theorem 3.3. *Let the state of a composite quantum system QR be represented by the density matrix ω_{QR}. Let $\omega_Q = \text{Tr}_R[\omega_{QR}]$ and $\omega_R = \text{Tr}_Q[\omega_{QR}]$ be the reduced states for subsystems Q and R, respectively. Suppose that the measurement of an observable of R, represented by the self-adjoint operator $A_R \in \mathscr{L}(\mathcal{H}_R)$ with spectral decomposition $A_R = \sum_\mu \mu \Pi_R^A(\mu)$, is performed. Then, the probability that an outcome μ is observed is given by $\text{Tr}\left[\omega_R \Pi_R^A(\mu)\right]$. Correspondingly, if an outcome μ is observed, the subsystem Q is left in a state, which depends on μ according to the following formula:*

$$\omega_Q(\mu) = \frac{1}{\text{Tr}\left[\omega_R \Pi_R^A(\mu)\right]} \text{Tr}_R\left\{\omega_{QR}\ \left[\mathbb{1}_Q \otimes \Pi_R^A(\mu)\right]\right\}. \tag{3.17}$$

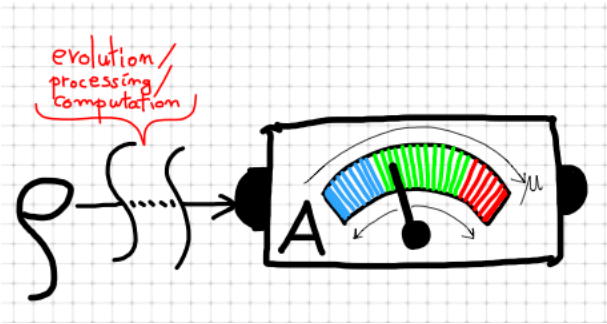

Figure 5: A measurement is the ending point of a quantum process. Here a measurement of the physical quantity represented by the self-adjoint operator $A = \sum_\mu \mu \Pi^A(\mu)$ is performed after an arbitrary quantum processor acted upon the input state ρ. *Before* performing the measurement, it is generally impossible to predict the measurement outcome: only the probabilities with which each outcome will be obtained, i.e. $\Pr\{A = \mu \| \rho\} = \mathrm{Tr}[\rho\, \Pi^A(\mu)]$, can be computed. Only *after* the measurement has been performed and the outcome $\bar{\mu}$ has been obtained, one can say that the physical quantity A is equal to $\bar{\mu}$, and that the result of the computation is obtained.

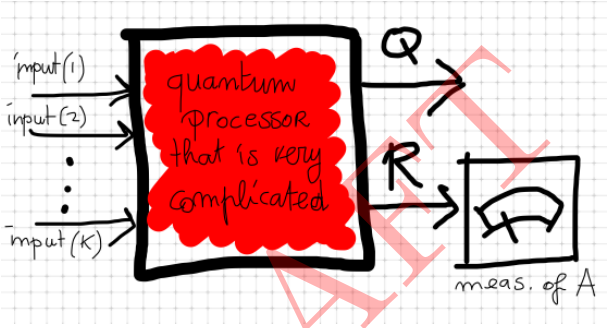

Figure 6: A measurement of the physical quantity A performed only on one branch of the computation. The state that correctly describes the remaining branch Q in general will depend on the outcome of the measurement performed on R. The formula to compute such a state is given in Theorem 3.3.

Example 3.4 (Averaging outcomes = Discarding)**.** We already encountered the concept of *expectation value* of an observable, see equation (2.6), which has been defined as the average of the possible outcomes of a measurement. Theorem 3.3 suggests that it is possible to compute, in a similar way, the *average state* in which subsystem Q is left after a measurement performed

on R. Such a state can be computed as follows:

$$\begin{aligned}
\overline{\omega_Q} &= \sum_\mu \Pr\{A_R = \mu \| \omega_R\} \omega_Q(\mu) \\
&= \sum_\mu \mathrm{Tr}\left[\omega_R \, \Pi_R^A(\mu)\right] \frac{1}{\mathrm{Tr}\left[\omega_R \, \Pi_R^A(\mu)\right]} \mathrm{Tr}_R\left\{\omega_{QR} \, \left[\mathbb{1}_Q \otimes \Pi_R^A(\mu)\right]\right\} \\
&= \sum_\mu \mathrm{Tr}_R\left\{\omega_{QR} \, \left[\mathbb{1}_Q \otimes \Pi_R^A(\mu)\right]\right\} \\
&= \mathrm{Tr}_R\left\{\omega_{QR} \, \left(\mathbb{1}_Q \otimes \left[\sum_\mu \Pi_R^A(\mu)\right]\right)\right\} \\
&= \mathrm{Tr}_R[\omega_{QR}],
\end{aligned}$$

where the last step follows from the fact that $\sum_\mu \Pi_R^A(\mu) = \mathbb{1}_R$, *always*, as argued in equation (2.1). Remarkably, the density matrix $\overline{\omega_Q}$ does not depend on the observable A measured on R. We conclude that averaging the state of Q over the measurement outcomes is completely equivalent to discarding subsystem R, without performing any measurement on it.

The setup represented in Figure 6, composed by a quantum processor and a partial measurement, can also be seen as a new kind of quantum processor: we call this a *quantum instrument.* A quantum instrument is a quantum processor with two outputs: a *quantum* output (the system Q in Figure 6), which can be fed into the next computational step, and a *classical* output (the measurement outcome μ), which can be used to condition (in a structure like `if[...] then[...]`) the next computational step. A typical quantum circuit with quantum processors and quantum instruments will look as in Figure 7.

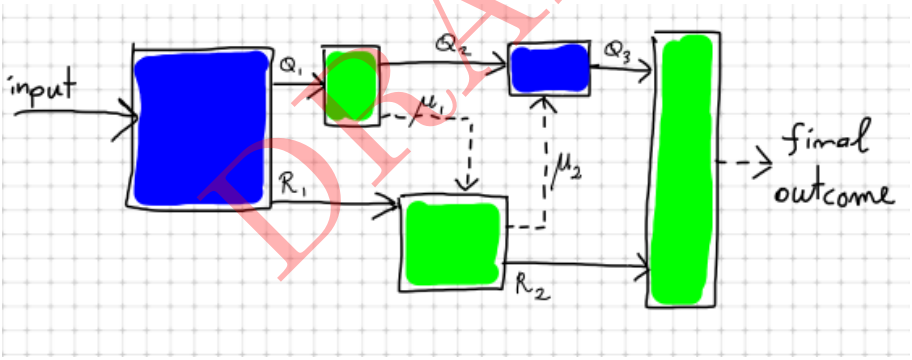

Figure 7: A typical quantum circuit composed by quantum processors (blue) and quantum instruments (green). Continuous lines correspond to computational branches carrying quantum systems, dashed lines correspond to computational branches carrying classical information (i.e. a measurement outcome).

☞ Of course, a quantum processor can be seen as a particular quantum instrument, i.e. a quantum instrument with *only one* possible outcome. Hence, we can treat quantum processors and quantum instruments on the same footing – however, for the sake of clarity, we will keep calling them with different names.

We conclude by stating, without proof, the following characterization of quantum instruments, generalizing the Kraus representation for quantum processors given in Theorem 3.2:

Theorem 3.4 (Kraus form for quantum instruments)**.** *A* quantum instrument *is defined by:*

1. *an m-dimensional input quantum system (associated with $\mathbb{C}^m$) and an n-dimensional output quantum system (associated with $\mathbb{C}^n$);*
2. *a set $\mathcal{X}$ of possible outcomes μ;*
3. *for each outcome μ, a family of $m \times n$ matrices $\{E_{\mu,k}; 1 \leq k \leq K_\mu\}$ such that*

$$\sum_{\mu \in \mathcal{X}} \sum_{k=1}^{K_\mu} E_{\mu,k}^\dagger E_{\mu,k} = \mathbb{1}_m. \tag{3.18}$$

Then, if the system fed into the quantum instrument is in state $\rho \in \mathscr{L}(\mathbb{C}^m)$, the probability of obtaining an outcome μ is given by

$$\Pr\{\mu \| \rho\} = \mathrm{Tr}\left[\rho \sum_{k=1}^{K_\mu} E_{\mu,k}^\dagger E_{\mu,k}\right], \tag{3.19}$$

and the corresponding output quantum state $\rho'(\mu) \in \mathscr{L}(\mathbb{C}^n)$ is given by

$$\rho'(\mu) = \frac{1}{\Pr\{\mu \| \rho\}} \sum_{k=1}^{K_\mu} E_{\mu,k} \rho E_{\mu,k}^\dagger. \tag{3.20}$$

☞ It should be clear that, in the case of an instrument with only one possible outcome (i.e. with the set $\mathcal{X}$ containing only one element), Theorem 3.4 above reduces to Theorem 3.2.

Appendix: the idea of indirect measurement model

Example 3.5 (Indirect measurement model)**.** According to our everyday intuition, we see that, when we measure a physical quantity of a physical system (like temperature, weight, etc), the physical system is there before, during, and after the measurement process. However, Postulate 3 of Quantum Theory tells us only about how to compute the probability distribution according to which outcomes are obtained, without mentioning what happens to the quantum system which is measured. Postulate 3 only states that, if a *direct* measurement is performed on a quantum system, an outcome is obtain with a certain probability: in a sense, then, the quantum system itself, after the measurement, is no more there in the formalism. Is this really the case?

The answer to this question comes from the calculations we just computed. The idea is that, in quantum theory, we always have to imagine to act *operationally*. What really happens during a measurement process, for example, the weighing of an apple? We take an apple, we put it on a (digital) scale, and we read a number on the screen. Hence, what we *directly* measure (i.e. *read*) when we weigh an apple is not the weight of the apple, but the numbers that appear on the scale's screen. We then hope that the numbers produced by the scale are *correlated* with the "true" weight of the objects we put on it, in the sense that, if we read 100 grams on the scale, we hope that the apple also weighs 100 grams (or something very close to such a value). Hence, when we weigh an apple using a scale, we are *indirectly* measuring the weight of the apple, by measuring (i.e. reading) the state of the scale's screen. (Actually—this, again, is not completely true, since the scale's screen itself is read via our retina, which collects the light

scattered on it; the retina, in turns, converts the light signals into electric signals in the optical nerve, and so on.....)

If we think carefully, then, every physical measurement process is in fact an indirect measurement process. The word "measurement" used in Postulate 3 is better to be understood as a mere *theoretical* abstraction and should be clearly distinguished from a *physical* measurement process. In order to do so, from now on we will use the terminology *direct measurement* to denote "abstract measurements", otherwise all measurement processes will be meant to be "physical" (i.e. indirect).

The operational picture is the following: suppose that we want to measure an observable of the quantum system Q (let's say, a "quantum apple"). In order to do that, we make the quantum apple interact with an auxiliary quantum system R (a "quantum scale"), and then we directly measure a suitable observable of the quantum scale, whose values we know being correlated with the weight of the quantum apple. Then, the auxiliary *direct* measurement performed on the quantum scale randomly produces an outcome (according to Postulate 3) and removes the quantum scale from the formalism, leaving only the quantum apple together with the outcome of the measurement of its weight. Therefore, it makes perfect sense to speak about the state of a quantum system *after* a measurement process. □

4 Two primitive protocols: quantum teleportation and quantum super-dense coding

4.1 Quantum teleportation

As a paramount application of what we have learned until now, we will describe the task of *quantum teleportation* in detail. Teleportation here has not to be understood as the teleportation appearing in some science-fiction movies, where objects (or persons!) are "teleported" from a place to another almost instantaneously. The process called quantum teleportation is able to transfer *the state* of a quantum system *here* to another quantum system *there*, however far this is, simply by communicating a limited amount of classical data.

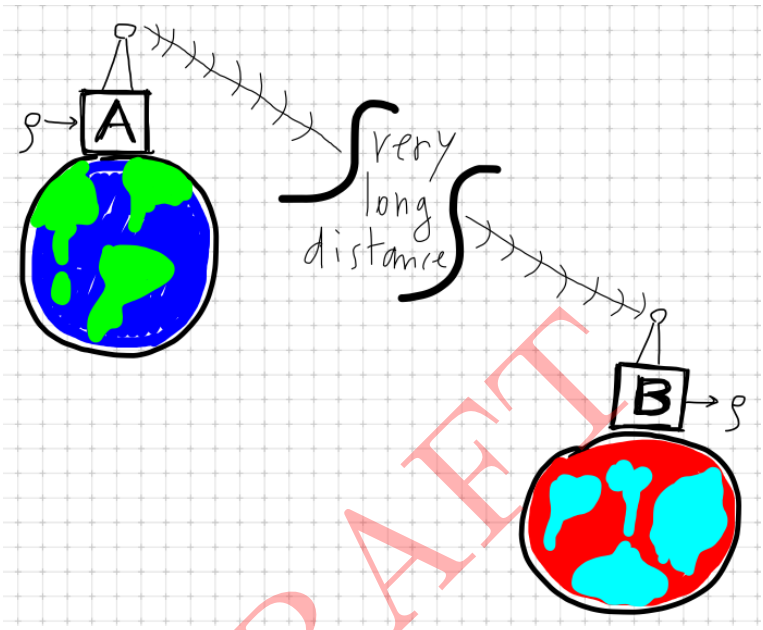

Figure 8: Quantum teleportation: two agents A and B have to transmit an unknown quantum state by using a classical communication channel only.

The scenario of quantum teleportation is the following: there are two agents A (Alice) and B (Bob) which have to successfully complete a mission. The mission starts at time $t = t_0$. At that time, the agent Alice is given a quantum system Q. Both Alice and Bob don't know the state (i.e. they do not know the density operator representing the state) of Q. The agents' mission is to *exactly* transmit the state of Q from Alice to Bob, by communicating only through a classical communication channel (i.e. Alice cannot "send" the quantum system Q to Bob).

Is there a way to accomplish the mission?

In order to model the protocol used by Alice and Bob, we will exploit the formalism of quantum processors and quantum instruments that we introduced in Section 3. The situation is depicted in Figure 9: Alice has the quantum system Q and her own quantum computer A, while Bob has only his own quantum computer B. They can apply any possible quantum processor and any possible quantum instrument on their respective systems, but they can exchange only classical information, i.e. they can exchange only the outcomes they obtain from the computations they choose (in the picture, the "wires" between Alice and Bob are only dashed wires, i.e. computational branches carrying only classical information, see Figure 7).

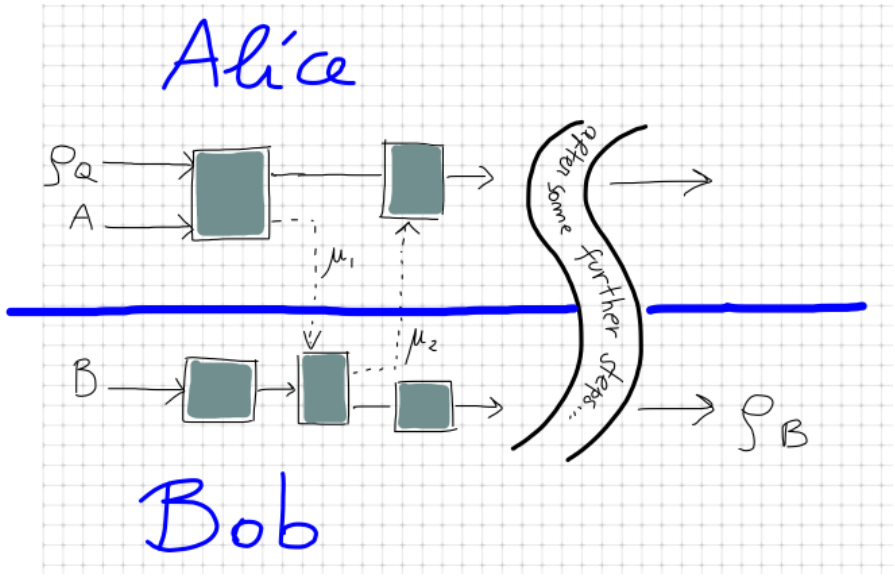

Figure 9: By running a computation that exchanges only classical information (i.e. measurement outcomes), Alice and Bob are required to "teleport" the state ρ from system Q (in Alice's hands) to system B (in Bob's hands).

Let us start with the simple case in which the quantum system Q is a qubit (see Example 2.1), i.e. a two-dimensional quantum system $\mathcal{H}_Q \cong \mathbb{C}^2$, and its state is pure, i.e. such that $\rho_Q = |\psi\rangle\langle\psi|_Q$ where $|\psi\rangle \in \mathbb{C}^2$ is a normalized vector (see Definition 2.1). Let us write

$$|\psi\rangle = \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}, \tag{4.1}$$

for some $c_0, c_1 \in \mathbb{C}$, such that $|c_0|^2 + |c_1|^2 = 1$.

☞ The key point is that the rules forbid Alice and Bob to exchange quantum states from time t_0 onwards—but nothing prevent Alice and Bob to meet *before* t_0, let's say at some time $t = t_{-1} < t_0$.

Let us imagine that, at time $t_{-1} < t_0$, Alice and Bob met, and prepared a composite (bipartite) quantum system AB, with Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B \cong \mathbb{C}^2 \otimes \mathbb{C}^2$, in the pure state $|z\rangle\langle z|$, where

$$|z\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{4.2}$$

(The density matrix $|z\rangle\langle z|$ has already been explicitly computed in equation (2.17).) After the quantum system AB has been prepared in the state $|z\rangle\langle z|$, Alice keeps the A subsystem, leaving the B subsystem to Bob. Alice and Bob then separate and reach their destination. We can imagine that this is actually the *prequel* to our story. How can Alice and Bob use the quantum system AB to achieve their mission?

The following construction, known as *quantum teleportation*, first appeared in 1993, in a very famous paper. At $t = t_0$, Alice holds systems Q (in the unknown state $|\psi\rangle\langle\psi|$) and A

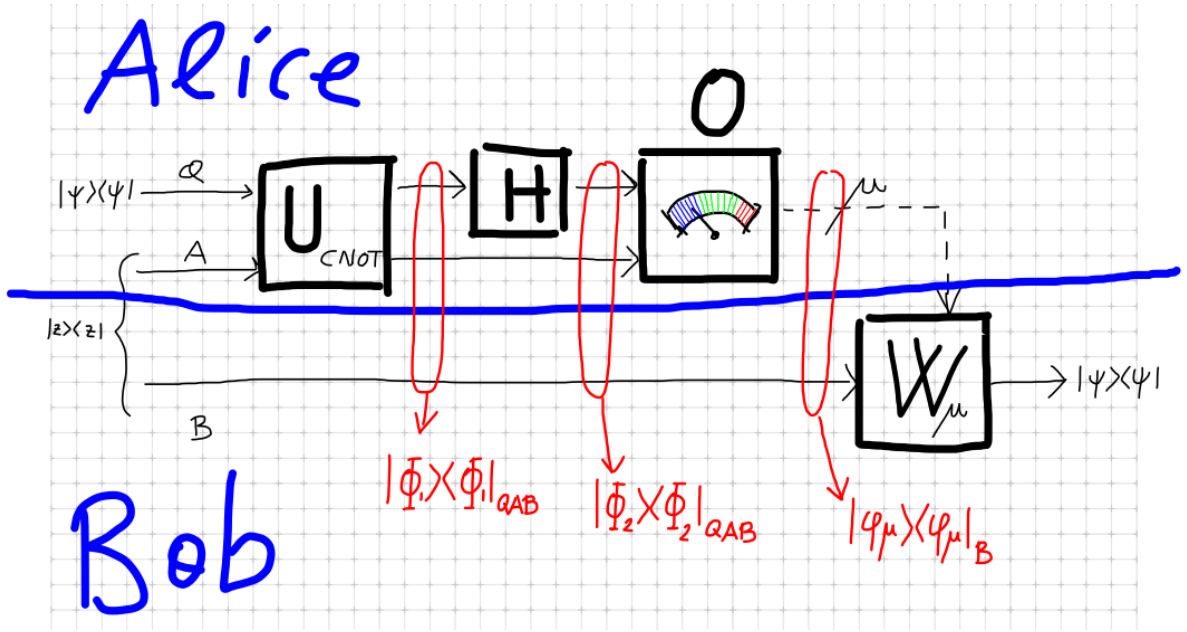

Figure 10: The circuit that teleports quantum states from Alice to Bob.

(her share of $|z\rangle\langle z|$). First, she applies on systems Q and A a control-NOT (introduced in Question 3.1). We recall that the unitary operator corresponding to a C-NOT gate is given by

$$U_{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

In the case considered here, the control qubit is represented by system Q, while the target qubit is represented by system A.

After Alice applied the C-NOT gate on her qubits, the state of the composite system QAB is given by (see Figure 10)

$$|\Phi_1\rangle\langle\Phi_1|_{QAB} = (U_{CNOT} \otimes \mathbb{1}_2)\ (|\psi\rangle\langle\psi|_Q \otimes |z\rangle\langle z|_{AB})\ (U_{CNOT} \otimes \mathbb{1}_2)^\dagger .$$

The explicit calculation of $|\Phi_1\rangle\langle\Phi_1|_{QAB}$ can be obtained from the form of the vector $|\Phi_1\rangle_{QAB}$,

which is given as follows:

$$\begin{aligned}
|\Phi_1\rangle_{QAB} =& U_{CNOT} \otimes \mathbb{1}_2 \left(|\psi\rangle_Q \otimes |z\rangle_{AB}\right) \\
=& U_{CNOT} \otimes \mathbb{1}_2 \left[\frac{c_0}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}_Q \otimes \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}_{AB} + \frac{c_1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}_Q \otimes \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}_{AB} \right] \\
=& U_{CNOT} \otimes \mathbb{1}_2 \left[\frac{1}{\sqrt{2}} \begin{pmatrix} c_0 \\ 0 \\ c_1 \\ 0 \end{pmatrix}_{QA} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_B + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ c_0 \\ 0 \\ c_1 \end{pmatrix}_{QA} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}_B \right] \\
=& \frac{1}{\sqrt{2}} \left[\begin{pmatrix} c_0 \\ 0 \\ 0 \\ c_1 \end{pmatrix}_{QA} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_B + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ c_0 \\ c_1 \\ 0 \end{pmatrix}_{QA} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}_B \right] \\
=& \frac{1}{\sqrt{2}} \left[c_0 \begin{pmatrix} 1 \\ 0 \end{pmatrix}_Q \otimes \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}_{AB} + c_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix}_Q \otimes \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}_{AB} \right].
\end{aligned} \tag{4.3}$$

After this, Alice further applies on system Q alone the unitary operation H represented by the matrix

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{4.4}$$

At this point, the state of composite system QAB is given by $|\Phi_2\rangle\langle\Phi_2|_{QAB}$, where the vector $|\Phi_2\rangle_{QAB}$ is given by the following equation

$$\begin{aligned}
|\Phi_2\rangle_{QAB} =& (H \otimes \mathbb{1}_4)|\Phi_1\rangle_{QAB} \\
=& \frac{1}{2} \left[c_0 \begin{pmatrix} 1 \\ 1 \end{pmatrix}_Q \otimes \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}_{AB} + c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix}_Q \otimes \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}_{AB} \right].
\end{aligned} \tag{4.5}$$

Simply by re-grouping the terms, the above state can also be written as follows:

$$\begin{aligned}
&|\Phi_2\rangle_{QAB} \\
=& \frac{1}{2} \left[\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}_{QA} \otimes \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}_B + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}_{QA} \otimes \begin{pmatrix} c_1 \\ c_0 \end{pmatrix}_B + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}_{QA} \otimes \begin{pmatrix} c_0 \\ -c_1 \end{pmatrix}_B + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}_{QA} \otimes \begin{pmatrix} -c_1 \\ c_0 \end{pmatrix}_B \right].
\end{aligned} \tag{4.6}$$

In the above equation, we recognize the vectors appearing in the QA subsystem as the standard basis of $\mathbb{C}^4$, which was denoted in Definition 1.4 as follows:

$$|e_{00}\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \; |e_{01}\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \; |e_{10}\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \; |e_{11}\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{4.7}$$

We moreover introduce in $\mathbb{C}^2$ the following vectors:

$$|\varphi_{00}\rangle = \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}, \ |\varphi_{01}\rangle = \begin{pmatrix} c_1 \\ c_0 \end{pmatrix}, \ |\varphi_{10}\rangle = \begin{pmatrix} c_0 \\ -c_1 \end{pmatrix}, \ |\varphi_{11}\rangle = \begin{pmatrix} -c_1 \\ c_0 \end{pmatrix}, \tag{4.8}$$

so that equation (4.6) can be rewritten as

$$|\Phi_2\rangle_{QAB} = \frac{1}{2} \sum_{\mu \in \{00,01,10,11\}} |e_\mu\rangle_{QA} \otimes |\varphi_\mu\rangle_B. \tag{4.9}$$

Notice that the vectors $|e_\mu\rangle$ are orthonormal, in the sense that $\langle e_\mu | e_{\mu'}\rangle = \delta_{\mu,\mu'}$ (for the definition of δ function, go back to Definition 1.4). The vectors $|\varphi_\mu\rangle$ are all normalized, since $\langle \varphi_\mu | \varphi_\mu \rangle = |c_0|^2 + |c_1|^2 = 1$ for all μ, however, they are *not* orthogonal. The state of the composite system QAB at this point can be written as follows:

$$|\Phi_2\rangle\langle\Phi_2|_{QAB} = \frac{1}{4} \sum_{\mu \in \{00,01,10,11\}} \sum_{\mu' \in \{00,01,10,11\}} |e_\mu\rangle\langle e_{\mu'}|_{QA} \otimes |\varphi_\mu\rangle\langle\varphi_{\mu'}|_B. \tag{4.10}$$

Next, Alice measures the observable O_{QA} on the composite system QA, represented by the self-adjoint matrix

$$O_{QA} = \begin{pmatrix} \ell_{00} & 0 & 0 & 0 \\ 0 & \ell_{01} & 0 & 0 \\ 0 & 0 & \ell_{10} & 0 \\ 0 & 0 & 0 & \ell_{11} \end{pmatrix}, \tag{4.11}$$

with $\ell_\mu \in \mathbb{R}$. The only condition we impose on O_{QA} is that $\ell_{00} \neq \ell_{01} \neq \ell_{10} \neq \ell_{11}$, so that the spectral projectors $\Pi^O(\ell_\mu)$ appearing in equation (1.13) are equal to $|e_\mu\rangle\langle e_\mu|$, for all $\mu \in \{00, 01, 10, 11\}$.
According to Postulate 3 and Postulate 5, Alice obtains the outcome ℓ_μ with probability

$$\begin{aligned} \Pr\{O_{QA} = \ell_\mu\} &= \mathrm{Tr}\left[(\Pi^O(\ell_\mu) \otimes \mathbb{1}_2) \ |\Phi_2\rangle\langle\Phi_2|_{QAB}\right] \\ &= \frac{1}{4}, \end{aligned} \tag{4.12}$$

for all $\mu \in \{00, 01, 10, 11\}$. The corresponding state at Bob's side (i.e. the state that describes Bob's particle *right after* Alice obtained the outcome ℓ_μ) is given by the formula (see Theorem 3.3)

$$\begin{aligned} \omega_B(\ell_\mu) &= \frac{1}{\Pr\{O_{QA} = \ell_\mu\}} \mathrm{Tr}_{QA}\left[(\Pi^O(\ell_\mu) \otimes \mathbb{1}_2) \ |\Phi_2\rangle\langle\Phi_2|_{QAB}\right] \\ &= 4 \, \mathrm{Tr}_{QA}\left[(\Pi^O(\ell_\mu) \otimes \mathbb{1}_2) \ |\Phi_2\rangle\langle\Phi_2|_{QAB}\right] \\ &= |\varphi_\mu\rangle\langle\varphi_\mu|_B. \end{aligned} \tag{4.13}$$

☞ We have now to stop for a while and clearly understand what is going on. In fact, it seems that, *exactly at the same moment* in which Alice reads the outcome ℓ_μ on her measurement apparatus, Bob's state is changed from $\mathrm{Tr}_{QA}[|\Phi_2\rangle\langle\Phi_2|_{QAB}] = \frac{\mathbb{1}_2}{2}$ (i.e. a mixed state), to $|\varphi_\mu\rangle\langle\varphi_\mu|_B$ (i.e. a pure state). It is like if Alice's decision of performing a measurement on her side can *instantaneously* influence the system at Bob's side! Can this be true? Of course not! Why?

The solution to the above question is a consequence of the fact that Bob does *not* know which outcome Alice obtained, before she tells him. This is due to the fact that, as we stressed already many times, the outcomes of a measurement are random, and cannot be computed in advance. So, until Bob does not receive from Alice the information about which outcome she obtained, it is correct to say that the quantum system B is either in state $|\varphi_{00}\rangle\langle\varphi_{00}|$, or $|\varphi_{01}\rangle\langle\varphi_{01}|$, or $|\varphi_{10}\rangle\langle\varphi_{10}|$, or $|\varphi_{11}\rangle\langle\varphi_{11}|_B$, each of them occurring with probability $p = 1/4$. In other words, the fact that Alice is performing a measurement on her share, makes the quantum system in Bob's hands to be a *random sample* (see Theorem 2.2) from the ensemble $\left(\left\{p_\mu = \frac{1}{4}\right\}, \{|\varphi_\mu\rangle\langle\varphi_\mu|\}\right)$. Then, Theorem 2.2 states that the quantum system at Bob's side, before Bob receives the information about Alice's outcome, is correctly described by the state

$$\begin{aligned}
\overline{\omega}_B &= \frac{1}{4}\sum_\mu |\varphi_\mu\rangle\langle\varphi_\mu|_B \\
&= \sum_\mu \mathrm{Tr}_{QA}\left[(\Pi^O(\ell_\mu)\otimes \mathbb{1}_2) \quad |\Phi_2\rangle\langle\Phi_2|_{QAB}\right] \\
&= \mathrm{Tr}_{QA}\left\{\left(\underbrace{\left[\sum_\mu \Pi^O(\ell_\mu)\right]}_{\mathbb{1}_4}\otimes \mathbb{1}_2\right) \quad |\Phi_2\rangle\langle\Phi_2|_{QAB}\right\} \\
&= \mathrm{Tr}_{QA}\left[(\mathbb{1}_4\otimes\mathbb{1}_2) \quad |\Phi_2\rangle\langle\Phi_2|_{QAB}\right] \\
&= \mathrm{Tr}_{QA}\left[\,|\Phi_2\rangle\langle\Phi_2|_{QAB}\,\right] \\
&= \frac{\mathbb{1}_2}{2}.
\end{aligned} \tag{4.14}$$

As the above calculation shows, the state of Bob's system is *not* changed simply by the fact that Alice is performing a measurement! Only if Alice communicates to Bob which outcome ℓ_μ she obtained, only then, after Bob *received* such information, he can correctly say that the state of quantum system B is $|\varphi_\mu\rangle\langle\varphi_\mu|_B$.

Question 4.1. In our story, how much information (measured in bits) must Alice send to Bob in order to communicate her measurement outcome?

So, let us suppose that Alice indeed communicates to Bob her outcome: she can do this, since Alice and Bob are allowed to sends email to each other. Now, to be practical, let Alice's outcome be, for example, ℓ_{10}. The corresponding state of Bob, after he knows the outcome value, is $|\varphi_{10}\rangle\langle\varphi_{10}|$, where $|\varphi_{10}\rangle = \begin{pmatrix} c_0 \\ -c_1 \end{pmatrix}$. We recall that the unknown state that Alice wants to send to Bob is $|\psi\rangle\langle\psi|$, with $|\psi\rangle = \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}$. So, what Bob has to do (when he knows that Alice's outcome was ℓ_{10}) is to apply on his share B the unitary operator represented by the matrix

$$W_{10} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{4.15}$$

so that

$$W_{10}|\varphi_{10}\rangle\langle\varphi_{10}|W_{10}^\dagger = |\psi\rangle\langle\psi|, \tag{4.16}$$

as required to accomplish the mission! In the general case in which Alice obtained the outcome ℓ_μ and Bob's quantum system is in state $|\varphi_\mu\rangle\langle\varphi_\mu|$, it is very easy to show that there always exists a unitary operator W_μ that Bob can apply so that the final state is $|\psi\rangle\langle\psi|$.

☞ Write the matrices representing the unitary operators W_μ, for all $\mu \in \{00, 01, 10, 11\}$.

Hence, quantum teleportation is possible! Summarizing, for the success of the protocol, we required two assumptions:

1. Alice and Bob must have prepared and share the entangled state $|z\rangle\langle z|$ in advance (i.e. before the protocol starts);
2. Alice must communicate to Bob the outcome of her measurement, and Bob has to apply an operation conditionally on the information he receives.

☞ Since quantum teleportation needs classical information to be transmitted (condition 2 above), and information can travel at most at the speed of light, this fact limits the speed of quantum teleportation to be less than (or at most equal to) the speed of light.

Question 4.2. We considered only the case in which Alice is given a *pure* state. Is it difficult to extend the previous discussion so to consider the general case of *mixed* states?

4.2 Quantum super-dense coding

In the previous section we studied how pre-shared entanglement, in the form of the density matrix $|z\rangle\langle z|$ in Eq. (2.17), enables two parties to transfer the state of a two-dimensional quantum system by sending two bits of classical information. The "resource balance" of quantum teleportation is the following:

pre-shared entanglement + transmission of 2 classical bits →
transmission of 1 qubit

We will now see that also the converse is true: pre-shared entanglement enables two parties to transfer two bits of classical information by sending only a two-dimensional quantum system. The protocol achieving this task is called "super-dense coding" and its resource balance is the following:

pre-shared entanglement + transmission of 1 qubits →
transmission of 2 classical bits

The task is the following: Alice and Bob (the same agents as before!) are far apart when, at time $t = t_0$, Alice is given two bits of classical information (i.e. an integer from 0 to 3, or, in binary notation, one pair among 00,01,10,11) and she is requested to communicate these to Bob by sending *only one* qubit. This means that Bob, receiving only one qubit from Alice, must be able to *perfectly* recover which one among the four alternatives $\{00, 01, 10, 11\}$ Alice was given at $t = t_0$.

First idea: Alice tries to carefully encode each alternative $ij \in \{00, 01, 10, 11\}$ on suitably chosen density matrices $\rho^{ij} \in \mathbb{M}(\mathbb{C}^2)$, prepare a qubit in the state ρ^{ij} corresponding to the message she receives at $t = t_0$, and send this to Bob. Bob, in order to read the message "ij", has to be able to *perfectly* distinguish among the four density matrices $\{\rho_A^{00}, \rho_A^{01}, \rho_A^{10}, \rho_A^{11}\}$ chosen by Alice. But, as we learned in Example 2.3, this cannot be done: the quantum system A is a qubit, i.e. a two-dimensional quantum system, for which *at most two* perfectly distinguishable states exist at a time.

A solution for Alice and Bob exists, and it is called super-dense coding. The protocol of super-dense coding works as follows: as we did already in the case of quantum teleportation, we imagine that, at a previous time $t_{-1} < t_0$, Alice and Bob meet and share a bipartite quantum system AB, with Hilbert space $\mathcal{H}_{AB} \cong \mathbb{C}^2 \otimes \mathbb{C}^2$, prepared in the pure state $|z\rangle\langle z|_{AB}$ of equation (2.17). (Notice that we already exploited the same state in the protocol of quantum teleportation!) After the system AB has been prepared in the state $|z\rangle\langle z|_{AB}$, Alice keeps the A subsystem, leaving the B subsystem to Bob.

Fast forward: we are now at time $t = t_0$, when Alice receives a letter $\mu \in \{00, 01, 10, 11\}$. Depending on the value of μ, Alice applies on her share A one of the unitary operators $U_\mu \in \mathbb{M}(\mathbb{C}^2)$ defined as follows:

$$U_{00} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; U_{01} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \; U_{10} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; U_{11} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{4.17}$$

After Alice applied the correct unitary operator, the state of the composite system AB is given by

$$|\Psi_\mu\rangle\langle\Psi_\mu|_{AB} = (U_\mu \otimes \mathbb{1}_B)\, |z\rangle\langle z|_{AB} \left(U_\mu^\dagger \otimes \mathbb{1}_B\right), \tag{4.18}$$

where the vectors $|\Psi_\mu\rangle \in \mathbb{C}^4$ are as follows:

$$|\Psi_{00}\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}, |\Psi_{01}\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\\-1\end{pmatrix}, |\Psi_{10}\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\1\\0\end{pmatrix}, |\Psi_{11}\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\-1\\1\\0\end{pmatrix}. \tag{4.19}$$

We recall that Alice is allowed to send one qubit to Bob. She then sends subsystem A, so that Bob holds the whole composite system AB (a four-dimensional quantum system), which at this point is in state $|\Psi_\mu\rangle\langle\Psi_\mu|_{AB}$, depending on the message given to Alice. How can Bob learn the value of μ?

Exercise 4.1. Prove that $\langle\Psi_\mu|\Psi_{\mu'}\rangle = \delta_{\mu,\mu'}$.

After Bob received the qubit A from Alice, he applies on AB the unitary matrix

$$W = \frac{1}{\sqrt{2}}\begin{pmatrix}1&0&0&1\\1&0&0&-1\\0&1&1&0\\0&1&-1&0\end{pmatrix}. \tag{4.20}$$

The action of W on the four vectors $|\Psi_\mu\rangle$ is the following:

$$W|\Psi_{00}\rangle = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, W|\Psi_{01}\rangle = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, W|\Psi_{10}\rangle = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, W|\Psi_{11}\rangle \begin{pmatrix}0\\0\\0\\1\end{pmatrix}, \tag{4.21}$$

i.e. the four vectors $|\Psi_\mu\rangle$ are transformed into the elements $|e_\mu\rangle$ of the standard basis of $\mathbb{C}^4$. After this transformation, Bob measures an observable O_{AB} represented by the self-adjoint matrix

$$O_{AB} = \begin{pmatrix}\ell_{00}&0&0&0\\0&\ell_{01}&0&0\\0&0&\ell_{10}&0\\0&0&0&\ell_{11}\end{pmatrix}, \tag{4.22}$$

with $\ell_\mu \in \mathbb{R}$. The only condition we impose on O_{AB} is that $\ell_{00} \neq \ell_{01} \neq \ell_{10} \neq \ell_{11}$, so that the spectral projectors $E^O(\ell_\mu)$ appearing in equation (1.13) are equal to $|e_\mu\rangle\langle e_\mu|$, for all $\mu \in \{00, 01, 10, 11\}$. (Notice that the same observable appeared also in the teleportation protocol!)

According to Postulate 3, if the message Alice transmitted was μ, Bob will obtain the outcome $\ell_{\mu'}$ with probability

$$\begin{aligned}\Pr\{O_{AB} = \ell_{\mu'}\} &= \mathrm{Tr}\left[E^O(\ell_{\mu'})\,|e_\mu\rangle\langle e_\mu|_{AB}\right]\\ &= \delta_{\mu,\mu'},\end{aligned} \tag{4.23}$$

i.e. if Bob obtains the outcome μ, he is *sure* that the message Alice wanted to transmit is μ. Again, mission accomplished!

Summarizing, the protocol we described above, called super-dense coding (see also Figure 11), is closely related to quantum teleportation—it is, in a sense, its *reverse*:

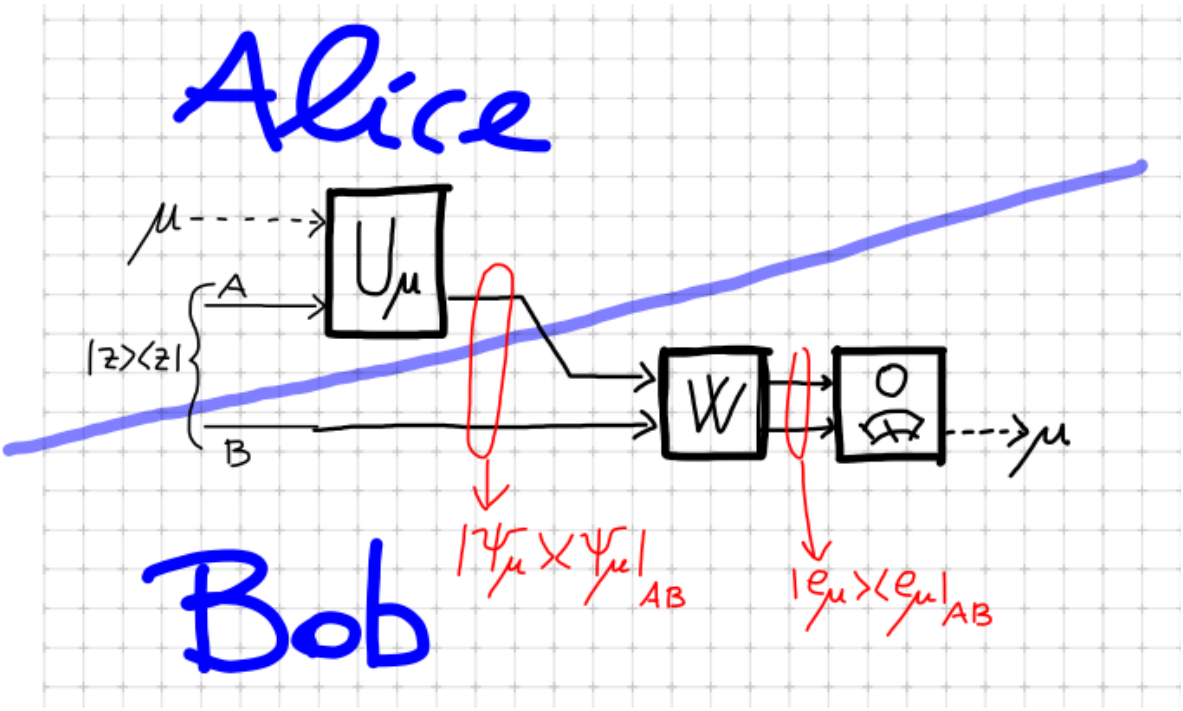

Figure 11: The circuit that implements super-dense coding from Alice to Bob.

- by quantum teleportation, Alice is able to transfer one qubit to Bob, by sending him only two bits of classical information. This is possible if Alice and Bob share in advance a suitable entangled pure state $|z\rangle\langle z|_{AB}$.
- by super-dense coding, Alice is able to communicate two bits of classical information to Bob, by sending him only one qubit. This is possible if Alice and Bob share in advance a suitable entangled pure state $|z\rangle\langle z|_{AB}$.

4.3 Optimality of quantum teleportation and quantum super-dense coding

proof by nesting