

Quantum mechanics and quantum information theory

Guillaume Aubrun

Chapter 1

The formalism of quantum mechanics

1.1 Mathematical preliminaries

Bra-ket notation

In all the notes, we denote by $\mathcal{H}$, $\mathcal{H}_1$, $\mathcal{H}_2$ complex Hilbert spaces of finite dimension. Many of the concepts we introduce make sense in infinite-dimensional spaces, but we choose to work in the simpler finite-dimensional setting

The inner product on $\mathcal{H}$ is denoted as $\langle \cdot | \cdot \rangle$ (or $\langle \cdot, \cdot \rangle$). Following the physics convention, it is anti-linear in the first variable and linear on the second variable:

$$\langle \lambda x | y \rangle = \bar{\lambda} \langle x | y \rangle, \quad \langle x | \lambda y \rangle = \lambda \langle x | y \rangle \quad \forall x, y \in \mathcal{H}, \forall \lambda \in \mathbb{C}.$$

Elements of $\mathcal{H}$ are denoted $|x\rangle$ (“ket”). The linear form given by $|y\rangle \mapsto \langle x | y \rangle$ is denoted $\langle x |$ (“bra”). The map

$$\mathcal{H} \rightarrow \mathcal{H}^*, \quad |x\rangle \mapsto \langle x|$$

is anti-linear: $\langle \lambda x | = \bar{\lambda} \langle x |$. The Hilbert space norm is $\|x\| = \sqrt{\langle x | x \rangle}$.

Exercise 1. Check the following formulas

$$|x\rangle\langle y| \cdot |z\rangle\langle t| = \langle y | z \rangle |x\rangle\langle t|, \quad |x\rangle\langle y|^* = |y\rangle\langle x|, \quad \text{Tr}|x\rangle\langle y| = \langle y | x \rangle.$$

Hermitian and positive operators

We denote by $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ denotes the linear operators from $\mathcal{H}_1$ to $\mathcal{H}_2$, and $\mathcal{B}(\mathcal{H}) = \mathcal{B}(\mathcal{H}, \mathcal{H})$. We also consider

$$\mathcal{B}(\mathcal{H})_{sa} = \{A \in \mathcal{B}(\mathcal{H}) : A^* = A\}$$

which is a real vector space.

Recal the *spectral theorem*: any self-adjoint operator $A \in \mathcal{B}(\mathcal{H})_{sa}$ can be written as

$$A = \sum_i \lambda_i |x_i\rangle\langle x_i|, \quad \lambda_i \in \mathbb{R},$$

where (x_i) is an orthonormal basis of $\mathcal{H}$. The λ_i are the eigenvalues, and $|x_i\rangle\langle x_i|$ is the rank 1 orthogonal projection onto $\mathbb{C}x_i$.

A self-adjoint operator A is said to be *positive* if its eigenvalues are ≥ 0 . We denote $\mathcal{B}(\mathcal{H})_+$ the set of positive operators. If A, B are self-adjoint, we write $A \leq B$ if $B - A$ is positive.

Exercise 2. Let $A \in \mathcal{B}(\mathcal{H})$ (not assumed to be self-adjoint). Then

$$A \text{ positive} \iff \langle x | Ax \rangle \text{ is real and } \geq 0 \quad \forall x \in \mathcal{H}.$$

[This is false over $\mathbb{R}$.]

Every positive operator has a positive square root $A^{1/2}$, defined by taking the square root of the eigenvalues:

$$A = \sum_i \lambda_i |u_i\rangle\langle u_i| \implies A^{1/2} := \sum_i \sqrt{\lambda_i} |u_i\rangle\langle u_i|, \quad (A^{1/2})^2 = A,$$

with $\lambda_i \geq 0$ and (u_i) an orthonormal basis. Actually $A^{1/2}$ is a polynomial in A , so every operator that commutes with A also commutes with $A^{1/2}$.

For every $A \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$, $A^*A \in \mathcal{B}(\mathcal{H}_1)_+$ is self-adjoint positive (because $\langle A^*Ax|x \rangle = \|Ax\|^2 \geq 0$), and we set

$$|A| = \sqrt{A^*A} \in \mathcal{B}(\mathcal{H}_1)_+.$$

Warning! In general

$$|A^*| \neq |A|, \quad |AB| \neq |A||B|, \quad |A+B| \not\leq |A| + |B|.$$

(The first two expressions do not even make sense for operators between different Hilbert spaces.)

Partial isometries, polar decomposition, singular value decomposition

An operator $U : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ is a *partial isometry* if there is a subspace $E \subset \mathcal{H}_1$ such that

$$\|Ux\| = \|x\| \quad \forall x \in E, \quad \|Ux\| = 0 \quad \forall x \in E^\perp.$$

E is the *initial space* of U and $F = \text{Ran}(U|_E) = \text{Ran } U$ is the *final space*.

Proposition 1. $U : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ is a partial isometry $\iff U^*U$ is an orthogonal projection (onto E) $\iff UU^*$ is an orthogonal projection (onto F).

We denote by $\mathcal{U}(\mathcal{H})$ the group of unitary operators on $\mathcal{H}$ (i.e., isometries, or operators $U \in \mathcal{B}(\mathcal{H})$ such that $U^*U = \text{Id}_{\mathcal{H}}$)

Theorem 2 (Polar decomposition). *For every $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ there exists a partial isometry $U : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ such that*

$$T = U|T|$$

We can choose U to have initial space $\text{Ran}|T|$ and final space $\text{Ran } T$; we also have $|T| = U^*T$.

Proof. Since for all $x, y \in \mathcal{H}_1$,

$$\langle |T|x, |T|y \rangle = \langle |T|^2x, y \rangle = \langle T^*Tx, y \rangle = \langle Tx, Ty \rangle,$$

we may define

$$U : \text{Ran}|T| \rightarrow \text{Ran } T, \quad |T|x \mapsto Tx.$$

One checks that the map U is well-defined and that $\|Uy\| = \|y\|$ for all $y \in \text{Ran}|T|$. We extend U to a map on $\mathcal{H}_1$ by setting $U = 0$ on $(\text{Ran}|T|)^\perp$. We have $U|T| = T$. $\square$

Theorem 3 (Singular value decomposition). *Any $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ can be written as*

$$T = \sum_i s_i |u_i\rangle\langle v_i|,$$

where (u_i) is an orthonormal family in $\mathcal{H}_2$, (v_i) is an orthonormal family in $\mathcal{H}_1$, (s_i) are positive numbers such that $s_1 \geq \dots \geq s_d > 0$ and $d = \text{rk}(T)$. We also have $T^* = \sum_i s_i |v_i\rangle\langle u_i|$.

The numbers $s_1 \geq \dots \geq s_d$ are called the *singular values* of A . A variant of the formulation is that every $T \in \mathcal{B}(\mathcal{H})$ can be written as

$$T = \sum_i s_i |u_i\rangle\langle v_i|$$

with $s_i \geq 0$ and $(u_i), (v_i)$ are orthonormal **bases** of $\mathcal{H}$.

Proof. By the spectral theorem applied to $|T|$ (note that $\text{rk } T = \text{rk } |T|$) after discarding eigenvalues equal to zero, we obtain an family (v_i) orthonormal basis of $\text{Ran } |T|$ and numbers $s_1 \geq \dots \geq s_d > 0$ such that

$$|T| = \sum_i s_i |v_i\rangle\langle v_i|$$

By the polar decomposition, there is a partial isometry U with initial space $\text{Ran } |T|$ such that $T = U|T|$. We then have

$$T = U|T| = \sum_i s_i |Uv_i\rangle\langle v_i|. \quad \square$$

If we denote $u_i = Uv_i$, the result follows by observing that $\langle u_i, u_j \rangle = \langle U^*Uv_i, v_j \rangle = \langle v_i, v_j \rangle$ since U^*U is the projection onto $\text{Ran } |T|$.

In matrix form, the singular values decomposition states that $A \in M_{d_1, d_2}$ can be written $A = U_1 \Delta U_2$ with $U_1 \in \mathcal{U}(d_1)$, $U_2 \in \mathcal{U}(d_2)$ and $\Delta_{ij} = 0$ if $i \neq j$.

Norms of operators

We introduce three important norms on operators. Given a linear map $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ with singular values $s_1(A) \geq \dots s_d(A) > 0$, we define

- **The operator norm** $\|A\|_{op}$

$$\|A\|_{op} = \sup_{\substack{x \in \mathcal{H}_1 \\ \|x\| \leq 1}} \|Ax\| = \sup_{\substack{x \in \mathcal{H}_1 \\ \|x\| \leq 1}} \sup_{\substack{y \in \mathcal{H}_2 \\ \|y\| \leq 1}} |\langle y | Ax \rangle|, \quad \|A\| = \|A^*\|.$$

We have $\|UAV\|_{op} = \|A\|_{op}$ if $U \in \mathcal{U}(\mathcal{H}_2)$ and $V \in \mathcal{U}(\mathcal{H}_1)$, and thus $\|A\|_{op} = s_1(A)$ (largest singular value).

- **The Hilbert–Schmidt norm** $\|A\|_{HS}$, which is the norm of A seen as a vector of dimension $\dim \mathcal{H}_1 \times \dim \mathcal{H}_2$. Given orthonormal bases (u_i) of $\mathcal{H}_1$ and (v_j) of $\mathcal{H}_2$, we define

$$\begin{aligned} \|A\|_{HS} &= \left(\sum_i \|Au_i\|^2 \right)^{1/2} = \left(\sum_{i,j} |\langle Au_i, v_j \rangle|^2 \right)^{1/2} \\ &= \text{Tr}(A^*A)^{1/2} = \text{Tr}(|A|^2)^{1/2} = \left(\sum_{i=1}^d s_i(A)^2 \right)^{1/2} \end{aligned}$$

We also have $\|UAV\|_{HS} = \|A\|_{HS}$ if $U \in \mathcal{U}(\mathcal{H}_2)$ and $V \in \mathcal{U}(\mathcal{H}_1)$. Other simple to verify inequalities are, for $B \in \mathcal{B}(\mathcal{H}_2, \mathcal{H}_3)$

$$\|AB\|_{HS} \leq \|A\|_{HS} \|B\|_{op}, \quad \|AB\|_{HS} \leq \|A\|_{op} \|B\|_{HS}$$

- **The trace norm** $\|A\|_1$

$$\|A\|_1 = \text{Tr}|A| = \sum_{i=1}^d s_i(A).$$

As before, we have $\|UAV\|_1 = \|A\|_1$ if $U \in \mathcal{U}(\mathcal{H}_2)$ and $V \in \mathcal{U}(\mathcal{H}_1)$.

The triangle inequality is not obvious!

Proposition 4. *For every operators $A, B \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$, we have*

$$\mathrm{Tr}|A + B| \leq \mathrm{Tr}|A| + \mathrm{Tr}|B|.$$

Proof. Introduce the polar decompositions $A = U|A|$, $B = V|B|$, $A + B = W|A + B|$. For an orthonormal basis (u_i) of $\mathcal{H}_1$, write

$$\mathrm{Tr}|A + B| = \sum_i \langle u_i | W^* (A + B) u_i \rangle = \sum_i \langle u_i | W^* U |A| u_i \rangle + \sum_i \langle u_i | W^* V |B| u_i \rangle.$$

Now

$$|\langle u_i | W^* U |A| u_i \rangle| = |\langle |A|^{1/2} U^* W u_i | |A|^{1/2} u_i \rangle| \leq \| |A|^{1/2} U^* W u_i \| \| |A|^{1/2} u_i \|,$$

so by Cauchy-Schwarz

$$\begin{aligned} \sum_i |\langle e_i | W^* U |A| e_i \rangle| &\leq \left(\sum_i \| |A|^{1/2} U^* W e_i \|^2 \right)^{1/2} \left(\sum_i \| |A|^{1/2} e_i \|^2 \right)^{1/2} \\ &\leq \| |A|^{1/2} U^* W \|_{HS} \| |A|^{1/2} \|_{HS} \\ &\leq \| |A|^{1/2} \|_{HS} \| |A|^{1/2} \|_{HS} = \mathrm{Tr}|A|. \end{aligned}$$

Using a similar computation for the B term, we obtain $\mathrm{Tr}|A + B| \leq \mathrm{Tr}|A| + \mathrm{Tr}|B|$. $\square$

Another useful inequality is

Proposition 5. *For every operators $A, B \in \mathcal{B}(\mathcal{H})$*

$$\|AB\|_1 \leq \|A\|_1 \|B\|_\infty.$$

Proof. We claim that the unit ball for the operator norm in $\mathcal{B}(\mathcal{H})$ is the convex hull of $\mathcal{U}(\mathcal{H})$. Since $\|AU\|_1 = \|A\|_1$ when $U \in \mathcal{U}(\mathcal{H})$, the proposition then follows from the triangle inequality for the trace norm.

We now prove the claim. Clearly, every $B \in \mathrm{conv} \mathcal{U}(\mathcal{H})$ satisfies $\|B\|_{op} \leq 1$ (since $U \in \mathcal{U}(\mathcal{H}) \Rightarrow \|U\|_{op} \leq 1$). For the converse, we use the fact that every number $s \in [0, 1]$ can be written as $s = \frac{\alpha + \beta}{2}$ with $\alpha, \beta \in \mathbb{C}$, $|\alpha| = |\beta| = 1$. The singular value decomposition of an operator B with $\|B\|_{op} \leq 1$ reads as $B = \sum_i s_i |u_i\rangle \langle v_i|$ for $0 \leq s_i \leq 1$ and $(u_i), (v_i)$ orthonormal bases. If we introduce complex numbers α_i and β_i of modulus one such that $s_i = \frac{\alpha_i + \beta_i}{2}$, then

$$B = \frac{1}{2} \left[\underbrace{\sum_i \alpha_i |u_i\rangle \langle v_i|}_{\in \mathcal{U}(\mathcal{H})} + \underbrace{\sum_i \beta_i |u_i\rangle \langle v_i|}_{\in \mathcal{U}(\mathcal{H})} \right],$$

proving our claim. $\square$

Exercise 3. If we identify $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)^*$ with $\mathcal{B}(\mathcal{H}_2, \mathcal{H}_1)$ via

$$A \mapsto (B \mapsto \mathrm{Tr}(AB)),$$

then the norms $\|\cdot\|_1$ and $\|\cdot\|_{op}$ are dual to each other.

Tensor products of vector and Hilbert spaces

We introduce tensor products in the most elementary way, using expansion of vectors into fixed bases. Given vector spaces V_1 and V_2 (on an arbitrary field $\mathbb{K}$) with bases $(u_k)_k$ and $(v_l)_l$, we define

$$V_1 \otimes V_2 := \left\{ \sum_{k,l} \alpha_{k,l} u_k \otimes v_l : \alpha_{k,l} \in \mathbb{K} \right\},$$

which is a vector space with a basis whose elements are labeled as $(u_k \otimes v_l)_{(k,l)}$.

Given elements $x \in V_1$ and $y \in V_2$, written as

$$x = \sum_k \lambda_k u_k \quad y = \sum_l \mu_l v_l$$

we then define

$$x \otimes y := \sum_{k,l} \lambda_k \mu_l \cdot u_k \otimes v_l \in V_1 \otimes V_2.$$

Observe that the map $t : (x, y) \mapsto x \otimes y$ defined this way is a bilinear map from $V_1 \times V_2$ to $V_1 \otimes V_2$. If W is any vector space and $b : V_1 \times V_2 \rightarrow W$ is a linear map, there is a unique linear map $\tilde{b} : V_1 \otimes V_2 \rightarrow W$ such that $b = \tilde{b}t$: set $\tilde{b}(u_k \otimes v_l) = b(u_k, v_l)$ and extend $\tilde{b}$ to $V_1 \otimes V_2$ by linearity. This «universal property» of the tensor product can be used to define it in a more abstract way, without relying on choices of bases. It is routine to show that the construction we introduce does not depend on the choice of bases (i.e., other choices of bases lead to other vector spaces which are canonically isomorphic to our definition of $V_1 \otimes V_2$).

Elements of $V_1 \otimes V_2$ which have the form $x \otimes y$ for $x \in V_1$ and $y \in V_2$ are called *pure tensors* or *elementary tensors*.

Assume now that H_1 and H_2 are (finite-dimensional complex) vector spaces and that $(u_k)_k$ and $(v_l)_l$ are orthonormal bases. We can then equip $H_1 \otimes H_2$ with a Hilbert space structure such that $(u_k \otimes v_l)_{(k,l)}$ is an orthonormal basis. More precisely, we define an inner product on $H_1 \otimes H_2$ by the formula

$$\langle x_1 \otimes x_2, y_1 \otimes y_2 \rangle := \langle x_1, y_1 \rangle \cdot \langle x_2, y_2 \rangle$$

(extended to a map on $(H_1 \otimes H_2) \times (H_1 \otimes H_2)$ by linearity and anti-linearity in each separate variable); it is routine to verify that this indeed defines an inner product.

All the considerations above can be readily extended to define the Hilbert space tensor product $H_1 \otimes \cdots \otimes H_k$ of finitely many Hilbert spaces $H_1, \dots, H_k$.

Schmidt decomposition

If $\mathcal{H}$ is a Hilbert space, the canonical map $|x\rangle \mapsto \langle x|$ from $\mathcal{H}$ to $\mathcal{H}^*$ is anti-linear. Once we choose an orthogonal basis (u_i) , we could define a linear isometry from $\mathcal{H}$ to $\mathcal{H}^*$ by $\sum \lambda_i |u_i\rangle \mapsto \sum \lambda_i \langle u_i|$, but such an isomorphism is not canonical and depends on a choice of basis.

Another possibility is to define a new Hilbert space structure on $\mathcal{H}$ by using the same vector addition, but defining a new scalar multiplication as $\lambda \cdot x = \bar{\lambda}x$ (the norm being unchanged). If we denote by $\overline{\mathcal{H}}$ the Hilbert space obtained this way, then $\mathcal{H}^*$ and $\overline{\mathcal{H}}$ are canonically isomorphic.

There is a canonical identification

$$\mathcal{H}_1 \otimes \mathcal{H}_2 \longrightarrow \mathcal{B}(\mathcal{H}_1^*, \mathcal{H}_2), \quad |x_1\rangle \otimes |x_2\rangle \longmapsto |x_2\rangle \langle x_1|$$

This map is an isometry from $(\mathcal{H}_1 \otimes \mathcal{H}_2, \|\cdot\|)$ to $(\mathcal{B}(\mathcal{H}_1^*, \mathcal{H}_2), \|\cdot\|_{HS})$ and allows to translate statements about operators to statements about tensors.

For example, here is the tensor version of the singular value decomposition, which is called the Schmidt decomposition.

Proposition 6. Every element $z \in \mathcal{H}_1 \otimes \mathcal{H}_2$ can be written as

$$z = \sum_i s_i x_i \otimes y_i$$

where (s_i) are positive numbers, (x_i) is an orthonormal family in $\mathcal{H}_1$ and (y_i) is an orthonormal family in $\mathcal{H}_2$. The numbers s_i are called the Schmidt coefficients of z .

Tensor products of operators

If $T_1 : V_1 \rightarrow W_1$ and $T_2 : V_2 \rightarrow W_2$ (V_i, W_i vector spaces), we define $T_1 \otimes T_2$ by

$$(T_1 \otimes T_2)(v_1 \otimes v_2) = T_1 v_1 \otimes T_2 v_2$$

and then extend to a linear map. Alternatively, one could use the universal property, since $(v_1, v_2) \mapsto T_1 v_1 \otimes T_2 v_2$ is bilinear..

Suppose now that T_1 and T_2 are operators between Hilbert spaces. If we have singular value decompositions

$$T_1 = \sum_i s_i |u_i\rangle\langle v_i|, \quad T_2 = \sum_j t_j |w_j\rangle\langle x_j|,$$

then

$$T_1 \otimes T_2 = \sum_{i,j} s_i t_j |u_i \otimes w_j\rangle\langle v_i \otimes x_j|$$

is a singular value decomposition for $T_1 \otimes T_2$. In particular, we obtain the formulas

$$\|T_1 \otimes T_2\|_{op} = \|T_1\|_{op} \|T_2\|_{op}, \quad \|T_1 \otimes T_2\|_1 = \|T_1\|_1 \|T_2\|_1, \quad \|T_1 \otimes T_2\|_{HS} = \|T_1\|_{HS} \|T_2\|_{HS}.$$

1.2 The postulates of quantum mechanics for closed systems and pure states

Postulate 1. A quantum system is modeled by a complex Hilbert space $\mathcal{H}$. A pure state is a unit vector in $\mathcal{H}$.

If $\mathcal{H} = \mathbb{C}^2$, such a system is a “qubit”, a quantum bit. It could be the spin of an electron, or the polarization of a photon. We don’t care, like in a classical computer we don’t care how the bit is implemented.

Postulate 2. An observable is modeled by a self-adjoint operator A on $\mathcal{H}$. The possible outcomes of the observation are the eigenvalues of A .

If $\sigma(A)$ is the spectrum of A , we have

$$A = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda, \tag{1.1}$$

where P_λ is the orthogonal projection onto the eigenspace of eigenvalue λ .

Postulate 3. If we observe a quantum state in state ψ using an observable A given by (1.1), the outcome is a random variable X such that

$$\mathbb{P}(X = \lambda) = \|P_\lambda \psi\|^2 \tag{1.2}$$

Moreover, after the outcome λ has been observed, the post-measurement state of the system is

$$\frac{P_\lambda \psi}{\|P_\lambda \psi\|}.$$

By Pythagoras theorem, we have $\sum_{\lambda} \|\mathbb{P}_{\lambda}\psi\|^2 = \|\psi\|^2 = 1$, so the formula (1.2) defines a probability distribution.

The description of a quantum systems comes with a distinguished observable called the Hamiltonian $H(t)$, which may depend on the time $t \geq 0$. Observing the Hamiltonian gives the energy of the system.

Postulate 4 (Schrödinger equation). The state at time $t \geq 0$ of a system governed by a Hamiltonian $H(t)$ in initial pure state ψ_0 is the solution of the differential equation

$$\psi'_t = -iH(t)\psi_t$$

If $H(t) = H$ does not depend on time, the solution to this equation is given by

$$\psi_t = U(t)\psi_0$$

with $U(t) = \exp(-itH)$ is a unitary transformation on $\mathcal{H}$.

Postulate 5. If two disjoint quantum systems are modeled by Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, the joint system is modeled by the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. If each systems are described by unit vectors $\psi_A \in \mathcal{H}_A$ and $\psi_B \in \mathcal{H}_B$, without interaction, the pure state of the joint system is $\psi_A \otimes \psi_B$.

We now describe the limitations of this formalism. Observing the system modeled by $\mathcal{H}_A$ using the observable M in state ψ_A is equivalent to an observing the joint system model by $\mathcal{H}_A \otimes \mathcal{H}_B$ using the observable $M \otimes \text{id}_B$ in state $\psi_A \otimes \psi_B$. By these we mean that the statistics of the observation outcomes are the same, and do not depend on ψ_B .

However, is the joint system is in a pure state $z \in \mathcal{H}_A \otimes \mathcal{H}_B$ which is not of the form $\psi_A \otimes \psi_B$, there is no pure state on $\mathcal{H}_A$ which gives the right statistics for every operation! This suggest that our formalism is not complete since it does not allow to describe the state a subsystem—a major issue !

Another minor issue is that there is no reason to distinguish the states ψ and $-\psi$ (or more generally $e^{i\alpha}\psi$ for α real), since the statistic outcomes for any observation are the same for both states. This suggest that the set of possible states of a system is not the unit sphere of the Hilbert space, but rather the associated projective space.

An elegant way to solve this minor issue is to replace a unit vector ψ by the corresponding rank 1 projection $|\psi\rangle\langle\psi|$ as our notion of pure state. Indeed, we have $|\psi\rangle\langle\psi| = |\phi\rangle\langle\phi|$ if and only if the unit vectors ϕ, ψ satisfy $\phi = e^{i\alpha}\psi$ for some $\alpha \in \mathbb{R}$. As we will see, this also solves the major issue.

1.3 The postulates of quantum mechanics for open systems and mixed states

Let us modify our notion of state.

Postulate 1'. A quantum system is modeled by a complex Hilbert space $\mathcal{H}$. A mixed state is a self-adjoint operator on $\mathcal{H}$ with trace 1.

It is common to denote mixed states with the letter ρ or σ . We recover the notion of pure states from Postulate 1 by declaring that we call pure states on $\mathcal{H}$ a mixed state which has the form $|\psi\rangle\langle\psi|$ for a unit vector $\psi \in \mathcal{H}$.

The mixed state formalism also describes in a natural way probabilistic scenarios. A situation such as “I toss a fair coin, if it turns Head I prepare my system in state ρ and if it turns Tails I prepare my system in state σ ” is equivalent, in terms of measurement statistics, to having a system which in state $\frac{1}{2}(\rho + \sigma)$.

By the spectral theorem, any mixed state ρ of a n -dimensional Hilbert space can be written as

$$\rho = \sum_{i=1}^n p_i |\psi_i\rangle\langle\psi_i|$$

where (ψ_i) is an orthonormal basis and (p_i) are numbers such that $p_i \geq 0$ and $p_1 + \dots + p_n = 1$, which can be interpreted as a probability density. Thus, any state can be identified with a probabilistic mixture of pure states.

We can also reformulate this fact by saying that the set of mixed states is the convex hull of the set of pure states.

We now define how the other postulates have to be modified at the level of mixed states. Postulate 2 is unchanged.

Postulate 3'. If we observe a quantum state in mixed state ρ using an observable A given by (1.1), the outcome is a random variable X such that

$$\mathbb{P}(X = \lambda) = \text{Tr}(P_\lambda \rho)$$

Moreover, after the outcome λ has been observed, the post-measurement state of the system is

$$\frac{P_\lambda \rho P_\lambda}{\text{Tr}(P_\lambda \rho)}.$$

Observe that if the mixed state ρ is a pure state $|\psi\rangle\langle\psi|$, this is consistent with Postulate 3.

Postulate 4'. The state at time $t \geq 0$ of a system governed by a Hamiltonian $H(t)$ in initial mixed state ρ_0 is the solution of the differential equation

$$\rho'_t = -i[H(t), \rho_t]$$

Here, we denote by $[A, B] = AB - BA$ the commutator of the operators A and B . Again, this is consistent with Postulate 4: if ψ_t is the solution of the equation $\psi'_t = -iH(t)\psi_t$, the associated mixed state $\rho_t = |\psi_t\rangle\langle\psi_t|$ satisfies

$$\rho'_t = |\psi'_t\rangle\langle\psi_t| + |\psi_t\rangle\langle\psi'_t| = -iH(t)|\psi_t\rangle\langle\psi_t| + i|\psi_t\rangle\langle\psi_t|H(t)$$

In the case of a time-independent Hamiltonian $H(t) = H$, the solution to the equation is

$$\rho_t = U(t)\rho_0 U(t)^*$$

with $U(t) = \exp(-itH)$.

1.4 Partial trace and purification of mixed states

We now introduce the concept of partial trace. Given a bipartite system $\mathcal{H}_A \otimes \mathcal{H}_B$, we implicitly use the isomorphism between $\mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and $\mathcal{B}(\mathcal{H}_A) \otimes \mathcal{B}(\mathcal{H}_B)$ in the following. We define the partial trace over $\mathcal{H}_B$ as the operation

$$\text{Tr}_B : \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B) \rightarrow \mathcal{B}(\mathcal{H}_A)$$

equal to $\text{Tr}_B = \text{id}_{\mathcal{B}(\mathcal{H}_A)} \otimes \text{Tr}$. In this last formula, Tr denotes the linear form, as a map from $\mathcal{B}(\mathcal{H}_B)$ to $\mathbb{C}$.

Equivalently, the partial trace can be defined as

$$\text{Tr}_B(X \otimes Y) = \text{Tr}(Y) \cdot C$$

for every $X \in \mathcal{B}(\mathcal{H}_A)$ and $Y \in \mathcal{B}(\mathcal{H}_B)$, extended by linearity to $\mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$.

The partial trace satisfies the following property, which is reminiscent to properties of the conditional expectation in probability: for every $X, Y \in \mathcal{B}(\mathcal{H}_A)$ and $Z \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$,

$$\text{Tr}[(X \otimes \text{id}_{\mathcal{H}_B})Z(Y \otimes \text{id}_{\mathcal{H}_B})] = X \cdot \text{Tr}_B Z \cdot Y$$

This formula is simple to check if Z is of the form $Z_A \otimes Z_B$ and the general case follows by linearity.

The following proposition shows that every mixed state has appears the partial trace of a pure state. As a warm up, we consider the following computation. If a vector $z \in \mathcal{H}_A \otimes \mathcal{H}_B$ has a Schmidt decomposition

$$z = \sum s_i x_i \otimes y_i$$

with (y_i) being orthonormal, then since $\text{Tr}|y_i\rangle\langle y_j| = \delta_{ij}$

$$\text{Tr}_{\mathcal{H}'}|z\rangle\langle z| = \sum_{i,j} s_i s_j \text{Tr}_B(|x_i\rangle\langle x_j| \otimes |y_i\rangle\langle y_j|) = \sum_i s_i^2 |x_i\rangle\langle x_i| \quad (1.3)$$

This computation shows that the partial trace of a positive (resp., self-adjoint) operator on $\mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is a positive (resp., self-adjoint) operator on $\mathcal{B}(\mathcal{H}_A)$.

Proposition 7. *Let $\rho \in \mathcal{B}(\mathcal{H})$. The following are equivalent*

1. ρ is a mixed state (i.e., is positive and has trace 1)
2. there is a Hilbert space $\mathcal{H}'$ and a unit vector $\psi \in \mathcal{H} \otimes \mathcal{H}'$ such that

$$\rho = \text{Tr}_{\mathcal{H}'}|\psi\rangle\langle\psi|$$

Proof. The implication $2 \implies 1$ is easy. To show $1 \implies 2$, consider a spectral decomposition of ρ

$$\rho = \sum \lambda_i |x_i\rangle\langle x_i|$$

where (x_i) is an orthonormal family in $\mathcal{H}$ and λ_i are positive numbers such that $\sum \lambda_i = 1$. Consider the vector $z \in \mathcal{H} \otimes \mathcal{H}$ defined by

$$z = \sum \sqrt{\lambda_i} x_i \otimes x_i$$

This vector has norm 1 and the computation (1.3) shows that the partial trace of $|z\rangle\langle z|$ equals ρ . $\square$

Chapter 2

Completely positive maps and quantum channels

2.1 The Choi isomorphism

We use freely the following canonical isomorphisms

$$\mathcal{H}_A \otimes \mathcal{H}_B \simeq \mathcal{H}_B \otimes \mathcal{H}_A, \quad (\mathcal{H}_A \otimes \mathcal{H}_B) \otimes \mathcal{H}_C \simeq \mathcal{H}_A \otimes (\mathcal{H}_B \otimes \mathcal{H}_C), \quad (\mathcal{H}_A \otimes \mathcal{H}_B)^* \simeq \mathcal{H}_A^* \otimes \mathcal{H}_B^*$$

Another canonical isomorphism is given by the trace duality: the spaces $\mathcal{B}(\mathcal{H}_A, \mathcal{H}_B)^*$ and $\mathcal{B}(\mathcal{H}_B, \mathcal{H}_A)$ are isomorphic if we identify $S \in \mathcal{B}(\mathcal{H}_B, \mathcal{H}_A)$ with the linear form $T \mapsto \text{Tr}[ST]$, which is an element of $\mathcal{B}(\mathcal{H}_A, \mathcal{H}_B)^*$.

However, it turns out to be very useful to consider a *non-canonical* isomorphism between the spaces $\mathcal{L}(\mathcal{B}(\mathcal{H}_A), \mathcal{B}(\mathcal{H}_B))$ and $\mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$. (We denote by $\mathcal{L}(V, W)$ the space of linear maps between the vector spaces V and W .) This isomorphism is not canonical since it depends on a choice of basis.

Consider an orthonormal basis (e_i) in $\mathcal{H}_A$. We denote $E_{ij} = |e_i\rangle\langle e_j|$. The family (E_{ij}) is a basis of the vector space $\mathcal{B}(\mathcal{H}_A)$.

Given $\Phi \in \mathcal{L}(\mathcal{B}(\mathcal{H}_A), \mathcal{B}(\mathcal{H}_B))$, we define its *Choi matrix* $C_\Phi \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ as

$$C_\Phi = \sum_{i,j} E_{ij} \otimes \Phi(E_{ij})$$

It is easy to see that the map $\Phi \mapsto C_\Phi$ is an isomorphism. Another useful formula for C_Φ can be written using the “maximally entangled vector” $\chi \in \mathcal{H}_A \otimes \mathcal{H}_A$ defined as

$$\chi = \sum_i e_i \otimes e_i$$

We leave as an exercise the verification of the following formula

$$C_\Phi = (\text{id}_{\mathcal{H}_A} \otimes \Phi)(|\chi\rangle\langle\chi|) \tag{2.1}$$

The choice of a basis (e_i) in $\mathcal{H}_A$ also allows to identify the spaces $\mathcal{B}(\mathcal{H}_A, \mathcal{H}_B)$ and $\mathcal{H}_A \otimes \mathcal{H}_B$ (again, this isomorphism is not canonical). Given $X \in \mathcal{B}(\mathcal{H}_A, \mathcal{H}_B)$, we define

$$\text{vec}(X) = \sum_i e_i \otimes X(e_i) \in \mathcal{H}_A \otimes \mathcal{H}_B$$

Lemma 8. *Let $X, Y \in \mathcal{B}(\mathcal{H}_A, \mathcal{H}_B)$ and consider the linear map $\Phi : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$ given by $M \mapsto XMY^*$. The Choi matrix of Φ is $C_\Phi = |x\rangle\langle y|$, where $x, y \in \mathcal{H}_A \otimes \mathcal{H}_B$ are defined by $x = \text{vec}(X)$ and $y = \text{vec}(Y)$*

Proof. This is an straightforward computation

$$\begin{aligned}
C_\Phi &= \sum_{i,j} E_{ij} \otimes \Phi(E_{ij}) \\
&= \sum_{i,j} |e_i\rangle\langle e_j| \otimes |Xe_i\rangle\langle Ye_j| \\
&= \sum_{i,j} |e_i \otimes Xe_i\rangle\langle e_j \otimes Ye_j| \\
&= |\text{vec}(X)\rangle\langle \text{vec}(Y)|
\end{aligned}$$

as needed. $\square$

2.2 Completely positive maps

We start with a simple observation. Given a $\mathbb{C}$ -linear map $\Phi : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$, the following are equivalent (a map satisfying this condition is said to be $*$ -preserving)

1. For every self-adjoint $X \in \mathcal{B}(\mathcal{H}_A)$, the operator $\Phi(X)$ is self-adjoint
2. For every $X \in \mathcal{B}(\mathcal{H}_A)$, we have $\Phi(X^*) = \Phi(X)^*$
3. The Choi matrix C_Φ is self-adjoint.

To show the equivalence: 2. $\implies$ 1. is obvious; assuming 1., we see that 2. is true if X is either self-adjoint or anti-self-adjoint (i.e., iX is self-adjoint) and the general case follows; 3. means that $\Phi(E_{ij}) = \Phi(E_{ji})^*$ for every i, j , which is equivalent to 2.

We also note that there is a one-to-one correspondence between

- $\mathbb{C}$ -linear maps from $\mathcal{B}(\mathcal{H}_A)$ to $\mathcal{B}(\mathcal{H}_B)$ which are $*$ -preserving
- $\mathbb{R}$ -linear maps from $\mathcal{B}(\mathcal{H}_A)_{sa}$ to $\mathcal{B}(\mathcal{H}_B)_{sa}$

If $\Phi : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$ is a linear map, we may consider its adjoint $\Phi^* : \mathcal{B}(\mathcal{H}_B) \rightarrow \mathcal{B}(\mathcal{H}_A)$ which is characterized by the property

$$\text{Tr}[X\Phi(Y)] = \text{Tr}[\Phi^*(X)Y]$$

for every $X \in \mathcal{B}(\mathcal{H}_B)$ and $Y \in \mathcal{B}(\mathcal{H}_A)$.

A useful convention (which we do not follow) from physics literature is to use different symbols to denote different notions of adjoint. The (Hilbert space) adjoint of an operator $X \in \mathcal{B}(\mathcal{H})$ is denoted $X^\dagger$, and the (vector space) adjoint of a linear map $\Phi : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ is denoted Φ^* .

Definition. A linear map $\Phi : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$ is said to be

1. *positive* if, for every positive operator $X \in \mathcal{B}(\mathcal{H}_A)$, the operator $\Phi(X)$ is positive
2. *n-positive* if the map $\Phi \otimes \text{id}_{M_n} : \mathcal{B}(\mathcal{H}_A \otimes \mathbb{C}^n) \rightarrow \mathcal{B}(\mathcal{H}_B \otimes \mathbb{C}^n)$ is positive
3. *completely positive* (CP) if it is n -positive for every $n \geq 1$.

One check that Φ is positive (resp., CP) if and only if Φ^* is positive (resp., CP).

We now give a very example of a map which is positive but not completely positive. Consider the map $T : \mathcal{B}(\mathbb{C}^n) \rightarrow \mathcal{B}(\mathbb{C}^n)$ which sends a matrix M to its transpose ${}^tM$, i.e.

$$T((M_{ij})_{i,j}) = (M_{ji})_{i,j}$$

(this map is not canonical, it is associated with the choice of the canonical basis as a distinguished orthonormal basis in $\mathbb{C}^n$). This map is positive, for example because for every $\psi \in \mathbb{C}^n$ we have the relation.

$$\langle \psi | {}^t M \psi \rangle = \langle \bar{\psi} | M \bar{\psi} \rangle$$

However (if $n \geq 2$), the map T is not 2-positive. The map $T \otimes \text{id}_{M_2} : \mathcal{B}(\mathbb{C}^n \otimes \mathbb{C}^2) \rightarrow \mathcal{B}(\mathbb{C}^n \otimes \mathbb{C}^2)$ acts on block matrices as

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \mapsto \begin{pmatrix} {}^t A & {}^t B \\ {}^t C & {}^t D \end{pmatrix}.$$

This map is not positive. Assume $n = 2$ for simplicity and consider the positive rank-one matrix

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

Its image is

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

which is not positive, since its middle 2×2 block

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

has eigenvalues 1 and -1 .

It turns out that the class of completely positive maps has a simple structure.

Theorem 9 (Choi). *Consider a $\mathbb{C}$ -linear map $\Phi : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$. The following are equivalent*

1. *The map Φ is completely positive*
2. *Its Choi matrix C_Φ is positive*
3. *There exist $A_1, \dots, A_N$ in $\mathcal{B}(\mathcal{H}_A, \mathcal{H}_B)$ such that*

$$\phi(\rho) = \sum_{i=1}^N A_i \rho A_i^* \tag{2.2}$$

Equation (2.2) is called a Kraus decomposition of Φ , the operators A_i are called the Kraus operators and the minimal such N is called the Kraus rank of Φ .

Proof. We show that 3. implies 1. For every $X \in \mathcal{B}(\mathcal{H}_A, \mathcal{H}_B)$, it is easy to see that the map Φ_X defined by $\rho \mapsto X \rho X^*$ is positive. It is even completely positive, since

$$\Phi_X \otimes \text{id}_{M_n} = \Phi_{X \otimes \text{id}_{\mathbb{C}^n}}$$

Since a sum of CP maps is CP, this proves 1.

The implication 1. $\implies$ 2. follows from the formula

$$C_\Phi = (\Phi \otimes \text{id}_{\mathcal{B}(\mathcal{H}_A)})(|\chi\rangle\langle\chi|)$$

Assuming 2., by the spectral theorem, there are a_i in $\mathcal{H}_A \otimes \mathcal{H}_B$ such that $C_\Phi = \sum |a_i\rangle\langle a_i|$. To conclude, we define operators A_i by the relation $a_i = \text{vec}(A_i)$ and use Lemma 8. $\square$

If Φ as a Kraus decomposition given by (2.2), the adjoint map Φ^* has a Kraus decomposition given by

$$\Phi^*(\sigma) = \sum_i A_i^* \sigma A_i$$

2.3 Quantum channels

A map $\Phi : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$ is said to be

- *unital* if $\Phi(\text{id}_A) = \text{id}_B$
- *trace-preserving* if $\text{Tr} \Phi(X) = \text{Tr} X$ for every $X \in \mathcal{B}(\mathcal{H}_A)$
- *bi-stochastic* if it is both unital and trace-preserving (this is only possible if $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B)$).

These properties are dual to each other: Φ is unital if and only if Φ^* is trace-preserving.

If a map Φ has a Kraus decomposition $\Phi(\rho) = \sum A_i \rho A_i^*$, then

$$\Phi \text{ is unital} \iff \sum A_i A_i^* = \text{id}_\Phi \text{ is trace-preserving} \iff \sum A_i^* A_i = \text{id}$$

Definition. A *quantum channel* is a map $\Phi : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$ which is trace-preserving and unital.

Quantum channels preserve quantum states, even when applied to part of larger system.

Theorem 10 (Stinespring). *Let $\Phi : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$ be a CP map. There exists a Hilbert space $\mathcal{H}_E$ and a map $V : \mathcal{H}_A \rightarrow \mathcal{H}_B \otimes \mathcal{H}_E$ such that, for every $X \in \mathcal{B}(\mathcal{H}_A)$*

$$\Phi(X) = \text{Tr}_E V X V^* \quad (2.3)$$

Moreover, Φ is trace-preserving if and only if V is an isometry.

Proof. Write a Kraus decomposition

$$\Phi(X) = \sum_{i=1}^N A_i X A_i^*$$

Set $\mathcal{H}_E = \mathbb{C}^N$, with its canonical basis (e_i) . Define $V : \mathcal{H}_A \rightarrow \mathcal{H}_B \otimes \mathcal{H}_E$ by the formula, for every $x \in \mathcal{H}_A$

$$Vx = \sum_{i=1}^N A_i x \otimes e_i$$

We compute

$$V|x\rangle\langle x|V^* = |Vx\rangle\langle Vx| = \sum_{i,j=1}^N A_i |x\rangle\langle x| A_j^* \otimes |e_i\rangle\langle e_j|$$

It follows that

$$\text{Tr}_E V|x\rangle\langle x|V^* = \sum_{i=1}^N A_i |x\rangle\langle x| A_i^* = \Phi(|x\rangle\langle x|)$$

By linearity, formula (2.3) holds for every $X \in \mathcal{B}(\mathcal{H}_A)$.

The map V is an isometry iff $V^*V = \text{id}_{\mathcal{H}_A}$. The last part of the theorem follows since

$$V^*V = \sum_{i,j} A_i^* A_j \langle e_i | e_j \rangle = \sum_i A_i^* A_i \quad \square$$

Here is the variant when the domain and range are equal. For every quantum channel $\Phi : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$, there exist a Hilbert space $\mathcal{H}'$, a unit vector $\psi \in \mathcal{H}'$ and a unitary operator $U \in \mathcal{U}(\mathcal{H} \otimes \mathcal{H}')$ such that

$$\Phi(\rho) = \text{Tr}_{\mathcal{H}'} U(\rho \otimes |\psi\rangle\langle\psi|)U^*$$

Choose $\psi \in \mathcal{H}'$ arbitrarily and identify $\mathcal{H}$ with the subspace $\mathcal{H} \otimes \psi \subset \mathcal{H} \otimes \mathcal{H}'$. The isometry $V : \mathcal{H} \otimes \psi \rightarrow \mathcal{H} \otimes \mathcal{H}'$ given by Stinespring theorem can be extended to a unitary operator.

We now give some examples of quantum channels. A *unitary channel* is a channel of the form $\rho \mapsto U\rho U^*$ for a unitary operator U .

Exercise 4. Unitary channels are the only channels which send pure states to pure states, and also the only channels such which are bijective and such that the inverse map is also a channel.

A *mixed-unitary channel* is a convex combination of unitary channels. Mixed unitary channels are unital.

The *completely depolarizing channel* is the map $\Phi : M_d \rightarrow M_d$ given by

$$\Phi(X) = \text{Tr}(X) \frac{\text{id}}{d}$$

The channel

$$X \mapsto (1 - p)X + p \text{Tr}(X) \frac{\text{id}}{d}$$

can be used as a model of noise, which with probability p destroys the information.

The *dephasing channel* is the map which sends a matrix to its diagonal part (i.e. replaces non-diagonal entries by 0).

Exercise 5. Show that the depolarizing and dephasing channels are mixed-unitary channels.