Quantum Science and Technology

a.a. 2022-23
Elisa Ercolessi

These notes contains a detailed description of the Syllabus of the course.
Some Sections are complete while other Sections must be supplemented
with the topics treated in the book:
M.A. Nielsen, I.L. Chuang, Quantum Computation and Quantum Information
as indicated in the text.

Contents

Chapter 1

The Postulates of Quantum Mechanics

This Chapter covers all topics seen in class. Students can find a synthetic presentation also in Sect.ns 2.2.1 through 2.6.6 of [NC].

This Chapter introduces the postulates of quantum mechanics, which gives the fundamental ingredients to describe a quantum system, namely what we mean by: *states, observables, evolution, measurements* in the quantum context.

Just for comparison, let us recall their definition in *classical* mechanics:

i) The **state** of a physical system with N degrees of freedom, at a certain time t_0, is specified by N generalized coordinates $q_i(t_0)$ and N conjugate momenta $p_i(t_0)$. The space of coordinates and corresponding momenta is a symplectic manifold called phase space.

ii) **Observables** are given by real-valued (and at least continuous) functions f of $\{q_i, p_i\}$.

iii) The **time evolution** of a state $\{q_i(t), p_i(t)\}$ is fixed by the equations of motion, which are known if the Hamiltonian function $H(q_i, p_i)$ of the system is known; the state satisfies the Hamilton equations

$$\frac{dq_i}{dt} = \frac{\partial H_i}{\partial p_i}$$
$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}$$

The state is precisely known at any instant t if the initial state (initial condition) $\{q_i(0), p_i(0)\}$ is known.

iv) The **measure** of an observable f at time t yields the result $f(q_i(t), p_i(t))$.

In what follows, we will use Dirac notation, which is described in Appendix A.

1.1 States in QM

In quantum mechanics a state of a physical quantum system (at a certain time t) is given by a vector $|\psi\rangle$ belonging to an appropriate Hilbert space $\mathcal{H}$. This is the main and crucial difference with classical physics: states are not points in a generic manifold, but in a Hilbert space, which is a space with two fundamental structures:

- $\mathcal{H}$ is a vector space, so that every linear combination $\lambda|\chi\rangle + \mu|\psi\rangle$ of vectors $|\chi\rangle, |\psi\rangle \in \mathcal{H}$ (with $\lambda, \mu \in \mathbb{C}$) is still a state. In other words, the *superposition principle* holds.

- $\mathcal{H}$ is endowed with a scalar product, i.e. a positive, sesquilinear, symmetric form: $\langle\chi|\psi\rangle \in \mathbb{C}$, for all $|\chi\rangle, |\psi\rangle \in \mathcal{H}$.
  Because of the probabilistic interpretation that we will discuss below, a state is defined up to a normalization factor, meaning that we should assume the normalization condition: $\langle\psi|\psi\rangle = 1$, that is simply telling us that the total probability must be 1.

Even when the normalization condition is satisfied, the state remains defined up to a phase, since all vectors differing by a phase factor must be treated as equal: $|\psi\rangle \approx e^{i\varphi}|\psi\rangle$. It is possible to remove this ambiguity by describing states not as vectors but as projection operators. In fact, we can associate in a unique way to each state $|\psi\rangle$ the operator

$$P_\psi = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle}$$

which projects onto the (one-dimensional) linear subspace spanned by the vector $|\psi\rangle$ itself.

Exercise: verify that P_ψ does not depend on the phase factor.

The fundamental properties of projection operators are described in Appendix B.

THE QUBIT.
We now want to describe the quantum equivalent of a classical bit, which

is a system that is described by means of a binary variable, 0 or 1. Such a system is called *qubit* and is the simplest quantum system we can think of.

The state space of a qubit is described by a two-dimensional Hilbert space, therefore spanned by two orthonormall vectors that we call $|0\rangle$ and $|1\rangle$.
A pure[1] state of a qubit is therefore described by a vector of the form:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad , \quad |\alpha|^2 + |\beta|^2 = 1 \tag{1.1}$$

which depends on two complex coefficients α, β that have a probabilistic interpretation, that we will discuss below.

By taking into account that the overall phase of the vector is not important and the normalization condition $|\alpha|^2 + |\beta|^2 = 1$, we can rewrite eq. (1.1) as:

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle \quad s.t. \;\; \theta \in [0, \pi[\;, \;\; \phi \in [0, 2\pi] \tag{1.2}$$

Thus the coefficients α, β of the state of a qubit span a two dimensional sphere of unit radius, with coordinates θ, ϕ. This is called *Bloch sphere*, and it is represented in Fig. 1.1.

The z-axis intersects the Bloch sphere in two points, the North ($\theta = 0$) and the South ($\theta\pi$) poles that corresponds to the vectors $|0\rangle$ and $|1\rangle$.
The x-axis intersects the Bloch sphere in two points, for $\theta = \pi/2, \phi = 0$ and $\theta = \pi/2, \phi = \pi$ that that respectively give the two vectors:

$$|\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}} \tag{1.3}$$

Finally, the y-axis intersects the Bloch sphere in two points, for $\theta = \pi/2, \phi = \pi/2$ and $\theta = \pi/2, \phi = 3\pi/2$, that respectively give the two vectors:

$$|\pm i\rangle = \frac{|0\rangle \pm i|1\rangle}{\sqrt{2}} \tag{1.4}$$

All these three cases yield an orthonormal basis for the Hilbert space of a qubit; the states $|0\rangle, |1\rangle$ are said to form the computational basis.

Any two-levels system gives an experimental realization of a qubit. For example:

- an electron, where the two levels are given by the two spin states: $|0\rangle = |\uparrow\rangle, |1\rangle = |\downarrow\rangle$;

[1] Later on we will distinguish between pure and mixed states.

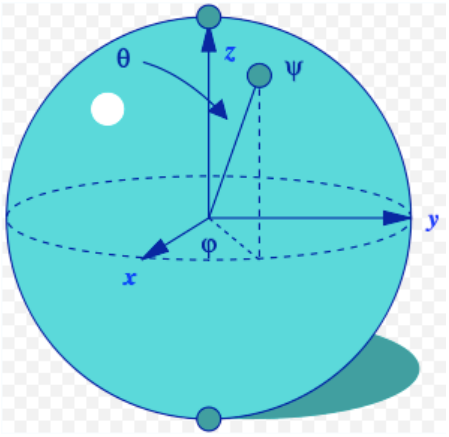

Figure 1.1: Bloch sphere.

- a photon, where the two levels are given by the two independent polarization states, e.g. horizontal and vertical: $|0\rangle = |H\rangle, |1\rangle = |V\rangle$;
- an atom, if the ground state and the first energy level are relatively close to each other and well separated from the higher ones: $|0\rangle = |g\rangle, |1\rangle = |e\rangle$.

When explicit calculations are needed, we will identify the computational basis in $\mathcal{H}$ with the canonical basis of $\mathbb{C}^2$:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \ , \ |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1.5}$$

Exercise. Verify that

$$\langle 0| = \begin{pmatrix} 1 & 0 \end{pmatrix} \ , \ \langle 1| = \begin{pmatrix} 0 & 1 \end{pmatrix} \tag{1.6}$$

and

$$P_0 = |0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \ , \ P_1 = |1\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{1.7}$$

Also, the matrices

$$\sigma_\pm = \frac{\sigma_x \pm \imath\sigma_y}{2} \tag{1.8}$$

are given by:

$$\sigma^+ = |0\rangle\langle 1| = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \ , \ \sigma^- = |1\rangle\langle 0| = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{1.9}$$

1.2 Observables in QM

In quantum mechanics, observables are identified by self-adjoint operators A on $\mathcal{H}$. This requirement is due to the fact that these operators are "diagonalizable", with a real spectrum (spectral theorem).
Sticking to finite-dimensions, this follows form the fact that the eigenvalue equation

$$A|\psi_i\rangle = \lambda_i|\psi_i\rangle$$

admits only real eigenvalues, $\lambda_i \in \mathbb{R}$, and that the corresponding eigenvectors $|\psi_i\rangle$ form an orthonormal basis for $\mathcal{H}$.
We can rephrase these properties in terms of the projection operators $P_i = |\psi_i\rangle\langle\psi_i|$ that project onto the subspace of each eigenvalue:

- orthonormality condition:

$$\text{the vectors } |\psi_i\rangle \text{ are orthonormal} \Leftrightarrow P_j P_k = \delta_{jk} P_j \ ;$$

- completeness relation:

$$\text{the set } \{|\psi_i\rangle\}_i \text{ is complete} \Leftrightarrow \textstyle\sum_{i=1}^{N} P_i = \mathbb{I};$$

- spectral decomposition:

$$\text{the vector } |\psi_i\rangle \text{ is an eigenvector of } A \text{ with eigenvalue } \lambda_i \Leftrightarrow$$

$$A = \sum_{i=1}^{N} \lambda_i P_i. \tag{1.10}$$

Given a (normalized) state ψ, we define *the mean value* of an observable A as (see Appendix C for the definition and the properties of trace):

$$\langle A\rangle \equiv \langle\psi|A|\psi\rangle = Tr[P_\psi\, A] \tag{1.11}$$

We can also define its standard deviation as:

$$\Delta A = \sqrt{\langle (A - \langle A\rangle)^2\rangle} = \sqrt{\langle A^2\rangle - \langle A\rangle^2} \tag{1.12}$$

PAULI MATRICES.
Observables for a qubit are 2×2 self-adjoint matrices. A (real) basis for such matrices is given by the identity $\mathbb{I}$ and the Pauli matrices $\sigma_x, \sigma_y, \sigma_z$:

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \ , \ \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ , \ \sigma_y = \begin{pmatrix} 0 & -\imath \\ \imath & 0 \end{pmatrix} \ , \ \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.13}$$

All three Pauli matrices $\sigma_x, \sigma_y, \sigma_z$ have eigenvalues ± 1, with eigenvectors given respectively by $|\pm\rangle$, $|\pm i\rangle$, $|0\rangle$ and $|1\rangle$.
They cannot share a common basis of eigenvectors because they do not commute: $[\sigma_\alpha, \sigma_\beta] = 2\imath\epsilon_{\alpha\beta\gamma}\sigma_\gamma$. As a consequence of these relations, one can also derive a Heisenberg uncertainty relation that we will discuss in a following section.
Instead, Pauli matrices anticommutes, since they satisfy: $\{\sigma_\alpha, \sigma_\beta\} = 2\delta_{\alpha\beta}\mathbb{I}$. In particular, one has: $(\sigma_\alpha)^2 = \mathbb{I}$.
One can also easily check that: $\sigma^\pm = (\sigma_x \pm i\sigma_y)/2$.

1.3 Evolution in QM

The time evolution of a quantum state is fixed once the Hamiltonian operator H of the system is known. It is determined by the Schrödinger equation:

$$\imath\hbar\frac{\partial}{\partial t}|\psi\,(t)\rangle = H|\psi\,(t)\rangle$$

Notice that the equation is linear, consistently with the superposition principle, and contains only first order time derivatives. This implies $|\psi(t)\rangle$ is -in principle- completely determined once an initial condition $|\psi(0)\rangle$ has been fixed. In this respect Schrödinger equation is fully deterministic.

If H does not depend on t, we can exponentiate this equation and reformulate the problem by saying that there must exist an evolution operator $U(t)$ so that

$$|\psi\,(t)\rangle = U(t)|\psi(0)\rangle$$

This operator is given by

$$U(t) = e^{-\frac{\imath tH}{\hbar}}$$

It is easy to verify that, since H is self-adjoint, $U(t)$ is unitary:

$$U(t)^\dagger = e^{-\frac{-\imath tH}{\hbar}} = U(-t) = U(t)^{-1}$$

This means that $U(t)$ preserves the scalar product and hence the probability interpretation that we will discuss below.

In general, any unitary operator U can be written as $U = e^{-\frac{\imath\lambda A}{\hbar}}$, for a suitable self-adjoint operator A and a real parameter λ (which can be absorbed in the definition of A itself). Therefore any unitary operator can represent the dynamics of a quantum system.

QUBIT EVOLUTION.
Since $\mathbb{I}, \sigma_x, \sigma_y, \sigma_z$ give a basis for self-adjoint 2×2 matrices

$$A = a_0\mathbb{I} + a_x\sigma_x + a_y\sigma_y + a_z\sigma_z = a_0\mathbb{I} + \vec{a}\cdot\vec{\sigma} \ , \ \ \vec{a} = (a_x, a_y, a_z)$$

(with real coefficients), the most general evolution of a qubit is given by the a unitary operator of the form:

$$\begin{aligned} U \ &= e^{\imath A} = e^{\imath\alpha} e^{\imath\frac{\theta}{2}\hat{n}\cdot\vec{\sigma}} = e^{\imath\alpha}(\cos\tfrac{\theta}{2}\,\mathbb{I} + \imath\sin\tfrac{\theta}{2}\,\hat{n}\cdot\vec{\sigma}) \\ &with \ \ \hat{n}\cdot\vec{\sigma} = \begin{pmatrix} n_z & n_x - \imath n_y \\ n_x + \imath n_y & -n_z \end{pmatrix} \end{aligned} \tag{1.14}$$

where we have set $a_0 = \alpha$, $\vec{a} = \frac{\theta}{2}\hat{n}$ with $\|\vec{a}\| = \frac{\theta}{2}\,\|\hat{n}\| = 1$.
The first term is just a multiplication by a global phase, while the second term originates a rotation on the Bloch sphere of an angle θ about the axis $\hat{n}$. In particular, the following matrices represent a rotation on the Bloch sphere of an angle θ about the x, y, z-axis respectively:

$$R_x(\theta) = e^{\imath\theta\sigma_x/2} = \begin{pmatrix} \cos\frac{\theta}{2} & -\imath\sin\frac{\theta}{2} \\ -\imath\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \tag{1.15}$$

$$R_y(\theta) = e^{\imath\theta\sigma_y/2} = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \tag{1.16}$$

$$R_z(\theta) = e^{\imath\theta\sigma_z/2} = \begin{pmatrix} e^{-\imath\frac{\theta}{2}} & 0 \\ 0 & e^{-\imath\frac{\theta}{2}} \end{pmatrix} \tag{1.17}$$

1.4 Measurements in QM

At odds with what happens in classical mechanics, measurements in QM are probabilsitic and destructive, as we will see now.

1.4.1 Projection Valued Measurements

A Projection Valued Measurement (PVM) depends on a family $\{P_m\}_m$ of (one-dimensional) projection operators such that:

- they are orthonormal: $P_jP_k = \delta_{jk}P_j$;
- and satisfy the completeness relation: $\sum_m P_m = \mathbb{I}$

and a corresponding set of real numbers $\{\lambda_m\}_m$.
If a system is in a state $|\psi\rangle$, we can use the the o.n basis of vectors $\{|\psi_m\rangle\}_m$ corresponding to $P_m = |\psi_m\rangle\langle\psi_m|$ to write the decomposition: $|\psi\rangle = \sum_m c_m|\psi_m\rangle$, with $c_m = \langle\psi_m|\psi\rangle$ and $\sum_m |c_m|^2 = 1$.

The measurement on the system is then

- *probabilistic*, since the possible outcomes are given by:

$$\lambda_m \;\; \textit{with probability} \;\; p_m = \langle\psi|P_m^\dagger P_m|\psi\rangle = \langle\psi|P_m|\psi\rangle = |c_m|^2 \; ,$$

- *destructive*, since after the measurement with outcome λ_m the state of the system has collapsed into the state:

$$\frac{P_m|\psi\rangle}{\langle\psi|P_m^\dagger P_m|\psi\rangle} = \frac{c_m}{|c_m|^2}|\psi_m\rangle = e^{\imath\alpha_m}|\psi_m\rangle \; .$$

Notice that the normalization condition of the vector $|\psi\rangle$ ensures that the sum of all probability is: $\sum_m p_m = \sum_m |c_m|^2 = 1$.

Any observable A defines a PVM, thanks to te spectral decomposition: $A = \sum_m \lambda_m P_m$.
We can say that the results of a measurement of an observable A on a state are its eigenvalues λ_m. Each eigenvalue is found with probability $p_m = |c_m|^2$, where c_m is the coefficient of the decomposition of the state $|\psi\rangle$ in the the basis of the eigenvectors of A.
The average of the measurements of A is therefore given by:

$$\langle A\rangle \equiv \sum_m \lambda_m p_m = \sum_m \lambda_m |c_m|^2 = \langle\psi|A|\psi\rangle \; .$$

This justifies the definition in eq.s (1.11) and (1.12).

MEASUREMENTS on a QUBIT .
We will start by looking at the *measurement on the computational basis*, which corresponds -as we will see- to the measurement along the z-axis of the Bloch sphere.
This means to consider the complete set of orthonormal projection operators:

$$P_0 = |0\rangle\langle 0| \;\; , \;\; P_0 = |1\rangle\langle 1|$$

and two (different) real numbers λ_0, λ_1. In the following we will assume $\lambda_0 = +1$, $\lambda_0 = -1$, so that the projective measurement we are considering corresponds to the self-adjoint operator:

$$+|0\rangle\langle 0| - |1\rangle\langle 1| = \sigma_z$$

If we consider a generic qubit state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, with $|\alpha|^2 + |\beta|^2 = 1$, we get the the possible outcomes of a measurement of σ_z are:

$$+1 \quad with \;\; p_0 = |a|^2$$

$$-1 \quad with \;\; p_0 = |b|^2$$

Exercise. Measurement along a generic direction $\hat{n}$ of the Bloch sphere.

Let us consider the observable $A = \hat{n} \cdot \vec{\sigma}$ with $\|\hat{n}\| = 1$.

a) Use Pauli matrices anticommutation relations to prove that $A^2 = \mathbb{I}$.

b) Show that A has eigenvalues equal to $+1$ and -1.

c) Define the two operators

$$P_\pm \equiv \frac{1}{2}\left(\mathbb{I} \pm \hat{n} \cdot \vec{\sigma}\right)$$

and show they are orthogonal projections. Show also that they give the resolution of the identity: $\mathbb{I} = P_+ + P_-$ and that the spectral decomposition of A is given by:

$$\mathbb{I} = P_+ - P_-$$

d) Given an arbitrary qubit state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, determine what are the possible values of the outcomes of a measure of A and the corresponding probabilities.

The distinguishablity problem.

A. Suppose a scientist, Alice, knows that a given apparatus is able to prepare a qubit either in the state $|0\rangle$ or in the state $|1\rangle$ and would like to determine the state in which the qubit is in, via a measurement.
To do so Alice can use a (projective) measurement of σ_z, since Alice will measure $+1$ (-1) with probability p=1 if the qubit is in the state $|\psi\rangle = |0\rangle$ $(|\psi\rangle = |1\rangle)$.

B. A similar kind of measurement can be constructed for any two orthogonal states $|\psi_1\rangle$ and $|\psi_2\rangle$, by considering the observable $A = +|\psi_1\rangle\langle\psi_1| - |\psi_2\rangle\langle\psi_2| = \hat{n} \cdot \vec{\sigma}$, for a suitable $\hat{n}$ (see previous exercise), as one can easily check.

C. Suppose now that Alice would like to distinguish the states $|\psi_1\rangle = |0\rangle$ and $|\psi_2\rangle = \frac{|0\rangle+|1\rangle}{2}$. Let's try to use a measurement of σ_z. The possible outcomes and corresponding probabilities are as follows:
- if the qubit is in the state $|\psi_1\rangle = |0\rangle$, Alice will measure ± 1 with probabilities $p_0 = 1$ and $p_1 = 0$,
- if the qubit is in the state $|\psi_2\rangle = \frac{|0\rangle+|1\rangle}{2}$, Alice will measure ± 1 with probabilities $p_0 = 1/1$ and $p_1 = 1/2$.

Hence, if the result of the measurement is -1, Alice can conclude with certainty that the qubit in the state $|\psi_2\rangle$, while if she measure $+1$, she cannot conclude anything. In particular there is no case in which Alice can conclude that the system is in the state $|\psi_1\rangle$.

One can show that no other PVM can do better, and this is due to the fact that the two states to be distinguished are not orthogonal. To improve this, one needs to consider a wider class of measurements,

1.4.2 Positive Operator Valued Measurements

The most general class of measurements in QM is give by Positive Operator Valued Measurements (POVM), of which PVMs represent a subclass. A POVM is identified by a a family $\{M_m\}_m$ of operators on which we ask only the condition:

$$\sum_m M_m^\dagger M_m = \mathbb{I}$$

The possible outcomes of the measurement are given by a family of real numbers $\{\lambda_m\}_m$. The measurement on a system in the state $|\psi\rangle$ is:

- *probabilistic*, with the possible outcomes given by:

$$\lambda_m \ \textit{with probability} \ p_m = \langle\psi|M_m^\dagger M_m|\psi\rangle \ ,$$

- *destructive*: after the measurement with outcome λ_m the state of the system has collapsed into the state:

$$\frac{M_m|\psi\rangle}{\langle\psi|M_m^\dagger M_m|\psi\rangle}$$

We exploit this definition to look back at the problem of distinguishing two states for a qubit, which can be in one of the two non-orthogonal states $|\psi_1\rangle = |0\rangle$ and $|\psi_2\rangle = \frac{|0\rangle+|1\rangle}{2}$. Alice can prepare a POVM by choosing three possibles outcomes $\lambda_0, \lambda_1, \lambda_2$ corresponding to the operators:

$$M_0 = |1\rangle\langle 1| \ , \ M_1 = |\phi\rangle\langle\phi| \ \textit{with} \ |\phi\rangle = \frac{|0\rangle - |1\rangle}{2}$$

$$M_2 \ s.t. \ M_2^\dagger M_2 = \mathbb{I} - M_0^\dagger M_0 - M_1^\dagger M_1$$

Notice that the states that are used to construct M_0 and M_1 are orthogonal to $|\psi_1\rangle$ and $|\psi_2\rangle$ respectively.

The possible outcomes and corresponding probabilities are as follows:
- if the qubit is in the state $|\psi_1\rangle = |0\rangle$, Alice will measure λ_0 with probability $p_0 = 0$, λ_1 with probability $p_1 \neq 0$, λ_2 with probability $p_2 \neq 0$;
- if the qubit is in the state $|\psi_2\rangle = \frac{|0\rangle+|1\rangle}{2}$, Alice will measure λ_0 with probability $p_0 \neq 0$, λ_1 with probabilitiy $p_2 = 0$, λ_2 with probability $p_2 \neq 0$;
Hence, if the result of the measurement is λ_0, Alice can conclude with certainty that the qubit in the state $|\psi_2\rangle$, while if she measure λ_1 she can conclude with certainty that the qubit in the state $|\psi_1\rangle$. If she measure λ_2, instead, she cannot conclude anything.
However this protocol is better than the one described above since now Alice can distinguish both states some of the times, without making errors of misclassification.

1.4.3 The Heisenberg principle

In a generic Hilbert space and for any two observables A, B, one can show the Heisenberg uncertainty principle:

$$(\Delta A)(\Delta B) \geq \frac{1}{2}\langle [A, B]\rangle \tag{1.18}$$

This principle is telling us that the product of the standard deviation of any two observables that do not commute cannot be zero, being limited from below by (one half of) the average of their commutator.
In other words, we can conclude that two operators that do not commute cannot be determined simultaneously with perfect accuracy. i.e. with certainty.

Exercise.
To prove the above equation, you can proceed as follows:
i) Show that, for any state $|\psi\rangle$:

$$\langle [A, B]\rangle \equiv \langle \psi|[A, B]|\psi\rangle = 2Re[\langle \psi|AB|\psi\rangle]$$

$$\langle \{A, B\}\rangle \equiv \langle \psi|\{A, B\}|\psi\rangle = 2\imath Im[\langle \psi|AB|\psi\rangle]$$

ii) Use Cauchy-Schwartz inequality: $|\langle \psi|AB|\psi\rangle]|^2 \leq \langle A^2\rangle\langle B^2\rangle$ to show that:

$$|\langle \psi|[A, B]|\psi\rangle]|^2 \leq 4\langle A^2\rangle\langle B^2\rangle$$

iii) Derive now eq. (1.18) by putting together i) and ii).

In particular this can be applied to the simultaneous measure of two Pauli matrices. One has:

$$(\Delta\sigma_\alpha)(\Delta\sigma_\beta) \geq \frac{1}{2}|\langle[\sigma_\alpha, \sigma_\beta]\rangle| = \epsilon_{\alpha\beta\gamma}\langle\sigma_\gamma\rangle \neq 0 \tag{1.19}$$

In the Stern-Gerlach experiment (see below), we can conclude that we cannot measure simultaneously with certainty the value of the spin along two arbitrary directions (e.g. x and z).
This also implies that in QM it is not possible to assign a given property (say, being spin up or down along the z-direction) a-priori; this is something that the quantum object acquires as a consequence of a specific measurement, which however does not belong intrinsically to the state of the particle.

1.5 Probability in QM

In classical mechanics, a system is described by a (pure) state x (x here denotes the coordinates of all degrees of freedom) which, in principle, can be determined with absolute precision by means of a suitable apparatus. Probability enters when a system can be prepared in a *mixed state*, that is in a (classical) *mixture* of possible states. The latter is identified by an *ensemble*, i.e. by a set of possible states in which the system can be in, with a corresponding probability.

$$\{x_j, p_j\}_j \quad with \quad p_j > 0\ ,\ \sum_j p_j = 1$$

The results of a series of n measurement on a mixture (for example we draw the dice n times), are given by the laws of classical probabilities, once we assume that the different measures are independent:

- the probability to obtain the ordered sequence $s = (x_{j_1}, x_{j_2}, \cdots, x_{j_n})$ is given by the product of the single outcome probabilities: $p_s = p_{j_1}p_{j_2}\cdots p_{j_n}$;
- the probability of a given event e is given by the sum of the probabilities of all sequences that yield the event e as result: $p_e = \sum p_s$.

For example, if the system is a perfect dice, we can assume $x_j = 1, 2, 3, 4, 5, 6$ with $p_1 = \cdots = p_6 = 1/6$. If we draw a dice twice and the event e is given by the fact that the sum of the two dices is 4, we should consider all series that lead to such event: $s_1 = (1, 3), s_2 = (3, 1), s_3 = (2, 2)$ which all

occur with probability $p_s = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$. The probability of the event is: $p_e = p_1 + p_2 + p_3 = 3 \cdot \frac{1}{36} = \frac{1}{12}$.

In QM the situation is different, since the (pure) state of a system can be a linear combination, which has intrinsically a probabilistic meaning. For example a qubit can be in the state $|\psi\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$ which is superposition of the state $|0\rangle$ and $|1\rangle$: the results of a measurement in the computational basis tell us that the qubit is found in the state $|0\rangle$ or in the state $|1\rangle$ with probabilities $p_0 = p_1 = 1/2$. This is called a *quantum mixture*.
This probabilities can be evaluated experimentally, by sending -for instance:
- a large number of spin 1/2 particles (electron or atom) through a Stern-Gerlach apparatus with the magnetic field along the $z-$direction; when particles are collected on a screen/detector (measure), we see that half of the particles group in the upper/lower spot;
- a large number of photons particles through a linear polarizer with the magnetic field along, say, the vertical direction; when particles are collected on a screen/detector (measure), we see that only half of the particles have gone through the polarizer.
However, in QM we can also have a situation identical to the classical one, which is called a *classical mixture* or *mixed state*. This is described by an ensemble of particles which are half in the state $|0\rangle$ and half in $|1\rangle$.
Notice that a Stern-Gerlach apparatus or a polarizer experiment with the magnetic field/polarizer axis in the z-direction is NOT able to distinguish between the classical and the quantum mixture. However it is not difficult to devise an experiment which is able to do so: for example we can choose the magnetic field/polarizer axis at 45^o with respect to the z-axis. Now, the experiment yields the probabilities of the two alternatives are equal (and equal to 1/2) only for the classical mixture; in the case of the quantum mixture instead, we will see all particles to have spin up/to go through the polarizer.

Exercise: the Game of the Coin

1. A classical coin.

Let us consider the following game: there are two players, Alice and Bob, with a (classical) coin that with head/tail. Starting from one of the two possibilities, and without looking at the coin: - first Alice can decided to flip the coin or not
- then Bob can decided to flip the coin or not - finally Alice can decide to flip the coin or not.
Alice (Bob) wins if the face of the coin is the same (opposite) at the beginning or at the end of the game.

Show that Alice and Bob have the same probabilities to win/lose.

2. 1. A quantum coin.

The coin now is a qubit and therefore can be in any superposition of the two classical states head and tail. Bob is a classical player and can only flip or not the two states. On the contrary, Alice is a quantum player and can act with any unitary evolution on the coin. Starting from one of the two possibilities, say head, and without looking at the coin:
- first Alice can decided to do nothing, to flip the coin (NOT) or to use the Hadamard (H) unitary transformations, that are defined as follows :

$$NOT : \alpha|0\rangle + \beta|1\rangle \mapsto \beta|0\rangle + \alpha|1\rangle \tag{1.20}$$

$$H : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha\frac{|0\rangle + |1\rangle}{\sqrt{2}} + \beta\frac{|0\rangle - |1\rangle}{\sqrt{2}} = \frac{\alpha+\beta}{\sqrt{2}}|0\rangle + \frac{\alpha-\beta}{\sqrt{2}}|1\rangle \tag{1.21}$$

- then Bob can decided to flip the coin or not,
- finally Alice can decide to do nothing or to use the flip or the Hadamard transformation.

Alice (Bob) wins the face of the coin is the same (opposite) at the beginning or at the end of the game.
Show that Alice can play in such a way to always win.

This example show that in QM, therefore, the rules of probabilities have to be changed. This is because we have now to deal with probability amplitudes that can be superimposed. Indeed, the possible events are represented by the chooce of a (say) projective measurement identified by the projectors $P_j = |e_j\rangle\langle e_j|$: a generic state of the system is given by the vector $|\psi\rangle = \sum_j c_j|e_j\rangle$ where c_j are the probability amplitudes, so that $p_j = |c_j|^2$ is the probability to measure the system in the state $|e_j\rangle$. Then, we have to apply the following rules:

- in a series of transformations, the probability amplitudes have to be multiplied;
- the final probability amplitude is obtained by summing up the probabilities amplitudes of all sequences that yield the same event (state $|e_j\rangle$);
- the final probability is obtained by squaring the corresponding final probability amplitude.

We will see these rules applied in two examples below.

The Stern-Gerlach experiment.
Let us consider the Stern-Gerlach experiment shown in Fig. 1.2

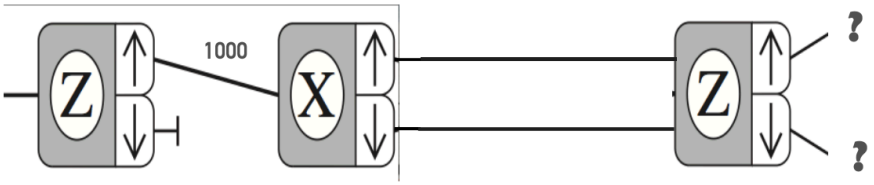

Figure 1.2: The Stern-Gerlach experiment.

Fig. 1.3 and Fig. 1.4 show the calculation of the final probability by applying the classical and the quantum rules respectively.

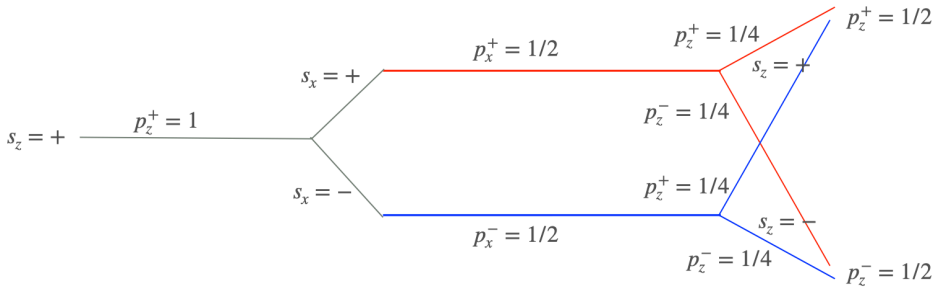

Figure 1.3: The Stern-Gerlach experiment: classical probabilities

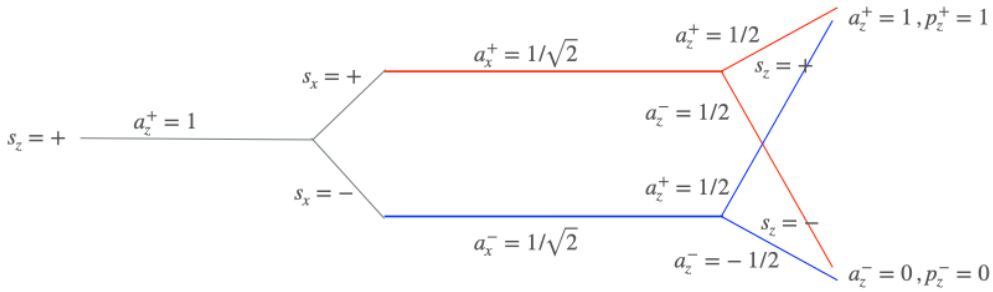

Figure 1.4: The Stern-Gerlach experiment: quantum probabilities

Let us notice that in the quantum case we have that the two alternatives correspond to sum two probability amplitudes that in one case are both positive, in the other are one positive and one negative, leading to constructive and destructive interference effects respectively. The amplitude can be

calculated by taking into account that:

$$|+\rangle_z = |0\rangle = \frac{|+\rangle_x + |-\rangle_x}{\sqrt{2}}$$

and

$$|+\rangle_x = \frac{|+\rangle_z + |-\rangle_z}{\sqrt{2}} = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

$$|-\rangle_x = \frac{|+\rangle_z - |-\rangle_z}{\sqrt{2}} = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

The Mach-Zender experiment.

This experiment is based on a device which is called beam splitter: it is, for example, a 50-50 mirror that, classically, reflects half of a light beam, allowing the other half to go through. From a quantum point of view, this means that a single photon has equal probabilities $p_1 = p_2 = 1/2$ to be reflected or to go through. Thus that the state $|v\rangle$ of the incoming photon is a superposition of the two states $|v_1\rangle$ and $|v_2\rangle$ corresponding respectively to be a reflected or a transmitted photon, as shown in Fig. 1.5.

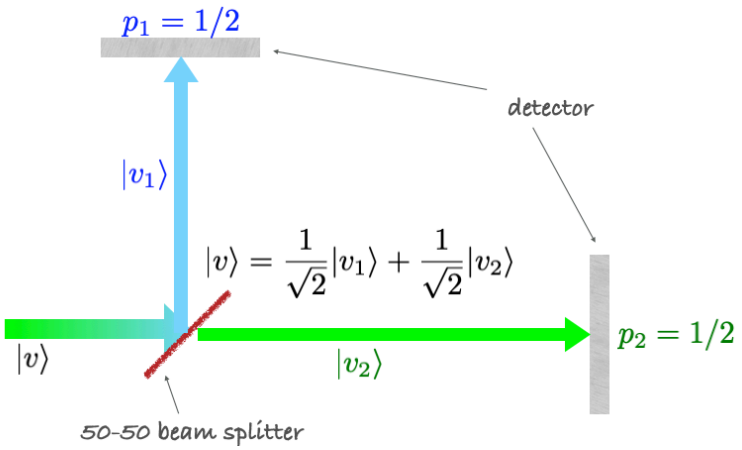

Figure 1.5: A 50-50 Beam Splitter

The Mach Zender apparatus contain a first beam splitter: both the reflected and the transmitted photon are then made go through a second beam splitter by means of suitable positioned mirrors, as shown in Fig. 1.6.
We leave as an exercise to show that the probability of finding the photon in detector 1 is $p_1 = 1$ while that of finding the photon in detector 2 is $p_2 = 0$. This is due to a quantum interference effect similar to the one that occurred in the Stern-Gerlach experiment.

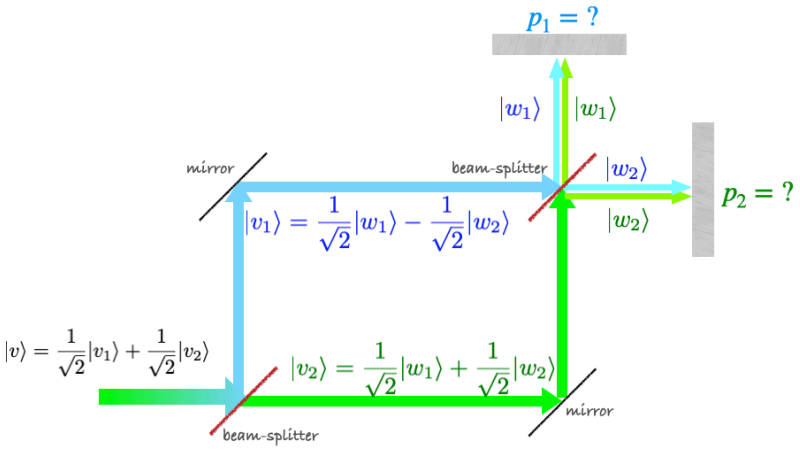

Figure 1.6: A 50-50 Beam Splitter

Chapter 2

Properties of composite systems

This Chapter covers all topics seen in class. Students can find a synthetic presentation also in Sect.ns 2.2.8, 2.4, 2.6 through 2.6.6 of [NC].

In this chapter we will introduce the description of composite quantum systems and their properties, including entanglement. We will also discuss the formalism of density matrices.

2.1 Composite systems

In classical mechanics, the total phase space of two (sub)systems is simply the cartesian product of the two phase spaces: $\mathcal{M}_t = \mathcal{M}_1 \times \mathcal{M}_2$ and has therefore dimension $d = d_1 + d_2$, if $dim(\mathcal{M}_{1,2}) = d_{1,2}$.
In quantum mechanics, instead, the total Hilbert space of two (sub)sytems is the *tensor product* of the two Hilbert spaces:

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

and therefore its dimension is $d = d_1 \cdot d_2$, if $dim(\mathcal{H}_{1,2}) = d_{1,2}$. Thus the dimension of the Hilbert spaces of N particles scales exponentially with the number of particles. This is the reason why quantum systems are in general hard to simulate in a classical computer and one would like to resort to quantum computation.

From the definition of tensor product of Hilbert spaces (see Appendix D), we have that a generic state of a state $|\psi\rangle \in \mathcal{H}$ is of the form:

$$|\psi\rangle = \sum_{\alpha\beta} |\psi_\alpha\rangle|\phi_\beta\rangle \ \ , \ |\psi_\alpha\rangle \in \mathcal{H}_1 \ , \ |\phi_\beta\rangle \in \mathcal{H}_2$$

If the sum contains just a single term so that the state can be written as a product of a vector in $\mathcal{H}_1$ and a vector in $\mathcal{H}_2$, the state is said to be a product

state or a *separable* state. If that is not possible, the state is called *entangled*. As we will see in the following, entanglement is a genuine quantum feature that can be exploited to develop protocols that are not allowed classically.

The state of a composite system can be evolved in different ways, by manipulating only subsystem 1 (or subsystem 2) by means of unitary operators of the form $U_1 \otimes \mathbb{I}_2$ ($\mathbb{I}_1 \otimes U_2$) or both by means of operators of the form $U_1 \otimes U_2$:

$$U_1 \otimes U_2 : |\psi\rangle = \sum_{\alpha\beta} |\psi_\alpha\rangle|\phi_\beta\rangle \mapsto \sum_{\alpha\beta} U_1|\psi_\alpha\rangle U_2|\phi_\beta\rangle$$

In all these cases vectors in $\mathcal{H}_{1,2}$ are transformed inside each Hilbert space $\mathcal{H}_{1,2}$ and do not mix inside the total Hilbert space $\mathcal{H}$. However we might also have unitary operators U that act on the total space that cannot be split in the tensor product of operators acting separately on the two subsytem Hilbert spaces. We will give examples when discussing quantum circuits.

A similar reasoning applies when considering observables: we have observables of the kind $A_1 \otimes A_2$ which act separately on the two Hilbert spaces, as well as global observables that can measure only physical quantities of the total system. As a consequence we can provide measurements on only one of the two subsystem (leaving the other unperturbed), or on both separately or, finally, on both in a joint manner. We will give examples below.

THE 2-QUBIT SYSTEM.
The Hilbert space of two qubits is $\mathcal{H}_t = \mathbb{C}^2 \otimes \mathbb{C}^2 \simeq \mathbb{C}^4$.
If, for each qubit we choose the computational basis $|0\rangle, |1\rangle$, we can construct the orthonormal basis on $\mathcal{H}$, which is called the computational basis:

$$|00\rangle,\ |01\rangle,\ |10\rangle,\ |11\rangle \tag{2.1}$$

where the first/second digit represents the state of the first/second qubit. Hence, a generic state of a 2-qubit system is of the form:

$$|\Psi\rangle = a_{00}|00\rangle + a_{01}|01\rangle + a_{10}|10\rangle + a_{11}|11\rangle \tag{2.2}$$

with $|a_{00}|^2 + |a_{01}|^2 + |a_{10}|^2 + |a_1|^2 = 1$. Let us consider some examples:

a) The state $|\psi\rangle = \frac{|00\rangle+|01\rangle}{\sqrt{2}} = |0\rangle \frac{|0\rangle+|1\rangle}{\sqrt{2}} = |0\rangle|+\rangle$ is an example of a product state.

b) The state $|\Psi^+\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$ is instead an example of an entangled state. Indeed, the following four states, which are called Bell states:

$$|\Psi^\pm\rangle = \frac{|00\rangle \pm |11\rangle}{\sqrt{2}} \quad , \quad |\Phi^\pm\rangle = \frac{|01\rangle \pm |10\rangle}{\sqrt{2}} \tag{2.3}$$

form an orthonormal basis of (maximally[1]) entangled states.

One important two-qubit transformation is the Controlled NOT (CNOT), which is defined as follows:

$$\begin{aligned} CNOT : a_{00}|00\rangle + a_{01}|01\rangle + a_{10}|10\rangle + a_{11}|11\rangle \\ \mapsto a_{00}|00\rangle + a_{01}|01\rangle + a_{11}|11\rangle + a_{11}|10\rangle \end{aligned} \tag{2.4}$$

Form this expression it is easy to see that the if the first qubit (control qubit) is zero, nothing is done on the second qubit (target qubit), while if the first qubit is one, the NOT operation (see eq. (1.20)) is applied on the second qubit. On the computational basis, this transformation is therefore given by the matrix:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{2.5}$$

Let us now consider some examples of measurements.
Suppose first that we perform a measurement on the computational basis only of subsystem 1 (Alice's lab) or 2 (Bob's lab). If we the state is $|\psi\rangle = \frac{|00\rangle+|01\rangle}{\sqrt{2}}$ Alice will measure 0 while Bob will measure either 0 or 1 with fifty-fifty probability. The results of Alice's and Bob's measuremnt are completed uncorrelated. This is summarized in Fig. **??**.
Now we perform a similar measure but on the Bell state $|\Psi^+\rangle$: the possible outcomes of the measurements are shown in Fig. **??**. One can see that both Alice and Bob cannot predict the results of their own measurement (for both there is a fifty-fifty probability to get either o or 1). However there is a perfect correlation between the measurements of Alice and Bob: after the measurement, each of them can predict with certainty the result of the other. Such correlation/intertwining of the properties of the two subsystems is due to entanglement and is a characteristic that cannot be found in classical physics.
Finally, we can perform some joint measures, on the state $|\Psi^+\rangle$ fro instance. We can choose first the computation basis of the total system, i.e. we consider the four projectors:

$$P_{00} = |00\rangle\langle 00| \ , \ P_{01} = |01\rangle\langle 01| \ , \ P_{10} = |10\rangle\langle 10| \ , \ P_{11} = |11\rangle\langle 11|$$

Now, the possible outcomes are either 00 or 11 with probabilities $p_{00} = p_{11} = 1/2$.

[1]We will discuss the notion of degree of entanglement in the last chapter.

If, instead, we choose to perform a measurement in the Bell's state basis, i.e. we consider the four projectors:

$$P_{\Psi^+} = |\Psi^+\rangle\langle\Psi^+| \ , \ P_{\Psi^-} = |\Psi^-\rangle\langle\Psi^-| \ , \ P_{\Phi^+} = |\Phi^+\rangle\langle\Phi^+| \ , \ P_{\Phi^-} = |\Phi^-\rangle\langle\Phi^-|$$

there is only one possible outcome corresponding to the first projector, with probability 1.

We end this section by giving some examples of states of a 3-qubit system. Clearly $\mathcal{H}_t = \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \simeq \mathbb{C}^6$ and the computational basis is given by the eight states :

$$|000\rangle \ , \ |001\rangle \ , \ |010\rangle \ , \ |011\rangle \ , \ |100\rangle \ , \ |101\rangle \ , \ |110\rangle \ , \ |111\rangle$$

Examples of entangled states are the Greenberger-Horne-Zeilinger (GHZ) state:

$$|GHZ\rangle = \frac{|000\rangle + |111\rangle}{\sqrt{2}}$$

and the W state:

$$|W\rangle = \frac{|001\rangle + |010\rangle + |100\rangle}{\sqrt{3}}$$

One says that the entanglement carried by the former is less robust than the one carried by the latter. This can be understood by looking at what happens if we perform a measure in the computational basis, say, on the last qubit. If we start from the GHZ state, we end up in a separable state of the othertwo wubits (either $|00\rangle$ or $|11\rangle$); if we start from the W state, instead we end up in a still entangled state of the other two qubits $((|00\rangle + |01\rangle)/\sqrt{2}$.

The Schrödinger cat.

Entanglement is at the heart of the infamous Schrödinger cat paradox, in which a cat is put in a room together with a two-level atom that can be in the ground state $|g\rangle$ or in the excited state $|e\rangle$. If the atom decays in the ground state, a hammer breaks a poison bottle, killing the cat, which can therefore be in the alive state $|A\rangle$ or dead state $|D\rangle$. The total system cat+atom is therfore described by the entangled state: $a|Ae\rangle + b|Dg\rangle$, where $|a|^2, |b|^2$ represent the probabilities of the atom to be in the excited/ground state. When we make a measurement (by opening the dooor or measuremnet the state of the atom), we force the system to collapse either in $|A\rangle$ or $|D\rangle$, thus determining whether the cat is dead or alive.

2.2 The EPR paradox and Bell inequalities

Schrodinger's cat paradox makes evident that, in quantum mechanics -at odds with what happens in classical physics-, we cannot attribute to a physical system a property which exists independently of measurements performed on the system. In their seminal paper[2] Einstein, Podolsky and Rosen described an ideal experiment to remark that the physical description provided by quantum mechanics cannot be complete and put forward the idea of hidden variables.

Their reasoning starts from the definition of when a theory can be considered to be complete: a physical theory is complete if it contains all *elements of reality*, where the latter are defined as follow:
"If, without in any way disturbing a system, we can predict with certainty (i.e. with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity". EPR also suppose that a physical theory should be local, i.e. that physical phenomena (evolution, measurements, ...) are independent one from the other if they occur at different time-space coordinates that cannot be connected by signals travelling at a speed smaller or equal than the speed of light.

We will give a version of the argument proposed by EPR due to to Clauser and Horne. Consider two particles which are prepared in the maximally entangled (singlet) Bell state: $|\Phi^-\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$. This has the property that if we perform the measurement of a generic rotation operator $R_n = \hat{n} \cdot \vec{\sigma}$ on both qubits, we always get opposite values: $(\hat{n} \cdot \vec{\sigma})| \pm \hat{n}\rangle = \pm| \pm \hat{n}\rangle$ To be complete, QM should attribute an intrinsic value to each R_n. In principle, we can choose $R = X$ or $R = Z$ and should conclude that the system should have an intrinsic value of the spin along either direction, but this is in contrast with Heisenberg's uncertainty principle.

Hence, EPR concluded that quantum mechanics is not a complete theory. Thus, there should be some hidden variables that we are not aware of, which determine such a probabilistic behaviour and contain all information about the properties of the system. Since then, this topics has been widely discussed in the phyical community, mainly strtating form theoretical and/or epistemological considerations. It was with the work by Bell that this discussion entered the realm of experimental verification.

To derive **Bell's inequality**, let us consider a two particle system, on which Alice and Bob can separately perform measurements. Alice can per-

[2] A. Einstein, B. Podolsky, N. Rosen, Phys. Rev. Letters 47 (1935) 777

form measures of two different binary properties: $Q = \pm 1, R = \pm 1$ which can occur with classical probabilities P_Q, P_R. Similarly, Bob can measure two different binary properties $S = \pm 1, T = \pm 1$ which can occur with classical probabilities P_S, P_T. We assume locality and that Alice and Bob perform the experiment at the same time, so there no light signal (i.e. no information) propagation.
The system can be prepared in one of the 16 the states given by $(Q = q, R = r, S = s, T = t)$ $(q, r, s, t = \pm 1)$ with probability $p(q, r, s, t)$. Thus the average of the observable $QS + RS + RT - QT$ is

$$\begin{aligned} &\mathbf{E}(QS + RS + RT - QT) \equiv \textstyle\sum_{q,r,s,t=\pm 1} p(q, r, s, t)\,(qs + rs + rt - qt) \quad (2.6)\\ &\leq \textstyle\sum_{q,r,s,t=\pm 1} p(q, r, s, t)\,|qs + rs + rt - qt| = \sum_{q,r,s,t=\pm 1} p(q, r, s, t) \times 2 = 2 \end{aligned}$$

since $|qs+rs+rt-qt| \leq 2$ (as it can be easily verified on the 16 combinations) and $\sum_{q,r,s,t=\pm 1} p(q, r, s, t) = 1$. Also:

$$\begin{aligned} \mathbf{E}(QS + RS + RT - QT) &= \textstyle\sum_{q,r,s,t=\pm 1} p(q, r, s, t)qs + \sum_{q,r,s,t=\pm 1} p(q, r, s, t)rs \\ &\quad + \textstyle\sum_{q,r,s,t=\pm 1} p(q, r, s, t)rt - \sum_{q,r,s,t=\pm 1} p(q, r, s, t)qt \\ &= \mathbf{E}(QS) + \mathbf{E}(RS) + \mathbf{E}(RT) - \mathbf{E}(QT) \quad (2.7) \end{aligned}$$

Comparing (2.6) and (2.7), we obtain the (Bell) inequality:

$$\mathbf{E}(QS) + \mathbf{E}(RS) + \mathbf{E}(RT) - \mathbf{E}(QT) \leq 2 \quad (2.8)$$

in the version due to Clauser, Horne, Shimony, Holt (CHSH inequality).
Let's see now what happens in a quantum system. We consider rqo qubits in the Bell state $|\Phi^-\rangle$ and we suppose that Alice can measure the spin of the first qubit along either the z and x direction: $Q = \sigma_z, R = \sigma_x$, while Bob measures the spin of the second qubit along the directions: $S = (-\sigma_z - \sigma_x)/\sqrt{2}, T = (\sigma_z - \sigma_x)/\sqrt{2}$. Some algebra shows that

$$\langle QS\rangle = \frac{1}{\sqrt{2}}; \langle RS\rangle = \frac{1}{\sqrt{2}}; \langle RT\rangle = \frac{1}{\sqrt{2}}; \langle QT\rangle = -\frac{1}{\sqrt{2}}. \quad (2.9)$$

so that

$$\mathbf{E}(QS + RS + RT - QT) = \langle QS\rangle + \langle RS\rangle + \langle RT\rangle - \langle QT\rangle = 2\sqrt{2} \quad (2.10)$$

proving that quantum mechanics violates Bell's inequality.

The first experiment in this direction was done by Clauser, but the violation of Bell's inequality was demonstrated in the historical experiments due to Aspect[3].

[3]Clauser, Aspect and Zeilinger got the Nobel prize "for experiments with entangled photons, establishing the violation of Bell inequalities and pioneering quantum information science" in 2022

2.3 Teleportation

Let us first show here an important theorem, the *No Cloning Theorem* that states that: it is impossible to find a unitary trasnformation whose only effect is to copy an arbitrary state $|\psi\rangle$.

Let us suppose such an operator exists: $U : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ such that $U : |\psi\rangle \to |\psi\rangle \otimes |\psi\rangle$. Choosing two arbitrary states $|\psi\rangle, |\phi\rangle$ we have:

$$\langle \psi \otimes \phi | \psi \otimes \phi \rangle = (\langle \psi | \phi \rangle)^2$$

by the definition of tensor product and

$$\langle \psi \otimes \phi | \psi \otimes \phi \rangle = \langle \psi | U^\dagger U | \psi \rangle = \langle \psi | \phi \rangle$$

by construction. Hence: $(\langle \psi | \phi \rangle)^2 = \langle \psi | \phi \rangle$ which is however possible only if $|\psi\rangle = |\phi\rangle$ or they are orthogonal, not for arbitrary states.

This theorem is at odds with classical mechanics where the copy of a state of a system is always possible. Indeed this is a necessary operation of classical computing algorithms.
We will see how this theorem can be "evaded" (but not violated) in the protocol of Quantum Information of Teleportation and in algorithms of Quantum Computation thanks to the Toffoli gate.

Quantum Teleportation is a protocol to recreate an unknown state at a distance. Notice that only states, not matter, are transported.

We will describe here the protocol to teletransport the unknown state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ from Alice to Bob. In order to do so, alice and Bob should share qubits of an entangled pair: $|\beta_{00}\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, so the initial state is:

$$|\psi_0\rangle = |\psi\rangle\,|\beta_{00}\rangle = \frac{1}{\sqrt{2}}[\alpha(|000\rangle + |011\rangle) + \beta(|100\rangle + |111\rangle)]$$

where the first two qubits are in Alice's lab, while the third one is in Bob's lab. Now Alice performs a CNOT operation (eq. (2.4)) with the unknown qubit as control, getting:

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}[\alpha(|000\rangle + |011\rangle) + \beta(|110\rangle + |101\rangle) \tag{2.11}$$

She then sends the first qubit through a Hadamard gate (eq. (1.21)) obtaining:

$$|\psi_2\rangle = \frac{1}{2}[\alpha(|0\rangle + |1\rangle)(|00\rangle + |11\rangle) + \beta(|0\rangle - |1\rangle)(|10\rangle + |01\rangle)] \tag{2.12}$$

Separating now the qubits the qubits in Alices' and Bob's hands, we can now regroup the different terms as:

$$|\psi_2\rangle = \frac{1}{2}[|00\rangle(\alpha|0\rangle + \beta|1\rangle) + |01\rangle(\alpha|1\rangle + \beta|0\rangle) + |10\rangle(\alpha|0\rangle - \beta|1\rangle) + |11\rangle(\alpha|1\rangle - \beta|0\rangle)] \tag{2.13}$$

Now Alice performs a measurements on her two qubits. The following table shows the four possible outcomes and the states in which the total system collapse:

outcome c_1c_2	*Alice's state*	*Bob's state*
00	$\|00\rangle$	$\alpha\|0\rangle + \beta\|1\rangle = \mathbb{I}(\alpha\|0\rangle + \beta\|1\rangle)$
01	$\|01\rangle$	$\beta\|0\rangle + \alpha\|1\rangle = X(\alpha\|0\rangle + \beta\|1\rangle)$
10	$\|10\rangle$	$\alpha\|0\rangle - \beta\|1\rangle = Z(\alpha\|0\rangle + \beta\|1\rangle)$
11	$\|10\rangle$	$[\alpha\|1\rangle - \beta\|0\rangle = ZX(\alpha\|0\rangle + \beta\|1\rangle)$

Now Alice uses a classical communication channel to send the results c_1c_2 to Bob who, to recover the unknown state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ has just to perform the trasformation: $X^{c_2}Z^{c_1}$. Thus, using this protocol we achieve a teletransportation of an unknown qubit state from Alice and Bob. Notice that this quantum protocols makes requires the use of an entangled pairs and is assisted by a classical communication channel.

As final remarks, let us notice that the teleportation protocol :
- does not violate the No Cloning Theorem, since the total protocol destroys the state in Alice's hands and therefore we have not duplicated (or copied) the unknown qubit state; - is not in contradiction with special relativity, because in the classical communication channel signals cannot be sent faster than light and therefore this is so also for the information that is stored in the teleported state.

The first experiment of quantum teleportation outside a lab has been performed by Zeilinger, teleporting photons from one side of the Danube to the other (approximately 800 m.). By now, the longest distance that have been reached is ground-satellite, of about 1400 km.

2.4 Density matrices

We have already introduced the concept of pure and mixed state. In this section we will discuss it further form a mathematical point of view and we will see how mixed states are necessary to discuss open quantum systems.

2.4.1 Pure states

We have already seen that every normalized state $|\psi\rangle \in \mathcal{H}$, $\langle\psi|\psi\rangle = 1$, can be associated to the projection operator $P_\psi = |\psi\rangle\langle\psi|$, which satisfies the following properties:

(i) P_ψ is bounded and, in particular, $||P_\psi|| = 1$

(ii) $P_\psi^\dagger = P_\psi$

(iii) P_ψ is positive: $(\langle\alpha|P_\psi|\alpha\rangle \geq 0)$, for every $|\alpha\rangle \in \mathcal{H}$

(iv) $Tr\left[P_\psi\right] = 1$

(v) $P_\psi^2 = P_\psi$

We have already seen that P_ψ verifies (ii) and (v), since it is a projection operator. As for (i), it is sufficient to observe that:

$$||P_\psi|\alpha\rangle||^2 = |\langle\psi|\alpha\rangle|^2 \;\; \forall|\alpha\rangle\langle\psi|\psi\rangle \leq \|\alpha\|^2$$

from which one can deduce $||P_\psi|| \leq 1$. Then, setting $|\alpha\rangle = |\psi\rangle$ in the previous expression, it can be proved that $||P_\psi|| = 1$. Regarding (iii), it is immediate to prove that :

$$\langle\alpha|P_\psi|\alpha\rangle = \langle\alpha|\psi\rangle\langle\psi|\alpha\rangle = |\langle\psi|\alpha\rangle|^2 \geq 0$$

Finally, to prove (iv), we choose an orthonormal basis $\{|e_n\rangle\}_n$ such that: $|e_1\rangle = |\psi\rangle$, thus $\langle\psi|e_n\rangle = 0$ if $n \neq 1$ and $\langle\psi|e_1\rangle = 1$. Hence we find

$$Tr\left[P_\psi\right] = \sum_n \langle e_n|P_\psi|e_n\rangle = \sum_n \langle e_n|\psi\rangle\langle\psi|e_n\rangle = \langle e_1|\psi\rangle\langle\psi|e_1\rangle = 1$$

This means that a pure state of a quantum system is associated to the matrix

$$\rho_\psi \equiv P_\psi = |\psi\rangle\langle\psi|$$

i.e. to a matrix that satisfies the properties from (i) to (v) listed above. It is called a *pure density matrix*.

The vice versa is also true: every density matrix ρ can be written as $|\psi_\rho\rangle\langle\psi_\rho|$ for a suitable normalized state $|\psi_\rho\rangle$. In fact, since ρ is bounded, self-adjoint and positive, it admits the spectral decomposition $\rho = \sum_n \lambda_n P_n$ with $P_n = |e_n\rangle\langle e_n|$, $|e_n\rangle$ being the eigenvector corresponding to the eigenvalue $\lambda_n \geq 0$. Moreover,

$$\rho^2 = \sum_n \lambda_n^2 P_n^2 = \sum_n \lambda_n^2 P_n \text{ is the same as } \rho = \sum_n \lambda_n P_n$$

if and only if $\lambda_n^2 = \lambda_n$, that is to say $\lambda_n = 0$ or $\lambda_n = 1$. However, in order to satisfy $Tr[\rho] = \sum_n \lambda_n = 1$ it must be that all λ_n but one, say λ_1, are equal to zero. Hence $\rho = |e_1\rangle\langle e_1|$.

We can therfore give the following alternative definition of a quantum state (first postulate):
a pure state of a quantum system is described by a density matrix ρ, namely a self-adjoint, positive, idempotent operator with unit norm and unit trace.
We can also easily rephrase the other three postulates.

An observable is still identified by a self-adjoint operator A. Now, its mean value on the state ρ_ψ is given by:

$$\langle A \rangle = Tr\left[\rho_\psi A\right]$$

In fact we have:

$$\begin{aligned}
\langle A \rangle &= \langle\psi|A|\psi\rangle = \langle\psi|A\left(\sum_m |e_m\rangle\langle e_m|\right)|\psi\rangle \\
&= \sum_m \langle\psi|A|e_m\rangle\langle e_m|\psi\rangle = \sum_m \langle e_m|\psi\rangle\langle\psi|A|e_m\rangle \\
&= \sum_m \langle e_m|\rho_\psi A|e_m\rangle = Tr\left[\rho_\psi A\right]
\end{aligned}$$

The time evolution of the density matrix ρ_ψ is given by the relation:

$$\rho_\psi(t) = e^{-\frac{\imath t}{\hbar}\hat{H}}\rho_\psi(0)e^{+\frac{\imath t}{\hbar}\hat{H}}$$

In fact:

$$\rho_\psi(t) = |\psi(t)\rangle\langle\psi(t)| = \left\{e^{-\frac{\imath t}{\hbar}\hat{H}}|\psi(0)\rangle\right\}\left\{\langle\psi(0)|e^{+\frac{\imath t}{\hbar}\hat{H}}\right\}$$

Finally, a *POVM* is described by a collection of operators $\{M_n\}_n$ indexed by all possible values α_n of the quantity that is being measured, satisfying the condition $\sum_n M_n^\dagger M_n = \mathbb{I}$ and such that:

i) the measure on a state $|\psi\rangle$ gives as a result the value α_n with a probability equal to
$$p_n = Tr\left[\rho_\psi M_n^\dagger M_n\right]$$

ii) immediately after the measurement, the system is described by the density matrix:
$$\widetilde{\rho} = \frac{1}{Tr\left[\rho_\psi M_n^\dagger M_n\right]} M_n \rho_\psi M_n^\dagger$$

In fact we have

$$p_n = \langle\psi|M_n^\dagger M_n|\psi\rangle = Tr\left[\rho_\psi\ M_n^\dagger M_n\right]$$

and

$$\begin{aligned}\widetilde{\rho} &= \frac{1}{\sqrt{\langle\psi|M_n^\dagger M_n|\psi\rangle}} M_m|\psi\rangle\langle\psi|M_n^\dagger \frac{1}{\sqrt{\langle\psi|M_n^\dagger M_n|\psi\rangle}} \\ &= \frac{1}{Tr\left[\rho_\psi M_n^\dagger M_n\right]} M_n\rho_\psi M_n^\dagger\end{aligned}$$

2.4.2 Mixed states

Sometimes we cannot establish a priori that a system is in a well-defined state, but we can only assert that it can be found in a state chosen from a set of states $\{|\psi_k\rangle\}_{k\in I}$ (compatible, for example, with the methodology used to prepare the system or with certain boundary conditions). The states $|\psi_k\rangle$, $k \in I$, are supposed to be normalized but aren't necessarily orthogonal. We assume that the system can be found in the state $|\psi_k\rangle$ with probability p_k. Clearly, the following conditions must hold:

$$\begin{aligned}&0 \leq p_k \leq 1 \\ &\sum_{k\in I} p_k = 1\end{aligned}$$

In that case, it is said that the system is in a *mixed state*.

The set $\{|\psi_k\rangle\,,\ p_k\}_{k\in I}$ is called *statistical ensemble* and defines a matrix, called density matrix of the mixed state, given by:

$$\rho = \sum_k p_k|\psi_k\rangle\langle\psi_k| = \sum_k p_k\rho_k$$

Similarly to the pure state case, it isn't hard to show that ρ satisfies the following properties:

(i) ρ is bounded

(ii) $\rho^\dagger = \rho$

(iii) ρ is positive

(iv) $Tr\left[\rho\right] = 1$

In general, however, $\rho^2 \neq \rho$. In fact the following theorem holds:

$$\rho^2 = \rho \Longleftrightarrow \rho = |\psi\rangle\langle|\psi| \text{ (that is, } \rho \text{ is a pure state)}$$

For simplicity, we will prove this theorem only in the case in which the states $|\psi_k\rangle$ are orthogonal. In this case:

$$\rho^2 = \sum_{k,k'} p_k p_{k'} |\psi_k\rangle\langle\psi_k|\psi_{k'}\rangle\langle\psi_{k'}| = \sum_k p_k^2 \rho_k$$

Thus, ρ^2 can be equal to $\rho = \sum_k p_k \rho_k$ if and only if the p_k are all zero except one, which must be equal to one (in other words, if ρ represents a pure state).

We leave to the reader the formulation of the postulates about observables, time evolution and measurement in the case of mixed states. Here we simply notice that the mean value of an observable A on a mixed state $\rho = \sum_k p_k \rho_k$ is defined as the average of the mean values of A on the pure states ρ_k weighted with the probabilities p_k, that is:

$$\langle A \rangle \equiv \sum_k p_k Tr\left[\rho_k A\right] = Tr\left[\rho A\right]$$

Remark. The decomposition of ρ on an ensemble of pure states is not, in general, unique. For example, consider the Hilbert space of a qubit and the density matrix:

$$\rho = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1|$$

representing a statistical ensemble in which the system is found in the state $|0\rangle$ with probability $p_0 = 3/4$ and in the state $|1\rangle$ with probability $p_1 = 1/4$. Let's now consider the density matrix:

$$\rho' = \frac{1}{2}|a\rangle\langle a| + \frac{1}{2}|b\rangle\langle b|$$

representing a statistical ensemble in which the system is found in the state $|a\rangle$ or in the state $|b\rangle$ with equal probabilities: $p_a = p_b = 1/2$. Apparently, it may seem that $\rho \neq \rho'$, but a little bit of algebra shows that by choosing:

$$\begin{cases} |a\rangle = \sqrt{\frac{3}{4}}|0\rangle + \sqrt{\frac{1}{4}}|1\rangle \\ |b\rangle = \sqrt{\frac{3}{4}}|0\rangle - \sqrt{\frac{1}{4}}|1\rangle \end{cases}$$

one finds $\rho = \rho'$.

MIXED STATES OF A QUBIT.
Let's consider a density matrix for a qubit: this a 2×2 matrix satisfying properties $(i) - (iv)$ above. In particular it must be a hermitean, trace-one matrix, so it can be written as:

$$\rho = \frac{1}{2} \begin{pmatrix} 1 + n_z & n_x - in_y \\ n_x + in_y & 1 - n_z \end{pmatrix} = \frac{\mathbb{I} + \vec{n} \cdot \vec{\sigma}}{2} \tag{2.14}$$

where $\vec{n} = (n_x, n_y, n_z)$. Now, ρ should also be positive: this means that its eigenvalues $\lambda_\pm = (1 \pm |\vec{n}|)/2 \geq 0$. i.e. $|\vec{n}| \leq 1$. If $|\vec{n}| \leq 1$ we have one zero and one unit eigenvalue: this means that the matrix ρ represents a pure state, while for $|\vec{n}| < 1$ we have a genuine mixed state.

We can conclude to say that the space of density matrices of a qubit is the volume of a two-dimensional sphere, the *Bloch sphere* $|\vec{n}| \leq 1$ and that we are on the surface/interior if the state is pure/mixed.

2.4.3 Density matrix for a composite system

Let's now consider a composite system: $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. A pure state is described by a density matrix of the type:

$$\rho^{AB} = |\psi^{AB}\rangle\langle\psi^{AB}| \ , \ \text{ with } |\psi^{AB}\rangle \in \mathcal{H}$$

It is important to notice that in general $\rho^{AB} \neq \rho^A \otimes \rho^B$, for suitable density matrices ρ_A, ρ_B defined on $\mathcal{H}_A, \mathcal{H}_B$ respectively. A little bit of algebra shows that the case $\rho^{AB} = \rho^A \otimes \rho^B$ occurs only when $|\psi^{AB}\rangle = |\psi^A\rangle \otimes |\psi^B\rangle$, that is, $|\psi^{AB}\rangle$ is separable. An entangled state is instead described by the matrix $\rho^{AB} = \sum p_k \rho_k^{AB}$ with ρ_k^{AB} defined as above.

Given ρ^{AB}, we define the *reduced density matrix* ρ^A for the subset A as the one obtained from ρ^{AB} after tracing out all the degrees of freedom of B:

$$\rho^A \equiv Tr_{\mathcal{H}_B}\left[\rho^{AB}\right] \tag{2.15}$$

where the trace on $\mathcal{H}_B$ of matrices of the type $\rho^{AB} = \rho^A \otimes \rho^B$ is given by

$$Tr_{\mathcal{H}_B}\left[\rho^{AB}\right] = \rho^A Tr_{\mathcal{H}_B}\left[\rho^B\right]$$

This can then be extended to any density matrix by using the linearity property.

Example Consider a system of two qubits and take $|\psi^{AB}\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$, one of Bell's states, that defines the density matrix:

$$\rho^{AB} = |\psi^{AB}\rangle\langle\psi^{AB}| = \frac{|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|}{2}$$

Then

$$\begin{aligned}\rho^A = Tr_{\mathcal{H}_B}\left[\rho^{AB}\right] &= \frac{1}{2}\left(|0\rangle\langle 0|Tr_{\mathcal{H}_B}\left[|0\rangle\langle 0|\right] + |0\rangle\langle 1|Tr_{\mathcal{H}_B}\left[|0\rangle\langle 1|\right]\right. \\ &+ \left.|1\rangle\langle 0|Tr_{\mathcal{H}_B}\left[|1\rangle\langle 0|\right] + |1\rangle\langle 1|Tr_{\mathcal{H}_B}\left[|1\rangle\langle 1|\right]\right) \\ &= \frac{1}{2}\left(\ |0\rangle\langle 0|\langle 0|0\rangle + |0\rangle\langle 1|\langle 0|1\rangle\right. \\ &+ \left.|1\rangle\langle 0|\langle 1|0\rangle + |1\rangle\langle 1|\langle 1|1\rangle\ \right) = \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2}\end{aligned}$$

We notice that the state of the composite system AB is pure, while the reduced state on A, obtained by tracing away the degrees of freedom related to B, is mixed. This is due to the fact that the state of the composite system is entangled.

The concept of a reduced density matrix arises naturally when we want to perform a measurement related only to the degrees of freedom of one of the two subsets., i.e. whenever we are considering an open quantum system: in this case A represents the system we want to study and B the environment. The above considerations are showing that, whenever the system and the environment are entangled, the system is described by a mixed density state.

Chapter 3

Concepts of Quantum Computation and Information

This Chapter contains only an outline of the topics seen in class.
The full treatment can be found in Sect.ns 1.2 through 1.4 and in Sect.ns 8.1 through 8.3 of [NC].

In this chapter we will discuss the notion of quantum circuit and introduce some notable examples.
We will discuss then quantum channels.

OUTLINE

- *Quantum circuits* are composed by qubits that store information, by quantum gates that represent the possible unitary tranformation we can perform on qubit to manipulate information, measurements to read the processed information.

- *Examples of quantum gates*:
  - truth table and matrix representation;
  - single qubit gates and in particular the X,Y,Z,H transformations;
  - examples of two qubit gates and in particular the CNOT and other controlled gates;
  - examples of multiple qubit gates: the Toffoli gate.

- *Examples of algorithms*:
  - SWAP operation;
  - preparation of Bell states and GHZ states;

- teleportation circuit;
- quantum parallelism and the Deutsch-Josza algorithm.

- Why and how it is possible to *perform classical computation* on a quantum computer.

- Open systems and general quantum operations or *quantum channels*; the operator sum representation.

- *Important examples* of quantum channels:
  - the bit and the phase flip channels;
  - depolarizing, amplitude and phase damping channels.

INTERLUDES (not compulsory: no questions in exam)

- *Universality of quantum computation*; only general ideas - no proofs (Sect. 4.5 of [NC])

- *The IBM Qiskit Tools* (Links available in Virtuale)

- *Cryptography* and quantum key distributions (Slides available in Virtuale)

Appendices

Appendix A: Dirac notation

In this appendix we will be reviewing the Dirac notation which is widely used in these notes.

Let $\mathcal{H}$ be a separable, finite- or infinite-dimensional Hilbert space. We will denote its scalar product with $\langle\cdot|\cdot\rangle$ and its vectors with the *ket* symbol $|v\rangle$.

Upon choosing an orthonormal basis, $\{|e_n\rangle\}_{n=1}^{\infty}$ and $\langle e_n|e_m\rangle = \delta_{nm}$, we can represent the vector $|v\rangle$ as a column vector:

$$|v\rangle = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \\ \vdots \end{pmatrix}$$

where v_n are components of $|v\rangle$ in this particular orthonormal basis.
To each vector $|w\rangle \in \mathcal{H}$ we can associate a vector from the dual space $\mathcal{H}^*$, called *bra*. It is denoted with the symbol $\langle w|$ and is defined by the relation $\langle w| = (|w\rangle)^\dagger$. In the dual basis $\langle e_m|$ it is represented by the row vector:

$$(w_1^* \, w_2^* \cdots w_n^* \cdots)$$

The scalar product between two vectors $|v\rangle, |w\rangle$ is therefore performed like a matrix multiplication:

$$\langle w|v\rangle = \begin{pmatrix} w_1^* & w_2^* & \cdots & w_n^* & \cdots \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \\ \vdots \end{pmatrix} = \sum_n w_n^* v_n$$

We notice that for $|v\rangle$ to be a vector of $\mathcal{H}$ it must satisfy the condition $\|v\|^2 = \langle v|v\rangle = \sum_n |v_n|^2 < \infty$.

Lastly we observe that the operator P_v projecting onto the one-dimensional subspace generated by $|v\rangle$ is given by the matrix:

$$P_v = \frac{|v\rangle\langle v|}{\langle v|v\rangle} = \frac{1}{||v||^2} \begin{pmatrix} v_1^* v_1 & v_1^* v_2 & \cdots & v_1^* v_n & \cdots \\ v_2^* v_1 & v_2^* v_2 & \cdots & v_2^* v_n & \cdots \\ \vdots & \vdots & \cdots & \vdots & \cdots \end{pmatrix}$$

Appendix B: Projection operators

In this appendix we briefly recall the properties of orthogonal projection operators.

Given a closed subspace $M \in \mathcal{H}$, we decompose the Hilbert space $\mathcal{H}$ in the direct orthogonal sum of M and his orthogonal complement $M^\perp$: $\mathcal{H} = M \oplus_\perp M^\perp$. In this way every vector $|v\rangle \in \mathcal{H}$ can be decomposed in $|v\rangle = |m\rangle + |m'\rangle$ with $|m\rangle \in M$, $|m'\rangle \in M^\perp$.

We will denote with P_M to the operator orthogonally projecting onto the subspace M:

$$P_M : |v\rangle \mapsto |m\rangle$$

We recall that an operator P can be considered a projection operator if and only if it is self-adjoint and idempotent.

$$P^\dagger = P \ , \ \ P^2 = P \tag{3.1}$$

Since the projection is orthogonal, the operator $\mathbb{I} - P_M$ is equal to the projection operator $P_{M^\perp}$, which projects onto $M^\perp$. We also notice the following equality: $P_M P_{M^\perp} = P_{M^\perp} P_M = 0$.

In general, if M_1, M_2 are two orthogonal subspaces ($M_1 \perp M_2$), the corresponding two projections P_1 and P_1 are orthogonal and: $P_1 P_2 = P_2 P_1 = 0$. As a consequence, it is easy to prove that the result of the operation $P_1 + P_2$ is also a projection operator, which projects onto the subspace $M_1 \oplus_\perp M_2$. Given an orthonormal base $\{|m_j\rangle\}_{j=1}^{N_M}$ for M, it is not hard to verify that its projector P_M is obtained by summing all the projectors $P_j = |m_j\rangle\langle m_j|$, on the one-dimensional subspaces M_j spanned by the basis vectors $|m_j\rangle$:

$$P_M = \sum_{j=1}^{N_M} |m_j\rangle\langle m_j| \tag{3.2}$$

In particular if $M = \mathcal{H}$ and $\{|e_n\rangle\}_{n=1}^N$ is an orthonormal basis, we have:

$$\sum_{n=1}^{N} |e_n\rangle\langle e_n| = P_{\mathcal{H}} = \mathbb{I} \tag{3.3}$$

Appendix C: Trace of an operator

If $\mathcal{H}$ is a finite dimensional Hilbert space we can define the trace of any operator A in the following way. Given an orthonormal basis $\{|e_n\rangle\}_n$ we can construct the matrix $(N \times N)$ which represents A:

$$M_{nm}^{(A)} = \langle e_n|A|e_m\rangle$$

and set

$$Tr\,[A] = \sum_n M_{nn} = \sum_n \langle e_n|A|e_n\rangle \tag{3.4}$$

The following properties (linearity and cyclicity) are satisfied[1]:

$$Tr\,[\lambda A_1 + A_2] = \lambda Tr\,[A_1] + Tr\,[A_2]\ \ , \lambda \in \mathbb{C}$$
$$Tr\,[A_1 A_2 \cdots A_n] = Tr\,[A_n A_1 \cdots A_{n-1}]$$

Exercise: $Tr[A]$ is independent of the o.n. basis used to compute it.

As an important example, consider an operator of the kind $A = |\psi\rangle\langle\varphi|$. Then:

$$Tr[A] = \sum_n \langle e_n|\psi\rangle\langle\varphi|e_n\rangle = \sum_n \langle\varphi|e_n\rangle\langle e_n|\psi\rangle = \langle\varphi|\left(\sum_n |e_n\rangle\langle e_n|\right)|\psi\rangle = \langle\varphi|\psi\rangle$$

In particular:

$$Tr[P_\psi] = Tr[|\psi\rangle\langle\psi|] = \langle\psi|\psi\rangle$$

Appendix D: Tensor product

Given two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ (for our purposes it is enough to consider finite-dimensional ones), we start by considering two orthonormal

[1] In an infinite dimensional Hilbert space, the series which defines the trace (3.4) may not be convergent. An operator A whose series is absolutely convergent, and such that its trace is well defined, is said to be *trace-class*. The linearity and cyclicity properties are still valid.

basis: $\{|e_j\rangle\}_j$ in $\mathcal{H}_1$ and $\{|f_k\rangle\}_k$ in $\mathcal{H}_2$. Then we define the tensor product Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ as the vector space generated by the basis $\{|e_j\rangle \otimes |f_k\rangle \equiv |e_j f_k\rangle\}_{jk}$ endowed by the scalar product:

$$\langle \psi\phi|\psi'\phi'\rangle_{\mathcal{H}} \equiv \langle \psi|\psi'\rangle_{\mathcal{H}_1}\langle \phi|\phi'\rangle_{\mathcal{H}_2}$$

on all vectors in $\mathcal{H}$ of the type $|\psi\phi\rangle$ with $|\psi\rangle \in \mathcal{H}_1$, $|\phi\rangle \in \mathcal{H}_2$ that can then be extended to any vector by linearity.
The dimension of the total space $\mathcal{H}$ is $d = MN$ if the dimensions of $\mathcal{H}_{1,2}$ are M and N. A generic vector in $\mathcal{H}$ is of the form:

$$|v\rangle = \sum_{j=1}^{M}\sum_{k=1}^{N} c_{jk}|e_j f_k\rangle\}_{jk}$$

and $\|v\|^2 = \langle v|v\rangle = \sum_{j=1}^{M}\sum_{k=1}^{N}|c_{jk}|^2$.

Given two operators A, B defined on $\mathcal{H}_{1,2}$ respectively, we can define the operator $A \otimes B$ defined on the total space $\mathcal{H}$ as:

$$(A \otimes B)|\psi\phi\rangle = A|\psi\rangle \otimes B|\phi\rangle$$

on any product vector, which is then extended to any vector by linearity. Operators of the kind $A \otimes B$ generate the whole algebra of operators of the total space, in the sense that a generic oprator O on H can be written as: $O = \sum_{\alpha\beta} A_\alpha B_\beta$ for suitable operators A_α, B_β defined on $\mathcal{H}_{1,2}$ respectively.

By induction, this definition can be extended to the tensor product of an arbitrary finite number of Hilbert spaces. In particular we will be interested in the tensor product of N copies of the same Hilbert space: $\mathcal{H}_t = \mathcal{H}^{\otimes N}$. We remark that the tensor product of operators is:
- non commutative: $A \otimes B \neq B \otimes A$;
- but associative: $(A \otimes B) \otimes C = A \otimes (B \otimes C) = A \otimes B \otimes C$.