Contents

1. Mathematical Background

1.1. Complex numbers

In mathematics, a **complex number** is an element of a number system that extends the real numbers with a specific element denoted i, called the **imaginary unit** and satisfying the equation $i^2 = -1$. The set of complex numbers is denoted by the symbol $\mathbb{C}$.

Every complex number z can be expressed in the form $a + bi$, where a and b are real numbers and are referred to as its **real part** and its **imaginary part**, respectively. The real part of a complex number z is denoted $\Re(z)$, the imaginary part $\Im(z)$. A complex number with imaginary part equal to 0 is simply a real number; a complex number with real part equal to 0 is said to be a **purely imaginary** number.

Addition, subtraction and multiplication of complex numbers can be naturally defined by using the rule $i^2 = -1$ along with the associative, commutative, and distributive laws.

A complex number z can be identified with the ordered pair of real numbers $(\Re(z), \Im(z))$, which may be interpreted as coordinates of a point in a Euclidean plane with standard coordinates, which is then called the **complex plane** or **Argand diagram**. The horizontal axis is generally used to display the real part, with increasing values to the right, and the imaginary part marks the vertical axis, with increasing values upwards.

Given a complex number $z = a + ib$, the **complex conjugate** of z is the number $z^* = a - ib$, obtained by changing the sign of the imaginary part of z. Geometrically, z is the "reflection" of z about the real axis. It is trivial to see that, for any complex number z, $(z^*)^* = z$. A complex number is real if and only if it equals its own conjugate.

The square root of the product between a complex number z and its complex conjugate z^* is a non negative real number called **modulus** or **magnitude**:

$$|z| = \sqrt{z \cdot z^*} = \sqrt{(\Re(z) + i\Im(z))(\Re(z) - i\Im(z))} = \sqrt{\Re(z)^2 + \Im(z)^2}$$

By Pythagoras' theorem, $|z|$ is the distance from the origin to the point representing the complex number z in the complex plane.

The **argument** of z (sometimes called the "phase" φ), denoted as $\arg(z)$, is the angle formed by the vector $(\Re(z), \Im(z))$ with the positive real axis in the complex plane:

$$\arg(z) = \tan^{-1}\left(\frac{\Im(z)}{\Re(z)}\right)$$

Note that any rotation of $2k\pi$ with $k \in \mathbb{Z}$ is equivalent to performing no rotation at all, therefore the argument is often required to be specified in the interval $(-\pi, \pi]$.

A complex number $z = a + ib$ is said to be written in **rectangular form**, or **algebraic form**. Another way to express a complex number is the **polar form**; given a complex number z with modulus $r = |z|$ and argument $\varphi = \arg(z)$, the polar form of z is:

$$z = r(\cos(\varphi) + i\sin(\varphi))$$

A third way to express complex numbers is the **exponential form**:

$$z = re^{i\varphi}$$

Where the complex exponential $e^{i\varphi}$ is also referred to as the **phase factor**.

The fourth way to express a complex number is the **matrix form.** Given a complex number $z = a + ib$, it can be written as a 2×2 matrix as follows:

$$a + ib = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

The complex conjugate of a complex number in matrix form is simply its matrix transpose.

It is also possible to transition immediately from the exponential form to the matrix form:

$$re^{i\varphi} = r\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = r\cos(\theta)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + r\sin(\theta)\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Which also means that the real unit and the imaginary unit are simply:

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad\qquad i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

This representation is consistent with respect to standard complex number operations. For example, to show that $zz^* = a^2 + b^2$:

$$zz^* = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} a & -b \\ b & a \end{pmatrix}^T = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} aa + (-b)(-b) & ab + (-b)a \\ ba + a(-b) & bb + aa \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & \cancel{ab} - \cancel{ab} \\ \cancel{ab} - \cancel{ab} & a^2 + b^2 \end{pmatrix} =$$
$$= \begin{pmatrix} a^2 + b^2 & 0 \\ 0 & a^2 + b^2 \end{pmatrix} = (a^2 + b^2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = a^2 + b^2$$

The inverse of a matrix representing a complex number is the reciprocal of the number itself. Given a complex number $z = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$:

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}^{-1} = \frac{1}{a^2 + b^2}\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \frac{\cancel{z^*}}{z\cancel{z^*}} = \frac{1}{z}$$

1.2. Hilbert spaces

Given a vector space V and two vectors $\boldsymbol{x}, \boldsymbol{y} \in V$, their **inner product** $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ is given by:

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^\dagger \boldsymbol{y} = (x_1^* \ \dots \ x_n^*)\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} x_i^* y_i$$

In the context of quantum mechanics, vector spaces are assumed to be on the field $\mathbb{C}$. In this context, the inner product is also referred to as the **scalar product** (because the value returned is a single number) or **dot product** (because it is sometimes denoted with a dot), even though in general the scalar/dot product is a special case of inner product.

The inner product is a mathematical operation that satisfies (at least) this three properties:

- Invariance with respect to conjugation: $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle^*$;
- Linearity in the second position: $\langle \boldsymbol{x}, \alpha\boldsymbol{y} + \beta\boldsymbol{z} \rangle = \alpha\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \beta\langle \boldsymbol{x}, \boldsymbol{y} \rangle$;
- Antilinearity in the first position: $\langle \alpha\boldsymbol{x} + \beta\boldsymbol{y}, \boldsymbol{z} \rangle = \alpha^*\langle \boldsymbol{x}, \boldsymbol{z} \rangle + \beta^*\langle \boldsymbol{y}, \boldsymbol{z} \rangle$;
- $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$ for any $\boldsymbol{x} \in \mathbb{C}$;
- $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ if $\boldsymbol{x} = \boldsymbol{0}$.

The square root of the inner product of a vector $\boldsymbol{x}$ with itself is called the **norm** of the vector:

$$||\boldsymbol{x}|| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} = \sqrt{(x_1^* \ \dots \ x_n^*)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}$$

Any vector space that possesses an inner product is called an **Hilbert space**. Hilbert spaces are so obiquitous that, when not specified, any vector space is assumed to be an Hilbert space.

In quantum mechanics have particular relevance bases of vector spaces that are **orthonormal**, meaning that all vectors of the basis are orthogonal to each other (except for itself) and have norm equal to 1. Given a vector space V and a basis $B = \boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ for V, The set B forms an orthonormal basis if:

$$\forall i, j \in \{1, \dots, n\}, \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \begin{cases} 0 \text{ if } i \neq j \\ 1 \text{ if } i = j \end{cases}$$

Finding an orthonormal basis for a vector space can be tedious. However, if a generic (non orthonormal) basis is known for a certain vector space, it is possible to apply an algorithm called **Gram-Schmidt procedure** that can construct an orthonormal basis from the known generic basis.

Let V be a vector space and let $B = \boldsymbol{v}_1, ..., \boldsymbol{v}_n$ be a basis for V. Let $C = \boldsymbol{w}_1, ..., \boldsymbol{w}_n$ be the orthonormal basis that the Gram-Schmidt procedure will construct. The procedure is as follows: the first vector $\boldsymbol{w}_1$ is given by the ratio $\boldsymbol{v}_1 / ||\boldsymbol{v}_1||$, whereas for any $1 \leq k \leq n$ the vector $\boldsymbol{w}_{k+1}$ is defined inductively:

$$\boldsymbol{w}_{k+1} = \frac{\boldsymbol{v}_{k+1} - \sum_{i=1}^{k} \langle \boldsymbol{w}_i, \boldsymbol{v}_{k+1} \rangle \boldsymbol{w}_i}{\left|\left| \boldsymbol{v}_{k+1} - \sum_{i=1}^{k} \langle \boldsymbol{w}_i, \boldsymbol{v}_{k+1} \rangle \boldsymbol{w}_i \right|\right|}$$

Since the Gram-Schmidt decomposition always works, for any vector space always exist at least an orthonormal basis.

Exercise 1.2.1: Consider the vector space $\mathbb{C}^n$ having basis $B = \{(1,0,0), (0,2i,i+1), (1,0,3)\}$. Construct an orthonormal basis C for $\mathbb{C}^n$ from B.

Solution: The first element of C is simply $(1,0,0)/||1,0,0|| = (1,0,0)$. The second element:

$$\boldsymbol{w}_2 = \frac{\boldsymbol{v}_2 - \sum_{i=1}^{1} \langle \boldsymbol{w}_i, \boldsymbol{v}_2 \rangle \boldsymbol{w}_i}{\left|\left| \boldsymbol{v}_2 - \sum_{i=1}^{1} \langle \boldsymbol{w}_i, \boldsymbol{v}_2 \rangle \boldsymbol{w}_i \right|\right|} = \frac{\boldsymbol{v}_2 - \langle \boldsymbol{w}_1, \boldsymbol{v}_2 \rangle \boldsymbol{w}_1}{||\boldsymbol{v}_2 - \langle \boldsymbol{w}_1, \boldsymbol{v}_2 \rangle \boldsymbol{w}_1||} = \frac{\begin{pmatrix} 0 \\ 2i \\ i+1 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2i \\ i+1 \end{pmatrix} \right\rangle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}{\left|\left| \begin{pmatrix} 0 \\ 2i \\ i+1 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2i \\ i+1 \end{pmatrix} \right\rangle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right|\right|} =$$

$$= \frac{\begin{pmatrix} 0 \\ 2i \\ i+1 \end{pmatrix} - 0 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}{\left|\left| \begin{pmatrix} 0 \\ 2i \\ i+1 \end{pmatrix} - 0 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right|\right|} = \frac{\begin{pmatrix} 0 \\ 2i \\ i+1 \end{pmatrix}}{\left|\left| \begin{pmatrix} 0 \\ 2i \\ i+1 \end{pmatrix} \right|\right|} = \frac{\begin{pmatrix} 0 \\ 2i \\ i+1 \end{pmatrix}}{\sqrt{0^* \cdot 0 + (2i)^* \cdot 2i + (i+1)^* \cdot (i+1)}} = \begin{pmatrix} 0 \\ \frac{2i}{\sqrt{6}} \\ \frac{i+1}{\sqrt{6}} \end{pmatrix}$$

As for the third element:

$$\boldsymbol{w}_3 = \frac{\boldsymbol{v}_3 - \sum_{i=1}^{2} \langle \boldsymbol{w}_i, \boldsymbol{v}_3 \rangle \boldsymbol{w}_i}{\left|\left| \boldsymbol{v}_3 - \sum_{i=1}^{2} \langle \boldsymbol{w}_i, \boldsymbol{v}_3 \rangle \boldsymbol{w}_i \right|\right|} = \frac{\boldsymbol{v}_3 - \langle \boldsymbol{w}_1, \boldsymbol{v}_3 \rangle \boldsymbol{w}_1 - \langle \boldsymbol{w}_2, \boldsymbol{v}_3 \rangle \boldsymbol{w}_2}{||\boldsymbol{v}_3 - \langle \boldsymbol{w}_1, \boldsymbol{v}_3 \rangle \boldsymbol{w}_1 - \langle \boldsymbol{w}_2, \boldsymbol{v}_3 \rangle \boldsymbol{w}_2||} =$$

$$= \frac{\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \right\rangle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \left\langle \begin{pmatrix} 0 \\ \frac{2i}{\sqrt{6}} \\ \frac{i+1}{\sqrt{6}} \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \right\rangle \begin{pmatrix} 0 \\ \frac{2i}{\sqrt{6}} \\ \frac{i+1}{\sqrt{6}} \end{pmatrix}}{\left|\left| \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \right\rangle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \left\langle \begin{pmatrix} 0 \\ \frac{2i}{\sqrt{6}} \\ \frac{i+1}{\sqrt{6}} \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \right\rangle \begin{pmatrix} 0 \\ \frac{2i}{\sqrt{6}} \\ \frac{i+1}{\sqrt{6}} \end{pmatrix} \right|\right|} =$$

$$= \frac{\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \left(0^* \cdot 1 + \left(\frac{2i}{\sqrt{6}}\right)^* \cdot 0 + \left(\frac{i+1}{\sqrt{6}}\right)^* \cdot 3\right) \begin{pmatrix} 0 \\ \frac{2i}{\sqrt{6}} \\ \frac{i+1}{\sqrt{6}} \end{pmatrix}}{\left|\left| \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \left(0^* \cdot 1 + \left(\frac{2i}{\sqrt{6}}\right)^* \cdot 0 + \left(\frac{i+1}{\sqrt{6}}\right)^* \cdot 3\right) \begin{pmatrix} 0 \\ \frac{2i}{\sqrt{6}} \\ \frac{i+1}{\sqrt{6}} \end{pmatrix} \right|\right|} = \frac{\begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} + \left(\frac{3i-3}{\sqrt{6}}\right) \begin{pmatrix} 0 \\ \frac{2i}{\sqrt{6}} \\ \frac{i+1}{\sqrt{6}} \end{pmatrix}}{\left|\left| \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} + \left(\frac{3i-3}{\sqrt{6}}\right) \begin{pmatrix} 0 \\ \frac{2i}{\sqrt{6}} \\ \frac{i+1}{\sqrt{6}} \end{pmatrix} \right|\right|} =$$

$$= \frac{\begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} - \begin{pmatrix} 0 \\ 1+i \\ 1 \end{pmatrix}}{\left|\left| \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} - \begin{pmatrix} 0 \\ 1+i \\ 1 \end{pmatrix} \right|\right|} = \frac{\begin{pmatrix} 0 \\ -1-i \\ 2 \end{pmatrix}}{\left|\left| \begin{pmatrix} 0 \\ -1-i \\ 2 \end{pmatrix} \right|\right|} = \frac{\begin{pmatrix} 0 \\ -1-i \\ 2 \end{pmatrix}}{\sqrt{0^* \cdot 0 + (-1-i)^*(-1-i) + 2^* \cdot 2}} = \begin{pmatrix} 0 \\ \frac{-1-i}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix}$$

Obtaining the set $C = \left\{(1,0,0), \left(0, \frac{2i}{\sqrt{6}}, \frac{i+1}{\sqrt{6}}\right), \left(0, \frac{-1-i}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right)\right\}$ ∎

The **direct sum** of two vectors spaces V and W having bases $A = \{\boldsymbol{\alpha_1}, ..., \boldsymbol{\alpha_n}\}$ and $B = \{\boldsymbol{\beta_1}, ..., \boldsymbol{\beta_m}\}$ respectively, denoted as $V \oplus W$, is the vector space spanned by the basis $A \cup B = \{\boldsymbol{\alpha_1}, ..., \boldsymbol{\alpha_n}, \boldsymbol{\beta_1}, ..., \boldsymbol{\beta_m}\}$.

Every element $\boldsymbol{x} \in V \oplus W$ can be written as $\boldsymbol{v} \oplus \boldsymbol{w}$, for some $\boldsymbol{v} \in V$ and some $\boldsymbol{w} \in W$. The dimension of $V \oplus W$ is simply given by $\dim(V) + \dim(W)$.

Addition and scalar multiplication are defined by performing the operation on the two component vector spaces separately and adding the results. When V and W are inner product spaces, the standard inner product on $V \oplus W$ is given by:

$$(\boldsymbol{v_2}^\dagger \oplus \boldsymbol{w_2}^\dagger)(\boldsymbol{v_1} \oplus \boldsymbol{w_1}) = \langle \boldsymbol{v_2}, \boldsymbol{v_1} \rangle + \langle \boldsymbol{w_2}, \boldsymbol{w_1} \rangle$$

The vector spaces V and W embed in $V \oplus W$ in the obvious canonical way, and the images are orthogonal under the standard inner product.

Theorem 1.2.1 (Direct sum decomposition): For any finite inner product vector space S of dimension n, there exist a set of orthogonal subspaces $\{V_1, ..., V_k\}$ for some $k \leq n$ such that $S = V_1 \oplus ... \oplus V_k$

The **tensor product** between two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ is the vector $\boldsymbol{v} \otimes \boldsymbol{w}$ constructed as:

$$\boldsymbol{v} \otimes \boldsymbol{w} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \otimes \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} = \begin{pmatrix} v_1 \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} \\ \vdots \\ v_n \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} \end{pmatrix} = \begin{pmatrix} v_1 w_1 \\ \vdots \\ v_1 w_m \\ \vdots \\ v_n w_1 \\ \vdots \\ v_n w_m \end{pmatrix}$$

The tensor product has the following properties:

- Is bilinear: $(\boldsymbol{v_1} + \boldsymbol{v_2}) \otimes \boldsymbol{v_3} = \boldsymbol{v_1} \otimes (\boldsymbol{v_2} + \boldsymbol{v_3}) = \boldsymbol{v_1} \otimes \boldsymbol{v_2} + \boldsymbol{v_1} \otimes \boldsymbol{v_3}$
- Is associative: $(a\boldsymbol{v_1}) \otimes \boldsymbol{v_2} = \boldsymbol{v_1} \otimes (a\boldsymbol{v_2}) = a(\boldsymbol{v_1} \otimes \boldsymbol{v_2})$
- Is multiplicative: $(\boldsymbol{v_1} \otimes ... \otimes \boldsymbol{v_n})(\boldsymbol{w_1} \otimes ... \otimes \boldsymbol{w_n}) = \boldsymbol{v_1}\boldsymbol{w_1} \otimes ... \otimes \boldsymbol{v_n}\boldsymbol{w_n}$

Exercise 1.2.2: What is the tensor product of the two following vectors?

$$\boldsymbol{v} = \begin{pmatrix} 1 \\ 7 \end{pmatrix} \qquad \boldsymbol{w} = \begin{pmatrix} 3 \\ 10 \end{pmatrix}$$

Solution:

$$\boldsymbol{v} \otimes \boldsymbol{w} = \begin{pmatrix} 1 \cdot 3 \\ 1 \cdot 10 \\ 7 \cdot 3 \\ 7 \cdot 10 \end{pmatrix} = \begin{pmatrix} 3 \\ 10 \\ 21 \\ 70 \end{pmatrix}$$

□

Given two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$, the tensor product $\boldsymbol{v} \otimes \boldsymbol{w}^\dagger$ is also referred to as the **outer product** between $\boldsymbol{v}$ and $\boldsymbol{w}$. The result of the outer product between two vectors is a matrix:

$$\boldsymbol{v} \otimes \boldsymbol{w}^\dagger = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \otimes \begin{pmatrix} w_1^* & \dots & w_m^* \end{pmatrix} = \begin{pmatrix} v_1 \begin{pmatrix} w_1^* & \dots & w_m^* \end{pmatrix} \\ \vdots \\ v_n \begin{pmatrix} w_1^* & \dots & w_m^* \end{pmatrix} \end{pmatrix} = \begin{pmatrix} v_1 w_1^* & \dots & v_1 w_m^* \\ \vdots & \ddots & \vdots \\ v_n w_1^* & \dots & v_n w_m^* \end{pmatrix}$$

The tensor product can be extended to matrices, generating a block matrix:

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \otimes \boldsymbol{B} = \begin{pmatrix} a_{1,1}\boldsymbol{B} & a_{1,2}\boldsymbol{B} & \dots & a_{1,n}\boldsymbol{B} \\ a_{2,1}\boldsymbol{B} & a_{2,2}\boldsymbol{B} & \dots & a_{2,n}\boldsymbol{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1}\boldsymbol{B} & a_{m,2}\boldsymbol{B} & \dots & a_{m,n}\boldsymbol{B} \end{pmatrix}$$

By definition, $\dim(\boldsymbol{A} \otimes \boldsymbol{B}) = \dim(A)\dim(B)$. It should be noted that the tensor product between matrices, like the row-column product, is not commutative, but unlike the row-column product it requires no assumption on the dimension of the matrices to be defined.

The tensor product of two vector spaces V and W, having bases $A = \{\boldsymbol{v_1}, \boldsymbol{v_2}, ..., \boldsymbol{v_n}\}$ and $B = \{\boldsymbol{w_1}, \boldsymbol{w_2}, ..., \boldsymbol{w_m}\}$ respectively, is a nm-dimensional vector space denoted as $V \otimes W$.

The basis of $V \otimes W$ is constituted by all of possible tensor products between the vectors of the two original bases. Explicitly, the basis of $V \otimes W$ is:

$$\{\boldsymbol{v_1} \otimes \boldsymbol{w_1}, ..., \boldsymbol{v_n} \otimes \boldsymbol{w_1}, ..., \boldsymbol{v_1} \otimes \boldsymbol{w_m}, ..., \boldsymbol{v_n} \otimes \boldsymbol{w_m}\}$$

With $\boldsymbol{v_1}, ..., \boldsymbol{v_n} \in A$ and $\boldsymbol{w_1}, ..., \boldsymbol{w_m} \in B$.

This means that any generic vector of $V \otimes W$ can be written as:

$$\lambda_{1,1}(\boldsymbol{v_1} \otimes \boldsymbol{w_1}) + ... + \lambda_{n,1}(\boldsymbol{v_n} \otimes \boldsymbol{w_1}) + ... + \lambda_{1,m}(\boldsymbol{v_1} \otimes \boldsymbol{w_m}) + ... + \lambda_{n,m}(\boldsymbol{v_n} \otimes \boldsymbol{w_m})$$

For nm coefficients $\lambda_{i,j}$ with $i \in \{1, ..., n\}$ and $j \in \{1, ..., m\}$.

If V and W are two vector spaces for whose an inner product is defined (like Hilbert spaces), then it is possible to define an inner product for $V \otimes W$ as the product of the inner products with respect to those spaces.

The tensor product has many properties that are of interest for quantum state analysis.

Lemma 1.2.1: The tensor product of two unit vectors is also a unit vector.

Lemma 1.2.2: Let V and W be two vector spaces having bases $A = \{\boldsymbol{v_1}, \boldsymbol{v_2}, ..., \boldsymbol{v_n}\}$ and $B = \{\boldsymbol{w_1}, \boldsymbol{w_2}, ..., \boldsymbol{w_m}\}$ respectively. If A and B are both orthonormal, then the basis:

$$C = \{\boldsymbol{v_1} \otimes \boldsymbol{w_1}, ..., \boldsymbol{v_n} \otimes \boldsymbol{w_1}, ..., \boldsymbol{v_1} \otimes \boldsymbol{w_m}, ..., \boldsymbol{v_n} \otimes \boldsymbol{w_m}\}$$

Of $V \otimes W$ is also orthonormal.

1.3. Operators

The **conjugate transpose** of a matrix A, denoted as $A^\dagger$, is the matrix obtained from transposing A and then applying complex conjugation to each element of the resulting matrix:

$$A^\dagger = A^{T^*} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}^{T^*} = \begin{pmatrix} a^*_{1,1} & a^*_{2,1} & \dots & a^*_{m,1} \\ a^*_{1,2} & a^*_{2,2} & \dots & a^*_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ a^*_{1,n} & a^*_{2,n} & \dots & a^*_{m,n} \end{pmatrix}$$

Exercise 1.3.1: What is the conjugate transpose of the following matrix?

$$A = \begin{pmatrix} 1 & -2-i & 5 \\ 1+i & i & 4-2i \end{pmatrix}$$

Solution:

$$A^{\dagger} = \begin{pmatrix} 1 & -2-i & 5 \\ 1+i & i & 4-2i \end{pmatrix}^{\dagger} = \begin{pmatrix} 1 & 1+i \\ -2-i & i \\ 5 & 4-2i \end{pmatrix}^{*} = \begin{pmatrix} 1 & 1-i \\ -2+i & -i \\ 5 & 4+2i \end{pmatrix}$$

□

A square matrix A is said to be **Hermitian** if $A^{\dagger} = A$. It is said to be **unitary** if $A^{\dagger} = A^{-1}$.

Lemma 1.3.1: Let M be a generic matrix. Then $M^{\dagger} M$ is Hermitian.

Lemma 1.3.2: Let A be a matrix and let $\boldsymbol{v}, \boldsymbol{w}$ two vectors. If A is Hermitian, then $\langle \boldsymbol{w}, A\boldsymbol{v} \rangle = \langle A\boldsymbol{w}, \boldsymbol{v} \rangle$.

Theorem 1.3.1: An Hermitian matrix has all real eigenvalues.

Proof: Let $O : V \to V$ be a linear self-ajoint operator in matrix representation, and let λ be one of its eigenvalues. Recall that, given an eigenvector $\boldsymbol{v} \in V$ associated to λ, $O\boldsymbol{v} = \lambda\boldsymbol{v}$. Consider $\lambda\langle \boldsymbol{x}, \boldsymbol{x} \rangle$:

$$\lambda\langle \boldsymbol{v}, \boldsymbol{v} \rangle = \langle \lambda\boldsymbol{v}, \boldsymbol{v} \rangle = \langle O\boldsymbol{v}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, O^{\dagger}\boldsymbol{v} \rangle = \langle \boldsymbol{v}, O\boldsymbol{v} \rangle = \langle \boldsymbol{v}, \lambda\boldsymbol{v} \rangle = \lambda^{*}\langle \boldsymbol{v}, \boldsymbol{v} \rangle$$

By definition of eigenvector, $\boldsymbol{v}$ cannot be the null vector. Also, since the inner product between a vector and itself is not null, it is allowed to simplify

$$\lambda\cancel{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} = \lambda^{*}\cancel{\langle \boldsymbol{v}, \boldsymbol{v} \rangle} \Rightarrow \lambda = \lambda^{*}$$

Which means that λ is a real number, because it is equal to its complex conjugate. □

Theorem 1.3.2: If A and B are two unitary matrices, then $(AB)^{\dagger} AB = I$. In other words, the set of unitary matrices (of a fixed dimension) is closed under multiplication.

Proof: By definition, $A^{\dagger} A = I$ and $B^{\dagger} B = I$. Then:

$$(AB)^{\dagger} AB = B^{\dagger} A^{\dagger} AB = B^{\dagger} \left(A^{\dagger} A\right) B = B^{\dagger} I B = B^{\dagger} B = I$$

□

Theorem 1.3.3: If U is a unitary matrix, then the columns/rows of U form an orthonormal for of $\mathbb{C}^{n}$ with respect to the inner product, and vice versa.

Lemma 1.3.3: Unitary matrices are inner product-invariant. In other words, given a unitary matrix U and two vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^{n}$, $\langle U\boldsymbol{u}, U\boldsymbol{v} \rangle = \langle \boldsymbol{u}, \boldsymbol{v} \rangle$.

Recall that, for a fixed basis, each operator can be associated to a matrix that, when multiplied to a vector, performs the same action. By extension, an operator in matrix form is simply referred to as an "operator" as well.

Since trasposing an operator in matrix form exchanges domain and codomain, the conjugate transpose of an operator $O : V \to W$ is the operator $O^\dagger : W \to V$. The matrix representation of $O^\dagger$ is, as expected, the conjugate transpose of the matrix representation of O. The conjugate transpose of an operator is also called its **adjoint operator**. If an operator is equal to its adjoint (if its matrix representation is Hermitian), it is said to be **self-adjoint**.

Hermitian matrices are also normal matrices, therefore it is possible to apply the spectral theorem and obtain their eigendecomposition. The spectral decomposition of an Hermitian matrix has a unique property.

Theorem 1.3.4: Let O be an Hermitian matrix, having eigenvector basis $E = \{e_{\lambda_1}, ..., e_{\lambda_n}\}$. Let $S_{\lambda_i} = \text{span}\{e_{\lambda_i}\}$ be the subspace spanned by the i-th eigenvector. The set E is an orthogonal set and:

$$S_{\lambda_1} \oplus S_{\lambda_2} \oplus ... \oplus S_{\lambda_n} = V$$

Moreover, for every such decomposition, there exists a Hermitian operator whose eigendecomposition is this decomposition.

Proof: First, note how two eigenspaces of a matrix are necessarely disjoint. Given a unit vector $\boldsymbol{x}$ and two distinct eigenvalues λ_1 and λ_2, if $O\boldsymbol{x} = \lambda_1 \boldsymbol{x}$ and $O\boldsymbol{x} = \lambda_2 \boldsymbol{x}$ then $\lambda_1 \boldsymbol{x} = \lambda_2 \boldsymbol{x}$, but this is only possible if $\lambda_1 = \lambda_2$.

For any Hermitian operator, the eigenvectors for distinct eigenvalues must be orthogonal. Let $\boldsymbol{v}_1$ be an eigenvector for λ_1 and $\boldsymbol{v}_2$ an eigenvector for λ_2. Then:

$$\lambda_1 \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = \langle \boldsymbol{v}_1^* O^\dagger, \boldsymbol{v}_2 \rangle = \langle \boldsymbol{v}_1, O\boldsymbol{v}_2 \rangle = \lambda_2 \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle$$

Since λ_1 and λ_2 were assumed to be distinct, $\lambda_1 \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = \lambda_2 \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle$ is possible only if $\langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = 0$.

□

An important class of Hermitian operators are the **orthogonal projectors**. An orthogonal projector is any operator that is self-adjoint and maps elements of a space to elements of one of its subspaces. The matrix representation of an orthogonal projector, say P, has the remarkable property that $P^n = P = P^\dagger$.

Consider a vector space V having an orthonormal basis $\boldsymbol{v}_1, ..., \boldsymbol{v}_n$. Let W be a subspace of V of dimension $k \leq n$. It is possible to construct an orthogonal projector P in matrix form employing said basis as:

$$P = \sum_{i=1}^{k} \boldsymbol{v}_i^\dagger \boldsymbol{v}_i$$

Which is Hermitian by virtue of Lemma 1.3.1.

1.4. Bra-ket notation

A more comfortable formalism for denoting vectors is the **bra-ket** notation. The **ket** associated to a label Ψ, denoted as $|\Psi\rangle$, is just its column vector representation. The **bra** of a label Ψ, denoted as $\langle\Psi|$, is the transposed conjugate of the corresponding ket:

$$|\Psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix} \qquad \langle\Psi| = |\Psi\rangle^\dagger = \begin{pmatrix} \psi_1^* & \psi_2^* & ... & \psi_n^* \end{pmatrix}$$

The meaning of the label Ψ is arbitrary: there has to be a way (implicit or explicitly stated) to assign labels to vectors. In general, if a vector is referred to as v, its representation as ket is $|v\rangle$ and the representation of its complex conjugate as bra is $\langle v|$. In the context of quantum mechanics, the label of a ket or of a bra generally refers to a certain quantum state.

Exercise 1.4.1: Represent the canonical basis of $\mathbb{R}^2$ using the Dirac notation. Recall that canonical basis vectors are often denoted as $e_1, e_2, ..., e_n$.

Solution:

$$|e_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad |e_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad |e_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \qquad |e_4\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

□

Any linear combination of vectors with respect to a basis $\{\beta_1, ..., \beta_n\}$ can be written seamlessly as a linear combination of kets with respect to the basis $\{|\beta_1\rangle, ..., |\beta_n\rangle\}$:

$$\sum_{i=1}^{n} \lambda_i \begin{pmatrix} \beta_{i,1} \\ \vdots \\ \beta_{i,n} \end{pmatrix} = \lambda_1 \begin{pmatrix} \beta_{1,1} \\ \vdots \\ \beta_{1,n} \end{pmatrix} + ... + \lambda_n \begin{pmatrix} \beta_{n,1} \\ \vdots \\ \beta_{n,n} \end{pmatrix} \Longleftrightarrow \sum_{i=1}^{n} \lambda_i \; |\beta_i\rangle = \lambda_1 \; |\beta_1\rangle + ... + \lambda_n \; |\beta_n\rangle$$

The inner product of two kets $|\Psi_1\rangle$ and $|\Psi_2\rangle$ is simply denoted as $\langle\Psi_1|\Psi_2\rangle$, and is called **braket**. The outer product of two kets $|\Psi_1\rangle$ and $|\Psi_2\rangle$ is simply denoted as $|\Psi_1\rangle\langle\Psi_2|$, and is called **ketbra**.

Exercise 1.4.2: Consider the following two vectors, written in bra-ket notation following the convention of Exercise 1.4.1. What are their braket and ketbra?

$$|A\rangle = \frac{3}{4} \; |e_1\rangle \qquad\qquad |B\rangle = \frac{\sqrt{5}}{2} \; |e_3\rangle$$

Solution:

$$\langle A|B\rangle = \frac{3}{4}\frac{\sqrt{5}}{2}\langle e_1|e_2\rangle = \frac{3\sqrt{5}}{8} \qquad |A\rangle\langle B| = \frac{3}{4} \; |e_1\rangle \otimes \frac{\sqrt{5}}{2}\langle e_2| = \frac{3\sqrt{5}}{8} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

□

If a state Ψ belongs to an Hilber space H, its conjugate Ψ^* belongs to the **dual space** H^* of the Hilbert space H. Another way of expressing it is that if a ket $|\Psi\rangle$ belongs to H, then the corresponding bra $\langle\Psi|$ belongs to H^*.

The dual space is constitued by linear functionals X over the kets in H: if $|\Psi\rangle \in \mathrm{H}$ then $\mathrm{X}(\Psi) \in \mathbb{C}$:

$$\mathrm{X}(\alpha_1 \; |\Psi_1\rangle + \alpha_2 \; |\Psi_2\rangle) = \mathrm{X}(\alpha_1 \; |\Psi_1\rangle) + \mathrm{X}(\alpha_2 \; |\Psi_2\rangle)$$

Even though matrix representation of vectors is useful from time to time, the bra-ket notation is much comfortable to work with in quantum mechanics. Not only because it is more compact but also because it represents the conjugate transpose of a vector, which is an operation performed very often, as "flipping" its symbol left-to-right. In particular, the inner product "flips" the ket on the left side of the operation, whereas the outer product "flips" the ket on the right side of the operation.

The bra-ket notation can also be used to represent operators in matrix form. This is particularly useful to denote the application of operators to states written as kets and bras. To do this, first recall that the ketbra $|v\rangle\langle w|$ represents the outer product between $|v\rangle$ and $|w\rangle$ or, equivalently, the tensor product between $|v\rangle$ and $\langle w|$.

In particular, consider using canonical basis vectors $|e_1\rangle, ..., |e_n\rangle$ as done in Exercise 1.4.1. The ketbra of any pair of kets $|e_i\rangle, |e_j\rangle$ is:

$$|e_i\rangle\langle e_j| = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & \dots & 1 & \dots & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & 1 & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

Which means that a generic $m \times n$ matrix A can be written as:

$$A = \sum_{i=1}^{m} \sum_{i=1}^{n} a_{i,j} \; |e_i\rangle\langle e_j|$$

Where $a_{i,j}$ is the entry of A in the i-th row and j-th column, $|e_i\rangle$ is the i-th canonical basis vector and $\langle e_j|$ is the conjugate transpose of the j-th canonical basis vector. This representation is particularly powerful in the case of *sparse matrices*, matrices having 0 in many of its entries, because all those entries do not need to be denoted explicitly.

Exercise 1.4.3: How would this matrix be written in bra-ket notation?

$$A = \begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{\sqrt{2}}{2} & 0 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & 0 & 0 \end{pmatrix}$$

Solution:

$$\begin{aligned} A = {} & \sum_{i=1}^{4} \sum_{i=1}^{4} a_{i,j} \; |i\rangle\langle j| = \\ & = \sum_{i=1}^{4} a_{i,0} \; |e_i\rangle\langle e_1| + a_{i,1} \; |e_i\rangle\langle e_2| + a_{i,2} \; |e_i\rangle\langle e_3| + a_{i,3} \; |e_i\rangle\langle e_4| = \\ & = \frac{\sqrt{2}}{2} \; |e_1\rangle\langle e_1| + 0 \; |e_2\rangle\langle e_1| + 0 \; |e_3\rangle\langle e_1| + 0 \; |e_4\rangle\langle e_1| + 0 \; |e_1\rangle\langle e_2| + 0 \; |e_2\rangle\langle e_2| + 0 \; |e_3\rangle\langle e_2| + 0 \; |e_4\rangle\langle e_2| + \\ & \quad \frac{\sqrt{2}}{2} \; |e_1\rangle\langle e_3| + 0 \; |e_2\rangle\langle e_3| + 0 \; |e_3\rangle\langle e_3| + 0 \; |e_4\rangle\langle e_3| + 0 \; |e_1\rangle\langle e_4| + \frac{\sqrt{2}}{2} \; |e_2\rangle\langle e_4| + 0 \; |e_3\rangle\langle e_4| + 0 \; |e_4\rangle\langle e_4| = \\ & = \frac{\sqrt{2}}{2} \; |e_1\rangle\langle e_1| + \frac{\sqrt{2}}{2} \; |e_1\rangle\langle e_3| + \frac{\sqrt{2}}{2} \; |e_2\rangle\langle e_4| \end{aligned}$$

□

This notion generalizes to the case of any basis $\{\varphi_1, ..., \varphi_n\}$. A matrix O defined in bra-ket notation with respect to this basis is written as:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} \; |\varphi_i\rangle\langle\varphi_j|$$

2. Qubits

2.1. Single qubit systems

Consider any physical system that can be observed in only two possible states. Systems such as these can be constructed into many different ways, such as inspecting the spin of the electron (only spin up and spin down exist, no other spins can be found) or inspecting the energy levels of the electrons of very simple atoms (only a ground state and an excited state exist, no other states can be found). Systems such as these are called **two-state quantum systems**, or just two-state systems.

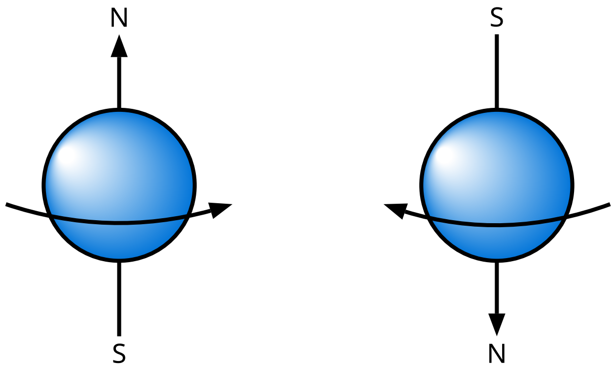


Figure 1: The spin is an intrinsic property of fundamental particles. Electrons have a value of spin that is either equal to $1/2$ or $-1/2$, also referred to as "up" and "down" respectively.

The term "two-state system" is somewhat misleading. Indeed, until measurement happens, the number of states in which a physical system can find itself is infinite; it is only *after* the measurement is performed that the system will be found in one of the two states. Therefore, "two-state" refers to the state of the system after the measurement has taken place.

More precisely, following the principles of quantum mechanics, these two "special" states, called **base states**, form a basis for the Hilbert space that contains the possible states in which the system can be *before* measurement happens. Any of these states can be constructed as a linear combination of the aforementioned basis, normalized according to Born's rule. In this respect, the term "two-state" refers to the number of dimensions of the Hilbert space

Let $|\varphi_1\rangle$ and $|\varphi_2\rangle$ be two base states. Any linear combination of the two is also a legitimate state $|\Psi\rangle$, as long as a normalization condition is respected:

$$|\Psi\rangle = \alpha\ |\varphi_1\rangle + \beta\ |\varphi_2\rangle, \text{with } \alpha, \beta \in \mathbb{C} \ \text{ such that } |\alpha|^2 + |\beta|^2 = 1$$

A two-state quantum system is also referred to as **qubit**. The name "qubit" comes from its classical counterpart, the bit, but while a bit is either 0 or 1, a qubit is in an indeterminate state until the measurement is performed[1].

It is therefore valid to refer to a state such as the $|\Psi\rangle$ described above as a qubit. In particular, being the result of a linear combination of basis, any state/qubit $|\Psi\rangle$ can be entirely represented (with respect to that basis) as the coefficients of the linear combination itself:

$$|\Psi\rangle = \alpha\ |\varphi_1\rangle + \beta\ |\varphi_2\rangle \Longleftrightarrow \begin{pmatrix} \alpha \\ \beta \end{pmatrix}_{\{|\varphi_1\rangle,\ |\varphi_2\rangle\}}$$

Any pair of states that form a basis for a two-dimensional Hilbert space and are also orthogonal to each other (in other words, form an orthonormal basis) can be used as base states. The simplest choice is the pair of vectors $\binom{1}{0}$ and $\binom{0}{1}$, commonly denoted as $|0\rangle$ and $|1\rangle$ respectively[2].

[1] A n-state quantum system is called a **qudit**, and it has the same computational power of a qubit.

[2] The name emphasises the analogy with the classical bit, but the choice of assigning these vectors to their respective symbols is completely arbitrary.

Theorem 2.1.1: The set $\{|0\rangle, |1\rangle\}$ forms an orthonormal basis for any two-dimensional Hilbert space.

Proof: The null vector of any two-dimensional Hilbert space is $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Constructing the null combination gives:

$$\mathbf{0} = k_1 \; |0\rangle + k_2 \; |1\rangle \Rightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix} = k_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Rightarrow \begin{cases} 0 = k_1 \cdot 1 + k_2 \cdot 0 \\ 0 = k_1 \cdot 0 + k_2 \cdot 1 \end{cases} \Rightarrow \begin{cases} 0 = k_1 \\ 0 = k_2 \end{cases}$$

Being linearly independent, they do form a basis. They are also orthonormal:

$$\langle 0|0\rangle = \langle 1|1\rangle = 1 \cdot 1 + 0 \cdot 0 = 1 \qquad \langle 0|1\rangle = \langle 1|0\rangle = 1 \cdot 0 + 0 \cdot 1 = 0$$

□

This basis is used obiquitously, and is therefore reffered to as the **standard basis**. When the basis at play is not specified, it is assumed that the basis is the standard basis. Denoting these two vectors as $|0\rangle$ and $|1\rangle$ is helpful because it allows one to intuitively associate these vectors to the classical bits 0 and 1.

Since orthonormality is a necessary condition for being a physically meaningful basis, when talking about a basis it will be implicitly assumed (unless stated otherwise) that the basis is orthonormal.

Another useful basis is the one denoted as $\{| +\rangle, | -\rangle\}$:

$$| +\rangle = \frac{\sqrt{2}}{2}(|0\rangle + |1\rangle) \qquad | -\rangle = \frac{\sqrt{2}}{2}(|0\rangle - |1\rangle)$$

Theorem 2.1.2: The set $\{| +\rangle, | -\rangle\}$ forms an orthonormal basis for any two-dimensional Hilbert space.

Proof: The basis $\{|0\rangle, |1\rangle\}$ can be written as a linear combination of $\{| +\rangle, | -\rangle\}$:

$$|0\rangle = \frac{\sqrt{2}}{2}(| +\rangle + | -\rangle) \qquad |1\rangle = \frac{\sqrt{2}}{2}(| +\rangle - | -\rangle)$$

Therefore, $\{| +\rangle, | -\rangle\}$ is a basis as well. It's also orthonormal:

$$\langle +| +\rangle = \langle -| -\rangle = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2}}{2} = 1 \qquad \langle +| -\rangle = \langle -| +\rangle = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2}}{2} = 0$$

□

Another relevant basis is $\{| \circlearrowleft\rangle, | \circlearrowright\rangle\}$, defined as:

$$| \circlearrowleft\rangle = \frac{\sqrt{2}}{2}(|0\rangle + i \; |1\rangle) \qquad | \circlearrowright\rangle = \frac{\sqrt{2}}{2}(|0\rangle - i \; |1\rangle)$$

Theorem 2.1.3: The set $\{| \circlearrowleft\rangle, | \circlearrowright\rangle\}$ forms an orthonormal basis for any two-dimensional Hilbert space.

Proof: This set is indeed a basis since the basis $\{|0\rangle, |1\rangle\}$ can be written as a linear combination of $\{| \circlearrowleft\rangle, | \circlearrowright\rangle\}$:

$$|0\rangle = \frac{\sqrt{2}}{2}(| \circlearrowleft\rangle + i \; | \circlearrowright\rangle) \qquad |1\rangle = \frac{\sqrt{2}}{2}(| \circlearrowleft\rangle - i \; | \circlearrowright\rangle)$$

Therefore, $\{| \circlearrowleft\rangle, | \circlearrowright\rangle\}$ is a basis as well. It's also orthonormal:

$$\langle \circlearrowleft \mid \circlearrowleft \rangle = \langle \circlearrowright \mid \circlearrowright \rangle = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2} \cdot i\frac{\sqrt{2}}{2} = 1 \quad \langle \circlearrowleft \mid \circlearrowright \rangle = \langle \circlearrowright \mid \circlearrowleft \rangle = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2} \cdot i\frac{\sqrt{2}}{2} = 0$$

□

According to Born's rule, the probability of finding $|\Psi\rangle$ in the state $|\varphi_1\rangle$ when measured is given by $|\alpha|^2$, wherease the probability of finding it in the state $|\varphi_2\rangle$ when measured is given by $|\beta|^2$.

As stated, this is an axiom of quantum mechanics, derived from experimental data and assumed to be true. Also, a measurement induces a wave function collapse, and the state of the qubit becomes one of the possible basis states. This means that the measurement process does not exist "in a vacuum", but is dependent on a chosen basis. This also means that any measurement performed afterwards will always give the same result with absolute certainty.

Note how:

$$|\langle \varphi_1 | \Psi \rangle|^2 = |\alpha \langle \varphi_1 | \varphi_1 \rangle + \beta \langle \varphi_1 | \varphi_2 \rangle|^2 = |\alpha|^2 \qquad |\langle \varphi_2 | \Psi \rangle|^2 = |\alpha \langle \varphi_2 | \varphi_1 \rangle + \beta \langle \varphi_2 | \varphi_2 \rangle|^2 = |\beta|^2$$

Later chapters will expand on this formalism, but the statement just given is, for the moment, sufficient.

Exercise 2.1.1: Write the state $|\Psi\rangle = \frac{\sqrt{2}}{2}(|0\rangle - |1\rangle)$ as a linear combination of the basis $\{| \circlearrowleft \rangle, | \circlearrowright \rangle\}$. What are the probabilities of obtaining the respective measurements?

Solution:

$$\begin{aligned} |\Psi\rangle &= \frac{\sqrt{2}}{2}(|0\rangle - |1\rangle) = \frac{\sqrt{2}}{2}\left(\frac{\sqrt{2}}{2}(| \circlearrowleft \rangle + i \mid \circlearrowright \rangle) - \frac{\sqrt{2}}{2}(| \circlearrowleft \rangle - i \mid \circlearrowright \rangle)\right) = \\ &= \frac{1}{2}(| \circlearrowleft \rangle + i \mid \circlearrowright \rangle) - \frac{1}{2}(| \circlearrowleft \rangle - i \mid \circlearrowright \rangle) = \cancel{\frac{1}{2} \mid \circlearrowleft \rangle} + \frac{i}{2} \mid \circlearrowright \rangle - \cancel{\frac{1}{2} \mid \circlearrowleft \rangle} + \frac{i}{2} \mid \circlearrowright \rangle = \\ &= i \mid \circlearrowright \rangle \end{aligned}$$

Which means that the probability of getting $| \circlearrowright \rangle$ is $|0 + 1i|^2 = 1$ and the probability of getting $| \circlearrowleft \rangle$ is 0.□

Any vector that results from a non-trivial linear combination of a basis, that is, when both coefficients of the linear combination are not zero, is said to be in a **superposition** of the states that comprise the basis. A basis is always necessary to be specified when talking about superposition: a state can be the result of a superposition with respect to a certain basis but not with respect to another basis.

Exercise 2.1.2: Consider the states $|\Psi_1\rangle$ and $|\Psi_2\rangle$. Are they in a superposition with respect to the basis $\{|0\rangle, |1\rangle\}$?

$$|\Psi_1\rangle = \frac{\sqrt{2}}{2}(| +\rangle + | -\rangle) \qquad |\Psi_2\rangle = \frac{\sqrt{3}}{2} \mid +\rangle - \frac{1}{2} \mid -\rangle$$

Solution: Note how:

$$\begin{aligned} |\Psi_1\rangle &= \frac{\sqrt{2}}{2}(| +\rangle + | -\rangle) = \frac{\sqrt{2}}{2}\left(\frac{\sqrt{2}}{2}(|0\rangle + |1\rangle) + \frac{\sqrt{2}}{2}(|0\rangle - |1\rangle)\right) = \frac{1}{2}(|0\rangle + |1\rangle) + \frac{1}{2}(|0\rangle - |1\rangle) = \\ &= \frac{1}{2} \mid 0\rangle + \cancel{\frac{1}{2} \mid 1\rangle} + \frac{1}{2} \mid 0\rangle - \cancel{\frac{1}{2} \mid 1\rangle} = |0\rangle \end{aligned}$$

This means that $|\Psi_1\rangle$ is just one of the two elements of the standard basis, and therefore there is no superposition. Indeed, $|\Psi_1\rangle$ written as a linear combination of the standard basis would be $|\Psi_1\rangle = 1 \; |0\rangle + 0 \; |1\rangle$, which is a trivial combination. On the other hand:

$$
\begin{aligned}
|\Psi_2\rangle &= \frac{\sqrt{3}}{2}\ |+\rangle - \frac{1}{2}\ |-\rangle = \frac{\sqrt{3}}{2}\left(\frac{\sqrt{2}}{2}(|0\rangle + |1\rangle)\right) - \frac{1}{2}\left(\frac{\sqrt{2}}{2}(|0\rangle - |1\rangle)\right) = \\
&= \frac{\sqrt{6}}{4}(|0\rangle + |1\rangle) - \frac{\sqrt{2}}{4}(|0\rangle - |1\rangle) = \frac{\sqrt{6}}{4}\ |0\rangle + \frac{\sqrt{6}}{4}\ |1\rangle - \frac{\sqrt{2}}{4}\ |0\rangle + \frac{\sqrt{2}}{4}\ |1\rangle = \\
&= \frac{\sqrt{6}-\sqrt{2}}{4}\ |0\rangle + \frac{\sqrt{6}+\sqrt{2}}{4}\ |1\rangle
\end{aligned}
$$

Which is a non-trivial combination. □

When a measurement is not performed, the system is in a superposition of base states, and the state in which the system is found when measured can be predicted only within a certain probability. Nevertheless, it is possible to extract at most a single bit of information from a qubit. Indeed, the state in which the qubit is prior to measurement is unknown and unknowable, and when measurement happens the value of the qubit is always one out of two allowed values. It would therefore be incorrect to state that a qubit holds an infinite amount of information.

That the same quantum state is represented by more than one vector means that there is a critical distinction between the complex vector space in which qubit values are written and the quantum state space itself. In particular, any unit vector multiplied by a phase factor is equivalent to the original vector, and therefore represents the same state.

The multiple by which two vectors representing the same quantum state differ is called the **global phase** and has no physical meaning. The notation $|v\rangle \sim |v'\rangle$ denotes the fact that the two vectors are equivalent up to a global phase $e^{i\varphi}$, that is $|v\rangle = e^{i\varphi}\ |v'\rangle$. The space in which two two-dimensional complex vectors are considered equivalent if they are multiples of each other is called **complex projective space** of dimension one.

Two complex vectors that differ from a phase factor belong to the same equivalence class with respect to the aforementioned relation. Each of these equivalence classes are the members of a quotient space, denoted as $\boldsymbol{CP}^1$:

$$\boldsymbol{CP}^1 = \{\alpha\ |\varphi_1\rangle + \beta\ |\varphi_2\rangle\}/\sim$$

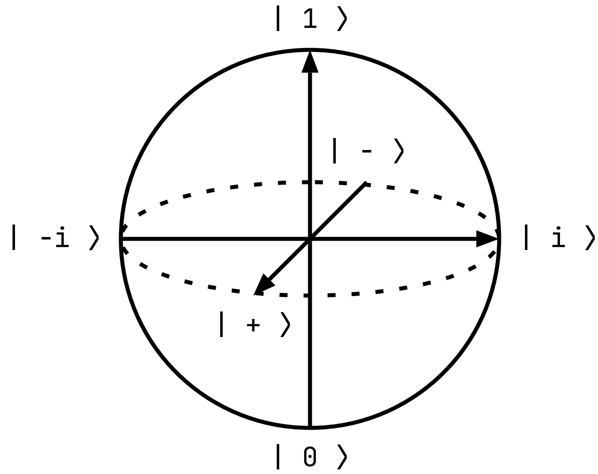
Therefore, the quantum state space for a single-qubit system is in one-to-one correspondence with the points of the complex projective space $\boldsymbol{CP}^1$.

A physical quantity that, unlike the global phase, is *not* irrelevant, is the **relative phase** of a single-qubit state. The relative phase of a superposition $\alpha\ |v_1\rangle + \beta\ |v_2\rangle$ is a measure of the angle in the complex plane between the two complex numbers α and β. More precisely, the relative phase is the complex number $e^{i\varphi}$ (that is, having modulus equal to one) such that:

$$\frac{\alpha}{\beta} = e^{i\varphi}\frac{|\alpha|}{|\beta|} \Rightarrow e^{i\varphi} = \frac{\alpha|\beta|}{|\alpha|\beta}$$

Two superpositions $\alpha\ |v_1\rangle + \beta\ |v_2\rangle$ and $\alpha'\ |v_1\rangle + \beta'\ |v_2\rangle$ whose amplitudes have the same magnitudes but that differ in a relative phase represent different states. On the other hand, if two superpositions (with respect to the same basis) have the same relative phase, they represent the same state.

Bases can be represented graphically as coordinates on a sphere, called **Bloch sphere**:



2.2. Multiple qubits systems

Consider two Hilbert spaces H_1 and H_2, each describing the state space of a single qubit. Clearly, there is interest in describing the Hilbert space of the entire system, encompassing both qubits at once.

It might be tempting to describe this merging of spaces simply by computing the direct sum of the two single qubit Hilbert spaces, like it is done in classical physics, but it would be incorrect. This is because the resulting space might not be a valid description of quantum states.

The Hilbert space describing the state of two qubits is given instead by the tensor product of the Hilbert spaces of the respecting qubits. Recall that, due to Lemma 1.2.1 and Lemma 1.2.2, the tensor product preserves both the norm of vectors and the orthonormality of bases. This means that the tensor product of two qubit state spaces is also guaranteed to be a valid state space.

More explicitly, let H_1 and H_2 be the state spaces of two qubits, having $\{|\varphi_1\rangle_1, |\varphi_2\rangle_1\}$ and $\{|\varphi_1\rangle_2, |\varphi_2\rangle_2\}$ respectively as bases. The basis of $H_1 \otimes H_2$, the state space of the system encompassing both qubits, is:

$$\{|\varphi_1\rangle_1 \otimes |\varphi_1\rangle_2, |\varphi_1\rangle_1 \otimes |\varphi_2\rangle_2, |\varphi_2\rangle_1 \otimes |\varphi_1\rangle_2, |\varphi_2\rangle_1 \otimes |\varphi_2\rangle_2\}$$

Which, in turn, means that any description of the state of two qubits at once is in the form:

$$\lambda_{1,1}\left(|\varphi_1\rangle_1 \otimes |\varphi_1\rangle_2\right) + \lambda_{1,2}\left(|\varphi_1\rangle_1 \otimes |\varphi_2\rangle_2\right) + \lambda_{2,1}\left(|\varphi_2\rangle_1 \otimes |\varphi_1\rangle_2\right) + \lambda_{2,2}\left(|\varphi_2\rangle_1 \otimes |\varphi_2\rangle_2\right)$$

Recall that the tensor product is not commutative, therefore the Hilbert spaces $H_1 \otimes H_2$ and $H_2 \otimes H_1$ are not the same. This means that it is strictly necessary to define an order on the qubits, otherwise the results would be inconsistent.

In the previous expression, the order of the qubits is given by the subscripts placed on the kets. This is fine, but cumbersome. A simpler approach is to drop the subscripts entirely and assume that the order of the kets mirrors the order of the qubits:

$$\lambda_{1,1}(|\varphi_1\rangle \otimes |\varphi_1\rangle) + \lambda_{1,2}(|\varphi_1\rangle \otimes |\varphi_2\rangle) + \lambda_{2,1}(|\varphi_2\rangle \otimes |\varphi_1\rangle) + \lambda_{2,2}(|\varphi_2\rangle \otimes |\varphi_2\rangle)$$

Therefore, any tensor product $|\varphi_i\rangle \otimes |\varphi_j\rangle$ is the tensor product between the base state $|\varphi_i\rangle$ of the first qubit and the base state $|\varphi_j\rangle$ of the second qubit.

It is also convenient to use the shortand $|\varphi_i\varphi_j\rangle$ for $|\varphi_i\rangle \otimes |\varphi_j\rangle$, allowing the basis of the two-qubit system to be written more compactly:

$$\{|\varphi_1\varphi_1\rangle, |\varphi_1\varphi_2\rangle, |\varphi_2\varphi_1\rangle, |\varphi_2\varphi_2\rangle\}$$

Again, in the notation $|\varphi_i\varphi_j\rangle$ it is implicitly assumed that $|\varphi_i\rangle$ is the state of the first qubit and $|\varphi_j\rangle$ is the state of the second qubit.

Exercise 2.2.1: Consider the two Hilbert spaces:

$$H_1 = \text{span}\{| \circlearrowleft\rangle, | \circlearrowright\rangle\} \qquad\qquad H_2 = \text{span}\{| +\rangle, | -\rangle\}$$

Each associated to a qubit. What is the Hilbert space $H_1 \otimes H_2$ that describes both qubits at once?

Solution: $H_1 \otimes H_2$ is spanned by the following four states:

$$|\varphi_1\rangle = | \circlearrowleft\rangle \otimes | +\rangle = \left(\frac{\sqrt{2}}{2}(|0\rangle + i\ |1\rangle)\right) \otimes \left(\frac{\sqrt{2}}{2}(|0\rangle + |1\rangle)\right) = \frac{1}{2}(|00\rangle + |01\rangle + i\ |10\rangle + i\ |11\rangle)$$

$$|\varphi_2\rangle = | \circlearrowleft\rangle \otimes | -\rangle = \left(\frac{\sqrt{2}}{2}(|0\rangle + i\ |1\rangle)\right) \otimes \left(\frac{\sqrt{2}}{2}(|0\rangle - |1\rangle)\right) = \frac{1}{2}(|00\rangle - |01\rangle + i\ |10\rangle - i\ |11\rangle)$$

$$|\varphi_3\rangle = | \circlearrowright\rangle \otimes | +\rangle = \left(\frac{\sqrt{2}}{2}(|0\rangle - i\ |1\rangle)\right) \otimes \left(\frac{\sqrt{2}}{2}(|0\rangle + |1\rangle)\right) = \frac{1}{2}(|00\rangle + |01\rangle - i\ |10\rangle - i\ |11\rangle)$$

$$|\varphi_4\rangle = | \circlearrowright\rangle \otimes | -\rangle = \left(\frac{\sqrt{2}}{2}(|0\rangle - i\ |1\rangle)\right) \otimes \left(\frac{\sqrt{2}}{2}(|0\rangle - |1\rangle)\right) = \frac{1}{2}(|00\rangle - |01\rangle - i\ |10\rangle + i\ |11\rangle)$$

□

This can be extended naturally to the case of a system having $n > 2$ qubits. If each of those qubits is represented by a Hilbert space H_i, with $i \in \{1, ..., n\}$, the state space of the entire system is $H_n \otimes H_{n-1} \otimes ... \otimes H_1$. Its basis can be written as:

$$\{|\varphi_1\varphi_1...\varphi_n\rangle, |\varphi_1\varphi_2...\varphi_n\rangle, ..., |\varphi_2\varphi_1...\varphi_n\rangle, |\varphi_2\varphi_2...\varphi_n\rangle, ..., |\varphi_n\varphi_1...\varphi_n\rangle, |\varphi_n\varphi_2...\varphi_n\rangle\}$$

As it is the case for a single qubit, a state of an n qubit system is in a superposition with respect to a certain basis if the linear combination:

$$\lambda_{1,1,...,n}\ |\varphi_1\varphi_1...\varphi_n\rangle + \lambda_{1,2,...,n}\ |\varphi_1\varphi_2...\varphi_n\rangle + ... + \lambda_{n,1,...,n}\ |\varphi_n\varphi_1...\varphi_n\rangle + \lambda_{n,2,...,n}\ |\varphi_n\varphi_2...\varphi_n\rangle$$

Is not trivial, meaning that at least two coefficients are not zero. As expected, superposition is basis-dependent.

When the basis under consideration is the standard basis, it is possible to write the basis of an n-qubit state space even more compactly. Recall that the two standard base states for a qubit are $|0\rangle = \binom{1}{0}$ and $|1\rangle = \binom{0}{1}$. For a two qubit system, it is possible to use integers $0, 1, 2, 3$ in the same way:

$$|00\rangle = |0\rangle = \begin{pmatrix}1\\0\\0\\0\end{pmatrix} \qquad |01\rangle = |1\rangle = \begin{pmatrix}0\\1\\0\\0\end{pmatrix} \qquad |10\rangle = |2\rangle = \begin{pmatrix}0\\0\\1\\0\end{pmatrix} \qquad |11\rangle = |3\rangle = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

In general, for a n qubit state it is possible to write its standard basis as $\{|0\rangle, |1\rangle, ..., |2^n - 2\rangle, |2^n - 1\rangle\}$, where each ket contains the integer representation of the binary number constructed by justapposing the binary digits of the state of each qubit:

$$|00...0\rangle = |0\rangle = \begin{pmatrix}1\\0\\0\\\vdots\\0\\0\\0\end{pmatrix} \quad |00...1\rangle = |1\rangle = \begin{pmatrix}0\\1\\0\\\vdots\\0\\0\\0\end{pmatrix} \quad \vdots \quad |11...0\rangle = |2^n - 2\rangle = \begin{pmatrix}0\\0\\0\\\vdots\\0\\1\\0\end{pmatrix} \quad |11...1\rangle = |2^n - 1\rangle = \begin{pmatrix}0\\0\\0\\\vdots\\0\\0\\1\end{pmatrix}$$

Note that, to use this shorthand notation, it is necessary to also specify the number of qubits of the state. Therefore, it should be employed when the number of qubits is either denoted explicitly or known from context.

Exercise 2.2.2: How would the matrix representation of the following state be?

$$\frac{1}{2}\ |00\rangle + \frac{1}{2}i\ |01\rangle + \frac{\sqrt{2}}{2}\ |11\rangle$$

Solution: Each ket has two digits, therefore this is a two qubit state:

$$\frac{1}{2}\ |00\rangle + \frac{1}{2}i\ |01\rangle + \frac{\sqrt{2}}{2}\ |11\rangle = \frac{1}{2}\ |0\rangle + \frac{1}{2}i\ |1\rangle + \frac{\sqrt{2}}{2}\ |3\rangle = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2}i \\ 0 \\ \frac{\sqrt{2}}{2} \end{pmatrix}$$

□

It is possible to use the n-dimensional standard basis to describe the state of n-qubits, but just like the case of a single qubit there are other bases that are useful to employ. One such basis for the case of a 2 qubit system is the following:

$$|\Phi^+\rangle = \frac{\sqrt{2}}{2}(|00\rangle + |11\rangle) \quad |\Phi^-\rangle = \frac{\sqrt{2}}{2}(|00\rangle - |11\rangle) \quad |\Psi^+\rangle = \frac{\sqrt{2}}{2}(|01\rangle + |10\rangle) \quad |\Psi^-\rangle = \frac{\sqrt{2}}{2}(|01\rangle - |10\rangle)$$

These states are called **Bell states**, and the basis that they form is called **Bell basis**.

The following 3-qubit states often appear in quantum experiments, respectively referred to as **Greenberger–Horne–Zeilinger state** and **W state**:

$$|GHZ\rangle = \frac{\sqrt{2}}{2}(|000\rangle + |111\rangle) \qquad\qquad |W\rangle = \frac{\sqrt{3}}{3}(|100\rangle + |010\rangle + |001\rangle)$$

Any unit vector of the 2^n-dimensional state space represents a possible state of an n-qubit system, but just as in the single-qubit case there is redundancy. In the multiple-qubit case, not only do vectors that are multiples of each other refer to the same quantum state, but properties of the tensor product also mean that phase factors distribute over tensor products; the same phase factor in different qubits of a tensor product represent the same state:

$$|v\rangle \otimes \left(e^{i\theta}\ |w\rangle\right) = e^{i\theta}(|v\rangle \otimes |w\rangle) = \left(e^{i\theta}\ |v\rangle\right) \otimes |w\rangle$$

Phase factors in individual qubits of a single term of a superposition can always be factored out into a single coefficient for that term.

Just as in the single-qubit case, vectors that differ only in a global phase represent the same quantum state; the space of distinct quantum states of an n-qubit system is a complex projective space of dimension $2^n - 1$, but in general it is easier to consider the Hilbert space directly but taking into account possible duplicate vectors.

The fact that state spaces describing more than one qubit are constructed from the tensor product has a crucial consequence. Even though the tensor product of n single qubit states represents a state of a n-qubit system, a state of a n-qubit system might not be decomposable into a tensot product of n single qubit states. This is because the dimension of a composite Hilbert spaces grows exponentially.

This means that there are states of n-qubit systems that cannot be conceived as simply the result of the combined contribution of each qubit, but are instead entities on their own. States like these are called **entangled states**. For example, the aforementioned Bell states are all entangled states. Indeed, the majority of states of a n-qubit system are entangled states.

Exercise 2.2.3: Consider this two 2-qubit states. Are they entangled states?

$$|\varphi_1\rangle = \frac{1}{8}\,|00\rangle + \frac{\sqrt{7}}{8}\,|01\rangle + \frac{\sqrt{7}}{8}\,|10\rangle + \frac{7}{8}\,|11\rangle \qquad |\varphi_2\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Solution: $|\varphi_1\rangle$ is not an entangled state, because it can be decomposed into two single (identical) qubit states as follows:

$$|\varphi_1\rangle = \frac{1}{8}\,|00\rangle + \frac{\sqrt{7}}{8}\,|01\rangle + \frac{\sqrt{7}}{8}\,|10\rangle + \frac{7}{8}\,|11\rangle = \left(\frac{1}{\sqrt{8}}\,|0\rangle + \sqrt{\frac{7}{8}}\,|1\rangle\right) \otimes \left(\frac{1}{\sqrt{8}}\,|0\rangle + \sqrt{\frac{7}{8}}\,|1\rangle\right)$$

On the other hand, the state $|\varphi_2\rangle$ is entangled. Attempting a decomposition gives:

$$(a\,|0\rangle + b\,|1\rangle) \otimes (c\,|0\rangle + d\,|1\rangle) = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \Rightarrow$$

$$ac\,|00\rangle + ad\,|01\rangle + bc\,|10\rangle + bd\,|11\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \Rightarrow$$

$$ac\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + ad\begin{pmatrix}0\\1\\0\\0\end{pmatrix} + bc\begin{pmatrix}0\\0\\1\\0\end{pmatrix} + bd\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \frac{1}{\sqrt{2}}\left(\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\right) \Rightarrow$$

$$\begin{pmatrix}ac\\ad\\bc\\bd\end{pmatrix} = \begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\0\\\frac{1}{\sqrt{2}}\end{pmatrix} \Rightarrow \begin{cases}ac = \frac{1}{\sqrt{2}}\\ ad = 0\\ bc = 0\\ bd = \frac{1}{\sqrt{2}}\end{cases}$$

Which is an impossible system of equations to solve. □

In the case of a two-qubit system, there is one and only decomposition (two one-qubit systems), therefore there is no ambiguity. In general, a state can be said to be entangled only with respect to a specific decomposition. More formally, a state $|\Psi\rangle$ member of a state space H decomposed as $H_1 \otimes H_2 \otimes \ldots \otimes H_n$ is said to be *separable* or *unentangled* if:

$$|\Psi\rangle = (|v_1\rangle \in H_1) \otimes (|v_2\rangle \in H_2) \otimes \ldots \otimes (|v_n\rangle \otimes H_n)$$

Otherwise, it is said to be entangled. When non specified, saying that a n-qubit state is entangled means "entangled with respect to the decomposition into n individual qubit states".

Despite being decomposition-dependent, entanglement is basis-independent, since the chosen basis plays no role in the definition of entanglement (even though some bases might be more comfortable than others when working with entangled states).

Measuring a n-qubit state with respect to its base states is no different than in the single qubit case. What could be more interesting is measuring only a part of the system; for example, in a 2-qubit system, measuring only the first qubit while leaving the second unmeasured, or vice versa. This is elaborated in the next chapter.

Exercise 2.2.4: Consider the 2-qubit state $|\Psi\rangle = \frac{1}{8}\,|00\rangle + \frac{\sqrt{7}}{8}\,|01\rangle + \frac{\sqrt{7}}{8}\,|10\rangle + \frac{7}{8}\,|11\rangle$. What is the probability of measuring each base state?

Solution: The probability of measuring $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ are, respectively: $\frac{1}{64}, \frac{7}{64}, \frac{7}{64}, \frac{49}{64}$. □

2.3. Qubit measurement

The kind of measurements presented until now consisted of measuring the state with respect to its basis. There is, however, a much richer way to measure quantum states, that goes beyond measuring bases.

First, recall that the outer product between two kets is a matrix (an operator). Since they are vectors like any other, this is valid for basis vectors as well: the outer product between two basis states is a matrix (an operator). However, basis vectors are orthonormal, making the representation of matrices in bra-ket notation particularly enticing.

Consider an n-qubit system spanned by the basis $\{\varphi_1, ..., \varphi_n\}$. An operator O defined in bra-ket notation with respect to this basis is written as:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij} \ |\varphi_i\rangle\langle\varphi_j|$$

Any ket $|\Psi\rangle \in V$ can be written as a linear combination $\sum_{k=1}^{N} p_k \ |\varphi_k\rangle$. Applying O to this ket gives:

$$O \ |\Psi\rangle = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij} \ |\varphi_i\rangle\langle\varphi_j|\right)\left(\sum_{k=1}^{n} p_k \ |\varphi_k\rangle\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} p_{ij}p_k \ |\varphi_i\rangle\langle\varphi_j|\varphi_k\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}p_k \ |\varphi_i\rangle$$

This is because, being all orthonormal, all bra-kets $\langle\varphi_j|\varphi_k\rangle$ with $j \neq k$ are 0 and those with $j = k$ are 1.

Exercise 2.3.1: Consider the operator $X = |10\rangle\langle00| + |00\rangle\langle10| + |11\rangle\langle11| + |01\rangle\langle01|$. What happens when applied to the state $|\Phi^+\rangle = \frac{\sqrt{2}}{2} \ |00\rangle + \frac{\sqrt{2}}{2} \ |11\rangle$?

Solution:

$$\begin{aligned}
X \ |\Phi^+\rangle &= \sum_{i=1}^{4}\sum_{j=1}^{4}\sum_{k=1}^{2} p_{ij}p_k \ |\varphi_i\rangle\langle\varphi_j|\varphi_k\rangle = \\
&= \sum_{i=1}^{4}\sum_{j=1}^{4} p_{ij}\frac{\sqrt{2}}{2} \ |\varphi_i\rangle\langle\varphi_j|00\rangle + p_{ij}\frac{\sqrt{2}}{2} \ |\varphi_i\rangle\langle\varphi_j|11\rangle = \\
&= \sum_{i=1}^{4} p_{i1}\frac{\sqrt{2}}{2} \ |\varphi_i\rangle\langle00|00\rangle + p_{i1}\frac{\sqrt{2}}{2} \ |\varphi_i\rangle\langle00|11\rangle + p_{i2}\frac{\sqrt{2}}{2} \ |\varphi_i\rangle\langle01|00\rangle + p_{i2}\frac{\sqrt{2}}{2} \ |\varphi_i\rangle\langle01|11\rangle + \\
&\quad p_{i3}\frac{\sqrt{2}}{2} \ |\varphi_i\rangle\langle10|00\rangle + p_{i3}\frac{\sqrt{2}}{2} \ |\varphi_i\rangle\langle10|11\rangle + p_{i4}\frac{\sqrt{2}}{2} \ |\varphi_i\rangle\langle11|00\rangle + p_{i4}\frac{\sqrt{2}}{2} \ |\varphi_i\rangle\langle11|11\rangle = \\
&= \sum_{i=1}^{4} p_{i1}\frac{\sqrt{2}}{2} \ |\varphi_i\rangle + p_{i4}\frac{\sqrt{2}}{2} \ |\varphi_i\rangle = 0\cdot\frac{\sqrt{2}}{2} \ |00\rangle + 0\cdot\frac{\sqrt{2}}{2} \ |00\rangle + 0\cdot\frac{\sqrt{2}}{2} \ |01\rangle + 0\cdot\frac{\sqrt{2}}{2} \ |01\rangle + \\
&\quad 1\cdot\frac{\sqrt{2}}{2} \ |10\rangle + 0\cdot\frac{\sqrt{2}}{2} \ |10\rangle + 0\cdot\frac{\sqrt{2}}{2} \ |11\rangle + 1\cdot\frac{\sqrt{2}}{2} \ |11\rangle = \frac{\sqrt{2}}{2} \ |10\rangle + \frac{\sqrt{2}}{2} \ |11\rangle
\end{aligned}$$

□

Again from Theorem 1.2.1, for any vector space V there exist a subspace S of V such that $V = S \oplus S^{\perp}$, where $S^{\perp}$ is the vector space that contains all vectors perpendicular to S. This means that any vector $|v\rangle \in V$ can be written in a unique way as $\boldsymbol{s}_1 + \boldsymbol{s}_2$, with $\boldsymbol{s}_1 \in S$ and $\boldsymbol{s}_2 \in S^{\perp}$.

For any S of the sort it is possible to construct a linear operator $P_S : S \rightarrow V$, called **projection operator**, or *projectors* for short, that maps $|v\rangle$ to s_1. To construct a projector, it is sufficient to take the outer product of a vector with itself. As a matter of fact, a projector simply "extracts" the "contribution" given by a basis to the state: the product between said basis and the component of the state with respect to the basis. Given one of the subspaces S and a basis $\{|\alpha_1\rangle, ..., |\alpha_s\rangle\}$ for this subspace, the projector P_S that returns the contribution given by $\{|\alpha_1\rangle, ..., |\alpha_s\rangle\}$ to $|\Psi\rangle$ is simply given by:

$$P_S = \sum_{i=1}^{s} |\alpha_i\rangle\langle\alpha_i| = |\alpha_1\rangle\langle\alpha_1| + \dots + |\alpha_s\rangle\langle\alpha_s|$$

In general, given a state space V, for any direct sum decomposition $V = S_1 \oplus \dots \oplus S_k$ into orthogonal subspaces, there exist i distinct projection operators P_i, each mapping a state $|v\rangle \in V$ to a vector $\boldsymbol{s}_i$ belonging the i-th subspace S_i. Framed this way, a measuring device with direct sum decomposition $V = S_1 \oplus \dots \oplus S_k$ acting on a state $|\Psi\rangle$ results in the state

$$|\phi\rangle = \frac{P_i\ |\Psi\rangle}{||P_i\ |\Psi\rangle||}$$

With probability $||P_i\ |\Psi\rangle||^2$. The ratio is necessary to (re)normalize the resulting vector in response to the measurement. A projector is well-defined, since applying a projector two times results in no difference.

When a measuring device with associated direct sum decomposition V = $S_1 \oplus \dots \oplus S_k$ interacts with an n-qubit system in state $|\Psi\rangle$, the interaction changes the state to one entirely contained within one of the subspaces, and chooses the subspace with probability equal to the square of the absolute value of the amplitude of the component of $|\Psi\rangle$ in that subspace.

Any projector is not only a projector in the mathematical sense, but is also self-ajoint. It is therefore easy to compute $||P_S\ |\Psi\rangle||^2$:

$$||P_S\ |\Psi\rangle||^2 = (P_S\ |\Psi\rangle)^\dagger P_S\ |\Psi\rangle = \langle\Psi|P_S|\Psi\rangle$$

Exercise 2.3.2: Consider the two-qubit state $|\Psi\rangle = \frac{1}{4}\ |00\rangle + \frac{1}{2}\ |01\rangle + \frac{1}{2}\ |10\rangle + \frac{\sqrt{7}}{4}\ |11\rangle$ and the subspace decomposition $S \oplus S^\perp$, where $S = \text{span}\{|00\rangle, |01\rangle, |10\rangle\}$. What is the projector with respect to this subspace? Which is the state reached after measurement? With which probability?

Solution: The projector associated to S is $P_S = |00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10|$. Applying the projector to $|\Psi\rangle$ gives:

$$P_S\ |\Psi\rangle = (|00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10|)\left(\frac{1}{4}\ |00\rangle + \frac{1}{2}\ |01\rangle + \frac{1}{2}\ |10\rangle + \frac{\sqrt{7}}{4}\ |11\rangle\right) =$$

$$= \frac{1}{4}\ |00\rangle\langle 00|00\rangle + \frac{1}{2}\ |00\rangle\langle 00|01\rangle + \frac{1}{2}\ |00\rangle\langle 00|10\rangle + \frac{\sqrt{7}}{4}\ |00\rangle\langle 00|11\rangle + \frac{1}{4}\ |01\rangle\langle 01|00\rangle +$$

$$= \frac{1}{2}\ |01\rangle\langle 01|01\rangle + \frac{1}{2}\ |01\rangle\langle 01|10\rangle + \frac{\sqrt{7}}{4}\ |01\rangle\langle 01|11\rangle + \frac{1}{4}\ |10\rangle\langle 10|00\rangle + \frac{1}{2}\ |10\rangle\langle 10|01\rangle +$$

$$= \frac{1}{2}\ |10\rangle\langle 10|10\rangle + \frac{\sqrt{7}}{4}\ |10\rangle\langle 10|11\rangle = \frac{1}{4}\ |00\rangle + \frac{1}{2}\ |01\rangle + \frac{1}{2}\ |10\rangle$$

Computing its Euclidean norm:

$$||P_S\ |\Psi\rangle|| = \left|\left|\frac{1}{4}\ |00\rangle + \frac{1}{2}\ |01\rangle + \frac{1}{2}\ |10\rangle\right|\right| = \sqrt{\left(\frac{1}{4}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \sqrt{\frac{9}{16}} = \frac{3}{4}$$

The new state reached with respect to this measurement process is:

$$|\phi\rangle = \frac{P_S\ |\Psi\rangle}{||P_S\ |\Psi\rangle||} = \frac{\frac{1}{4}\ |00\rangle + \frac{1}{2}\ |01\rangle + \frac{1}{2}\ |10\rangle}{\frac{3}{4}} = \frac{1}{3}\ |00\rangle + \frac{2}{3}\ |01\rangle + \frac{2}{3}\ |10\rangle$$

With probability:

$$||P_S\ |\Psi\rangle||^2 = \langle\Psi|P_S|\Psi\rangle = \left(\frac{3}{4}\right)^2 = \frac{9}{16}$$

□

As a matter of fact, the observation that any device has an associated direct sum decomposition is just a generalization of the single-qubit case. Every device measuring a single-qubit system has an associated orthonormal basis $\{|v_1\rangle, |v_2\rangle\}$ for the state space V of the system, having dimension 2. Each of these vectors generate a one-dimensional subspace, S_1 and S_2, consisting of all $\alpha\ |v_1\rangle$ and $\beta\ |v_2\rangle$ respectively, with $\alpha, \beta \in \mathbb{C}$, and $V = S_1 \oplus S_2$. Furthermore, the only nontrivial (with no subspaces of dimension 0) possible decompositions of V are the ones into two subspaces of dimension 1, and any choice of unit length vectors, one from each of the subspaces, yields an orthonormal basis.

Exercise 2.3.3: Rephrase the measurement of the single-qubit state $|\Psi\rangle = \frac{\sqrt{2}}{2}\ |0\rangle + \frac{\sqrt{2}}{2}\ |1\rangle$ under this formalism.

Solution: Let V be the vector space associated with said single-qubit system. A device that measures a qubit in the standard basis has associated the direct sum decomposition:

$$V = S_1 \oplus S_2 = \mathrm{span}\{|0\rangle\} \oplus \mathrm{span}\{|1\rangle\} = \mathrm{span}\{|0\rangle, |1\rangle\}$$

The state $|\Psi\rangle$ measured by such a device will be $|0\rangle$ with probability $\left|\frac{\sqrt{2}}{2}\right|^2 = \frac{1}{2}$, the amplitude of $|\Psi\rangle$ in the subspace S_1, and $|1\rangle$ with probability $\left|\frac{\sqrt{2}}{2}\right|^2 = \frac{1}{2}$, the amplitude of $|\Psi\rangle$ in the subspace S_2. □

This formalism allows one to construct measurements that don't simply refer to one or more base states.

Exercise 2.3.4: Consider the 2-qubit state $|\Psi\rangle = \frac{1}{8}\ |00\rangle + \frac{\sqrt{7}}{8}\ |01\rangle + \frac{\sqrt{7}}{8}\ |10\rangle + \frac{7}{8}\ |11\rangle$. What is the probability of observing the first qubit in state $|0\rangle$?

Solution: Measuring the first qubit in the state $|0\rangle$ is equivalent to measuring the overall system in a state that is either $|00\rangle$ or $|01\rangle$. Let V be the vector space associated with said 2-qubit system. A device that measures the first qubit in the standard basis has associated the direct sum decomposition:

$$\begin{aligned} V = S \oplus S^{\perp} &= (|0\rangle \otimes \mathrm{span}\{|0\rangle, |1\rangle\}) \oplus (|1\rangle \otimes \mathrm{span}\{|0\rangle, |1\rangle\}) = \mathrm{span}\{|00\rangle, |01\rangle\} \oplus \mathrm{span}\{|10\rangle, |11\rangle\} = \\ &= \mathrm{span}\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\} \end{aligned}$$

The projector associated to S is then $P_S = |00\rangle\langle 00| + |01\rangle\langle 01|$. Applying the projector to $|\Psi\rangle$ gives:

$$P_S\ |\Psi\rangle = (|00\rangle\langle 00| + |01\rangle\langle 01|)\left(\frac{1}{8}\ |00\rangle + \frac{\sqrt{7}}{8}\ |01\rangle + \frac{\sqrt{7}}{8}\ |10\rangle + \frac{7}{8}\ |11\rangle\right) = \frac{1}{8}\ |00\rangle + \frac{\sqrt{7}}{8}\ |01\rangle$$

Which means that the probability of measuring the first qubit in the state $|0\rangle$ is:

$$||P_S\ |\Psi\rangle||^2 = \langle\Psi|P_S|\Psi\rangle = \left|\left|\frac{1}{8}\ |00\rangle + \frac{\sqrt{7}}{8}\ |01\rangle\right|\right|^2 = \left(\frac{1}{8}\right)^2 + \left(\frac{\sqrt{7}}{8}\right)^2 = \frac{1}{64} + \frac{7}{64} = \frac{1}{8}$$

□

Explicitly writing out the direct sum decomposition associated to a measurement process can be tedious. A simpler way to specify the decomposition involves using Hermitian matrices. Recall how Theorem 1.3.4 states that the spectral decomposition of an Hermitian matrix is constituted by orthogonal subspaces, and for any decomposition it is possible to find an Hermitian matrix whose spectral decomposition is this decomposition.

Given a state space V, consider a subspace decomposition $V = S_1 \oplus \ldots \oplus S_k$. Let P_i be the projector associated to the subspace S_i and let $\{\lambda_1, \ldots, \lambda_k\}$ be a set of distinct real values. Then, the matrix:

$$O = \sum_{i=1}^{k} \lambda_i P_i = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_k \end{pmatrix}$$

Is an Hermitian operator (in matrix form) whose subspace decomposition is exactly $S_1 \oplus \dots \oplus S_k$. Thus, when describing a measurement, instead of directly specifying the associated subspace decomposition, it is sufficient to specify a Hermitian matrix whose spectral decomposition is that decomposition.

An Hermitian operator used for this purpose is also called **observable**. Any observable with the appropriate direct sum decomposition can be used to specify a given measurement; in particular, the values of the λ_i are irrelevant as long as they are distinct. The λ_i should be thought of simply as labels for the corresponding subspaces, or equivalently as labels for the measurement outcomes. In quantum mechanics, these labels are often chosen to represent a property, such as the energy, of the eigenstates in the corresponding eigenspace.

It is important to stress that it is not an Hermitian operator that acts on a state when measured, but instead the projectors associated to said operator. Indeed, applying a Hermitian operator to a quantum state might not even return a valid quantum state. The Hermitian operator is just a way, frequently used in quantum mechanics, to "pack" the projectors associated to a measuring apparatus in compact form, not to specify the result of the measurement itself.

Exercise 2.3.5: Write a measurement for a single-qubit system in the Hermitian operator form.

Solution: Let V be the vector space associated with a single-qubit system. A device that measures a qubit in the standard basis has associated the direct sum decomposition:

$$V = S_1 \oplus S_2 = \mathrm{span}\{|0\rangle\} \oplus \mathrm{span}\{|1\rangle\} = \mathrm{span}\{|0\rangle, |1\rangle\}$$

The projectors associated to the subspaces are $P_1 = |0\rangle\langle 0|$ and $P_2 = |1\rangle\langle 1|$ respectively. Picking two eigenvalues at random, say 1 and -1, the Hermitian operator associated to this measurement process is:

$$O = 1 \cdot P_1 + (-1) \cdot P_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

□

Exercise 2.3.6: Write the measurement for the first qubit in a two-qubit system in the Hermitian operator form.

Solution: Measuring the first qubit in the state $|0\rangle$ is equivalent to measuring the overall system in a state that is either $|00\rangle$ or $|01\rangle$. Let V be the vector space associated with a 2-qubit system. A device that measures the first qubit in the standard basis has associated the direct sum decomposition:

$$V = S_1 \oplus S_2 = (|0\rangle \otimes \mathrm{span}\{|0\rangle, |1\rangle\}) \oplus (|1\rangle \otimes \mathrm{span}\{|0\rangle, |1\rangle\}) = \mathrm{span}\{|00\rangle, |01\rangle\} \oplus \mathrm{span}\{|10\rangle, |11\rangle\} =$$
$$= \mathrm{span}\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$$

The projectors associated to the subspaces are $P_1 = |00\rangle\langle 00| + |01\rangle\langle 01|$ and $P_2 = |10\rangle\langle 10| + |11\rangle\langle 11|$ respectively. Picking two eigenvalues at random, say 1 and 0, the Hermitian operator associated to this measurement process is:

$$O = 1 \cdot P_1 + 0 \cdot P_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

□

Hermitian operators on higher dimensional vectors spaces can be merged with the tensor product. That is, if O_1 and O_2 are Hermitian operators defined for state spaces V_1 and V_2 respectively, then $O_1 \otimes O_2$ is an Hermitian operator defined for $V_1 \otimes V_2$. Furthermore, the eigenvalues of $O_1 \otimes O_2$ are the product of the eigenvalues of O_1 and O_2.

Not all Hermitian operators on the state space $V_1 \otimes V_2$ can be factored as tensor product of two Hermitian operators O_1 and O_2 acting on V_0 and V_1 respectively. Such factoring is possible only if each subspace in the subspace decomposition described by O can be written as $S_0 \otimes S_1$ for S_0 and S_1 in the subspace decompositions associated to O_1 and O_2 respectively.

A non-factorable Hermitian operator denotes a measurement on the state that cannot be conceived as measuring separately its smaller components.

Exercise 2.3.7: Can the following Hermitian operator be factored into smaller matrices?

$$O = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Solution: No. This is because the subspace decomposition described by the operator is $S_1 \oplus S_2$, where $S_1 = \text{span}\{|00\rangle, |01\rangle, |10\rangle\}$ and $S_2 = \text{span}\{|11\rangle\}$, and this decomposition is not factorable. □

2.4. Qubit manipulations

The manipulation of the state of an arbitrary-sized qubit system is done through **quantum transformations**. A quantum transformation is simply an operator that, when applied to a valid quantum state, results in a new quantum state that is still valid.

In order for this requirement to be followed, before introducing some examples of quantum transformations, it is necessary to clearly state what constraint a quantum transformation must abide to:

1. The Hilbert space of the possible qubit states should be the same before and after applying a transformation. A quantum transformation must then be an *endomorphism*, mapping elements from an Hilbert space to elements to the same space.
2. A quantum state is a linear combination (of base states), therefore a quantum transformation must be linear. In other words, given a state $a_1 \ |\varphi_1\rangle + ... + a_n \ |\varphi_n\rangle$ and a quantum transformation U, the following must hold:

$$U(a_1 \ |\varphi_1 + ... + a_n \ |\varphi_n\rangle\rangle = U(a_1 \ |\varphi_1\rangle) + ... + U(a_n \ |\varphi_n\rangle) = a_1 U(|\varphi_1\rangle) + ... + a_n U(|\varphi_n\rangle)$$

   This way, applying a quantum transformation to a superposition of states is the same as applying the transformation to each component of the superposition.
3. A quantum state is a unit vector, therefore a quantum transformation must return a state that is also a unit vector.
4. A quantum state is a linear combination of vectors from an orthonormal basis, therefore quantum transformations must map orthonormal bases to orthonormal bases.

Clearly, measurements cannot be considered quantum transformations.

Constraint number 2 imposes that the operator has a matrix representation with respect to a given basis. As stated in Theorem 1.3.3 the one and only kind of matrices that abide to constraint number 3 and 4 are the unitary matrices. This means that the set of unitary matrices and the set of valid quantum transformations are exactly the same set.

The fact that quantum transformations are unitary has important consequences. First, recall from Lemma 1.3.3 that unitary matrices are inner product-invariant. This means that measuring a state in the original basis and then applying a transformation to the outcome should give the same result as applying the transformation first and then measuring the resulting state in the transformed basis.

Second, recall from Theorem 1.3.2 that the product of two unitary matrices is also a unitary matrix. Therefore, applying more than one quantum transformation to a quantum state, which is equivalent to applying their function composition, will still result in a valid quantum state.

Any quantum state transformation that acts on only a small number of qubits is also referred to a **quantum gate**. Sequences of quantum gates are called **quantum gate arrays** or **quantum circuits**. Given a basic set of quantum gates, it is possible to combine them to construct elaborate transformations of varying complexity. For these reasons, the term "gate" and "transformation" tend to be used interchangeably.

The term "gate" is used to suggest a similarity with the classical logical gates, but does not necessarely entail that the physical implementation of a quantum transformation has to be a gate in the literal sense. Conceiving a quantum transformation as gates has the added advantage of abstracting the need to specify a basis when talking about operators.

It should also be noted that, whereas classical logical circuits can have loops (outputs that are fed back in the circuit as inputs), quantum circuits are said to be *acyclic*, meaning that they can't have loops.

Drawing quantum transformations as gates is obiquitous. Transformations are represented graphically by boxes, labeled by the transformation performed, that are to be read left-to-right. Boxes are connected with a line that represents the state of the qubit "flowing" through the circuit.

$$|\Psi\rangle - \boxed{U_1} - \boxed{U_2} - \dots - \boxed{U_n} - |\Psi'\rangle \qquad\qquad |\Psi'\rangle = U_n \dots U_2 U_1 \; |\Psi\rangle$$

Figure 2: On the left, a generic circuit acting on a state $|\Psi\rangle$ that returns a state $|\Psi'\rangle$. On the right, its analogous representation with matrix products.

When a new state is reached, there's most likely interest in sampling its value. Which is why the $|\Psi'\rangle$ symbol is often replaced by . Also, to denote a line that represents n states at once the shorthand notation $\diagup^{n}$ is used.

The simplest operator is the **identity operator**, denoted as I, leaves the state unchanged. With respect to the standard basis, the operator has this following form:

$$I = |0\rangle\langle 0| + |1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad\qquad - \boxed{I} -$$

Three matrices, called **Pauli matrices**[3], are also obiquitous:

$$X = |1\rangle\langle 0| + |0\rangle\langle 1| = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad Y = -|1\rangle\langle 0| + |0\rangle\langle 1| = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad Z = |0\rangle\langle 0| - |1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$- \boxed{X} - \qquad\qquad - \boxed{Y} - \qquad\qquad - \boxed{Z} -$$

Note how:

- X is equivalent to a classical `NOT` gate, since it changes $|0\rangle$ into $|1\rangle$ and vice versa. X has no effect on $|+\rangle$ and $|-\rangle$, therefore those states are its eigenvectors;
- Z changes the relative phase of a superposition in the standard basis. Z has no effect on $|0\rangle$ and $|1\rangle$, therefore those states are its eigenvectors;
- $Y = ZX$ is a combination of negation and phase change. Y has no effect on $|\circlearrowleft\rangle$ and $|\circlearrowright\rangle$, therefore those states are its eigenvectors.

Another useful operator is the **Hadamard operator**, denoted as H:

$$H = \frac{\sqrt{2}}{2}(|0\rangle\langle 0| + |1\rangle\langle 0| + |0\rangle\langle 1| - |1\rangle\langle 1|) = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad\qquad - \boxed{H} -$$

[3]Other sources refer to the Pauli matrices respectively as $\sigma_x, \sigma_y, \sigma_z$

Exercise 2.4.1: What happens when the Hadamard operator is applied to the states $|0\rangle, |1\rangle, |+\rangle, |-\rangle$?

Solution:

$$
\begin{aligned}
H\ |0\rangle &= \frac{\sqrt{2}}{2}(|0\rangle\langle 0| + |1\rangle\langle 0| + |0\rangle\langle 1| - |1\rangle\langle 1|)\ |0\rangle = \frac{\sqrt{2}}{2}(|0\rangle + |1\rangle) = |+\rangle \\
H\ |1\rangle &= \frac{\sqrt{2}}{2}(|0\rangle\langle 0| + |1\rangle\langle 0| + |0\rangle\langle 1| - |1\rangle\langle 1|)\ |1\rangle = \frac{\sqrt{2}}{2}(|0\rangle - |1\rangle) = |-\rangle \\
H\ |+\rangle &= \frac{\sqrt{2}}{2}(|0\rangle\langle 0| + |1\rangle\langle 0| + |0\rangle\langle 1| - |1\rangle\langle 1|)\frac{\sqrt{2}}{2}(|0\rangle + |1\rangle) = \frac{1}{2}(2\ |0\rangle) = |0\rangle \\
H\ |-\rangle &= \frac{\sqrt{2}}{2}(|0\rangle\langle 0| + |1\rangle\langle 0| + |0\rangle\langle 1| - |1\rangle\langle 1|)\frac{\sqrt{2}}{2}(|0\rangle - |1\rangle) = \frac{1}{2}(2\ |1\rangle) = |1\rangle
\end{aligned}
$$

□

The previous quantum transformations acted on single qubits, but extending the application of quantum transformations to more than one qubit at once is trivial. Suppose U is a unitary transformation that acts on a qubit: the unitary transformation $U^{\otimes n} = U \otimes U \otimes \dots \otimes U$ acts on n qubits at the same time.

In general, given n single-qubit transformations $U_1, U_2, \dots, U_n$, not necessarely identical, the unitary transformation $U_1 \otimes U_2 \otimes \dots \otimes U_n$ applies U_1 to the first qubit state, U_2 to the second qubit state, and so on, all at the same time. This also allows to have composite transformations that act on certain qubit states and ignore others, since applying the identity operator to a qubit state is equivalent to doing nothing.

Exercise 2.4.2: Consider a 3-qubit state. Apply a Hadamard transformation to the first qubit state and a Y transformation on the third qubit state. How would the resulting matrix look like?

Solution:

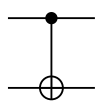
$$
\frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 0 & -i & 0 & 0 & 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 & i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i & 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 & 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 & 0 & i & 0 & 0 \\ i & 0 & 0 & 0 & -i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i & 0 & 0 & 0 & i \\ 0 & 0 & i & 0 & 0 & 0 & -i & 0 \end{pmatrix}
$$

□

Since it's not possible to conceive an entangled state simply as the sum of its parts, trasformations that act on single qubits cannot influence (create or destroy) the entanglement of states. Just as entangled states cannot be factorized into single qubit states, transformations that act on entangled states cannot be factorized into a tensor product between single-qubit transformations.

A qubit transformation such as these is the **controlled not gate**, or C_{not} for short. The gate acts on two qubits as follows:

$$
C_{\text{not}} = |00\rangle\langle 00| + |01\rangle\langle 01| + |11\rangle\langle 10| + |10\rangle\langle 11| = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}
$$



If the state of the single qubits are thought of as the values of classical bits, the C_{not} gate can be conceived as "flipping" the second bit (the state of the second qubit) if the first bit is 1 (if the first qubit is in state $|1\rangle$) and leave it unchanged otherwise.

Exercise 2.4.3: What is the effect of applying the C_{not} gate to the two-qubit state $\frac{\sqrt{2}}{2}(|00\rangle + |10\rangle)$?

Solution:

$$(|00\rangle\langle 00| + |01\rangle\langle 01| + |11\rangle\langle 10| + |10\rangle\langle 11|)\frac{\sqrt{2}}{2}(|00\rangle + |10\rangle) = \frac{\sqrt{2}}{2}(|00\rangle + |11\rangle)$$

The starting state is not entangled; the final state is. □

For its analogy with classical control gates, the state of the first qubit (the first bit) is also referred to as the **control bit**, whereas the state of the second qubit (the second bit) is also referred to as the **target bit**. However, this terminology might be misleading, since states expressed in different bases than the standard basis might result in the control bit becoming the target bit and viceversa, or having both bits changed.

Exercise 2.4.4: What happens when a C_{not} gate is applied to the state $|+-\rangle$?

Solution: Converting $|+-\rangle$ to the standard basis gives:

$$|+-\rangle = |+\rangle \otimes |-\rangle = \frac{\sqrt{2}}{2}(|0\rangle + |1\rangle) \otimes \frac{\sqrt{2}}{2}(|0\rangle - |1\rangle) = \frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle)$$

Applying C_{not} to this state gives:

$$(|00\rangle\langle 00| + |01\rangle\langle 01| + |11\rangle\langle 10| + |10\rangle\langle 11|)\frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle) = \frac{1}{2}(|00\rangle - |01\rangle - |10\rangle + |11\rangle)$$

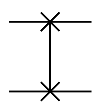
Converting it back:

$$\frac{1}{2}(|00\rangle - |01\rangle - |10\rangle + |11\rangle) = \frac{1}{2}(|0\rangle \otimes (|0\rangle - |1\rangle) - |1\rangle \otimes (|0\rangle - |1\rangle)) = \frac{1}{2}((|0\rangle - |1\rangle) \otimes (|0\rangle - |1\rangle)) =$$
$$= \frac{1}{2}\left(\sqrt{2}\,|-\rangle \otimes \sqrt{2}\,|-\rangle\right) = \frac{1}{2}(2\,|--\rangle) = |--\rangle$$

Which means that, in the Hadamard basis, the control bit and the target bit are in reversed! □

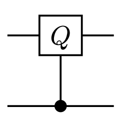
Another transformation that acts on two-qubit systems is the **swap gate**, that changes the state of the first qubit to be equal to the state of the second qubit and vice versa:

$$\text{SWAP} = |00\rangle\langle 00| + |01\rangle\langle 10| + |10\rangle\langle 01| + |11\rangle\langle 11| = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$



The controlled not gate and the swap gate are examples of a more general class of two-qubit controlled gates, where the gate performs a certain transformation Q on the second qubit when the first qubit is in state $|1\rangle$ and does nothing when the first qubit is in state $|0\rangle$. Any gate in this form can be written as $\bigwedge Q$:

$$\bigwedge Q = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes Q = \begin{pmatrix} I & 0 \\ 0 & Q \end{pmatrix}$$



For example, the controlled not gate is equivalent to $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes X$. Another example is the **controlled phase shift**, that changes the phase of the second qubit if the first qubit is in state $|1\rangle$ and does nothing otherwise:

$$\bigwedge e^{i\varphi} = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes Ie^{i\varphi} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{i\varphi} & 0 \\ 0 & 0 & 0 & e^{i\varphi} \end{pmatrix}$$

This gate is interesting because it employs a phase shift that when applied to single-qubit systems it just changes the global phase, and therefore is physically meaningless, whereas when applied as part of a conditional transformation it changes the relative phase between elements of a superposition, which is physically meaningful.

As a matter of fact, all single-qubit transformations can be reduced to a combination of three types of transformation: a *phase shift* $K(\delta)$, a *rotation* $R(\beta)$ and a *phase rotation* $T(\alpha)$:

$$K(\delta) = \begin{pmatrix} e^{i\delta} & 0 \\ 0 & e^{i\delta} \end{pmatrix} \qquad R(\beta) = \begin{pmatrix} \cos(\beta) & \sin(\beta) \\ -\sin(\beta) & \cos(\beta) \end{pmatrix} \qquad T(\alpha) = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & -e^{i\alpha} \end{pmatrix}$$

The transformation $R(\alpha)$ and $T(\alpha)$ corresponds to rotations by an angle of 2α along the y and z axis of the Bloch sphere respectively. For this reason, they are also referred to as *zenithal rotation* and *azimuthal rotation*.

Phase rotations of $\pi/2$ radiants and $\pi/4$ radiants are quite obiquitous, therefore they have been given proper names: S and T respectively:

$$P_{\frac{\pi}{2}} = S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \qquad P_{\frac{\pi}{4}} = T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix}$$

It would be useful to have a small set of elementary gates that could be used to construct an arbitrary complicated function. Sadly, this is not possible, because as stated some gates are not decomposable in smaller gates and must be implemented as-is. However, it is possible to employ a subset of gates that can approximate with sufficient accuracy any kind of gate, even those that cannot be decomposed.

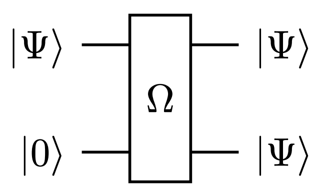
Another difference with classical computing is the concept of **cloning**, creating exact copies of states. A classical state (i.e. a Boolean variable) can be cloned effortlessly, both logically and physically (by splitting and, if needed, increasing the amplitude the electric current).

At first glance it might seem that quantum states would present no difference. Indeed, if a quantum state is in an eigenstate (that is, if it has been measured) there is no issue with cloning it. It is easy to notice how, for example, CNOT gates can be used to duplicate the state of the control qubit if its state is an eigenstate of the standard basis, since $\mathrm{CNOT}_{1,2}\ |00\rangle = |00\rangle$ and $\mathrm{CNOT}_{1,2}\ |10\rangle = |11\rangle$. However, when states are superpositions of base states, this doesn't hold anymore.

Theorem 2.4.1 (No-cloning theorem): Unknown (i.e. not measured) quantum states cannot be cloned.

Proof: The theorem can be proven by contradiction, considering for simplicity the case of a single qubit (the generalization to the n qubit case is trivial).

First, suppose that there exists a quantum transformation Ω that is capable of duplicating an unknown quantum state $|\Psi\rangle$. This transformation must be a unitary transformation with two inputs, $|\Psi\rangle$ and an ancillary qubit $|0\rangle$, and two outputs $|\Psi\rangle$ and $|\Psi\rangle$:



Since $|\Psi\rangle$ is not known, it can be expressed as a superposition $\alpha\ |\varphi_1\rangle + \beta\ |\varphi_2\rangle$ of two basis states $|\varphi_1\rangle$ and $|\varphi_2\rangle$, where α and β are both not null. As in the case of the CNOT gate for the $\{|0\rangle, |1\rangle\}$ basis, it can be assumed that Ω is capable of duplicating both basis states.

Since Ω is has two states as input and two states and output, its two single bit states for input and output must be merged using the tensor product. The input to omega is therefore $|\Psi\rangle \otimes |0\rangle$ and its output is $|\Psi\rangle \otimes |\Psi\rangle$.

If Ω were to exist, then applying it to $|\Psi\rangle \otimes |0\rangle$ would return $|\Psi\rangle \otimes |\Psi\rangle$. However:

$$
\begin{aligned}
&\Omega(|\Psi\rangle \otimes |0\rangle) = |\Psi\rangle \otimes |\Psi\rangle \Rightarrow \\
&\Omega((\alpha\ |\varphi_1\rangle + \beta\ |\varphi_2\rangle) \otimes |0\rangle) = (\alpha\ |\varphi_1\rangle + \beta\ |\varphi_2\rangle) \otimes (\alpha\ |\varphi_1\rangle + \beta\ |\varphi_2\rangle) \Rightarrow \\
&\Omega(\alpha\ |\varphi_1\rangle \otimes |0\rangle + \beta\ |\varphi_2\rangle \otimes |0\rangle) = (\alpha\ |\varphi_1\rangle + \beta\ |\varphi_2\rangle) \otimes (\alpha\ |\varphi_1\rangle + \beta\ |\varphi_2\rangle) \Rightarrow \\
&\Omega(\alpha\ |\varphi_1 0\rangle + \beta\ |\varphi_2 0\rangle) = \alpha^2\ |\varphi_1\varphi_1\rangle + \alpha\beta\ |\varphi_1\varphi_2\rangle + \alpha\beta\ |\varphi_2\varphi_1\rangle + \beta^2\ |\varphi_2\varphi_2\rangle \Rightarrow \\
&\alpha\Omega\ |\varphi_1 0\rangle + \beta\Omega\ |\varphi_2 0\rangle = \alpha^2\ |\varphi_1\varphi_1\rangle + \alpha\beta(|\varphi_1\varphi_2\rangle + |\varphi_2\varphi_1\rangle) + \beta^2\ |\varphi_2\varphi_2\rangle \Rightarrow \\
&\alpha\ |\varphi_1\varphi_1\rangle + \beta\ |\varphi_2\varphi_2\rangle = \alpha^2\ |\varphi_1\varphi_1\rangle + \alpha\beta(|\varphi_1\varphi_2\rangle + |\varphi_2\varphi_1\rangle) + \beta^2\ |\varphi_2\varphi_2\rangle \Rightarrow \\
&(\alpha^2 - \alpha)\ |\varphi_1\varphi_1\rangle + \alpha\beta(|\varphi_1\varphi_2\rangle + |\varphi_2\varphi_1\rangle) + (\beta^2 - \beta)\ |\varphi_2\varphi_2\rangle = 0
\end{aligned}
$$

Since both $|\varphi_1\rangle$ and $|\varphi_2\rangle$ are not null by definition, the only way for the left hand side to be null is to have the three coefficients $(\alpha^2 - \alpha)$, $\alpha\beta$, and $(\beta^2 - \beta)$ all equal to 0 at the same time. However, having $\alpha\beta$ equal to 0 means having either $\alpha = 0$ or $\beta = 0$, and both possibilities are not admissible by hypothesis.

Therefore, it must mean that there is no such thing as a unitary transformation with the properties of Ω and the theorem is proven. □

Note that, even though Theorem 2.4.1 forbids cloning unknown quantum states, it doesn't prevent an unknown quantum state to be *transferred* from one qubit to another, losing the information regarding the original state in the process (this aspect is explored further in the next chapter). Also, even though perfect clones cannot be created, there are techniques to create "imperfect" clones, even with an high degree of approximation.

3. Quantum Algorithms

3.1. Introduction to algorithms

Quantum algorithms are the counterpart to classical algorithms, a set of well-defined procedures that use quantum operators instead of classical operators.

Quantum algorithms, as expected, must operate on quantum information (elements of an Hilbert space), but real world information is generally classical information. Therefore, the first step in constructing a quantum algorithm is to device a method of representing classical information as quantum information. That is, defining a **quantum embedding**.

The simplest form of quantum embedding is **base embedding**, where classical bits are mapped to base states. This means that a binary string $b_1 b_2 ... b_n$ is mapped to the state $|b_1 b_2 ... b_n\rangle$. Of course, this embedding is only possible if the input is binary, but since all strings can be encoded into a binary alphabet in a unique way, theoretically speaking this is not restrictive. It should be noted, however, that this embedding might be wasteful and/or cumbersome, since to represent n classical bits, just as many qubits are needed.

Quantum transformations are carried out by unitary matrices, all having a defined inverse. This means that, when presented with an output, it is possible to recover the original input without any loss of information simply by multiplying the output with the inverse of the transformation. In other words, quantum computation is **reversible**.

Classical computation, on the other hand, is in general not reversible: if an output of a circuit is presented, it may not be possible to recover the original input. For example, whereas the logical `NOT` is reversible, the logical `AND` is not. This is not a setback, however, because any classical function can be adjusted to become reversible.

Exercise 3.1.1: Why is the logical `AND` not reversible?

Solution: Let A and B be two classical bits. Consider $A \wedge B$: if the output is 1, then it is known for sure that both A and B were equal to 1. On the other hand, if the output is 0, there are three possibilities: $A = 0$ and $B = 0$, $A = 1$ and $B = 0$, $A = 0$ and $B = 1$. Not having other prior information, all of these possibilities are equally probable. □

First, consider a reversible classical function with n input and n output bits. The output of this function is just a permutation 2^n of bit strings given in input. This means that for any classical reversible function there is a permutation $\pi : \boldsymbol{Z}_{2^n} \rightarrow \boldsymbol{Z}_{2^n}$ that maps input strings to output strings in the exact same way as the original function. This permutation can be used, without any additional modification, to define an equivalent quantum transformation:

$$U_\pi : \sum_{x=0}^{2^n-1} a_x \ |x\rangle \rightarrow \sum_{x=0}^{2^n-1} a_x \ |\pi(x)\rangle$$

Now consider a non-reversible classical function $f : \boldsymbol{Z}_{2^n} \rightarrow \boldsymbol{Z}_{2^m}$ with n input and m output bits. This function can be modified in a standard way to create a reversible function $\pi_f : \boldsymbol{Z}_{2^{n+m}} \rightarrow \boldsymbol{Z}_{2^{n+m}}$ that does the exact same thing, but is reversible.

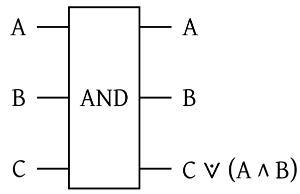
This function acts on two subset of bits, a set of n bits that contains the input and a set of m bits. Both sets are given to the function as input and both are present as output. Each pair (x, y) of input-output bits is mapped by the function to the pair $(x, y \veebar f(x))$, where $\veebar$ denotes the logical `XOR`[4] and f is the original, non reversible function. In other words, π_f simply applies the original function f to x and returns both the original input unchanged and the actual value of $f(x)$, stored in y.

[4] A much more common notation for the logical `XOR` is $\oplus$, but this notation is here avoided because it conflicts with the direct sum symbol.

Exercise 3.1.2: Construct a reversible version of the logical `AND`.

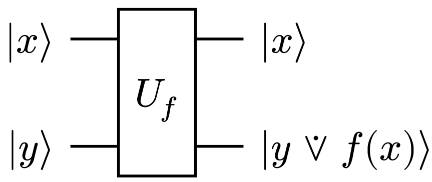
Solution:

Inputs			**Outputs**		
A	**B**	**C**	**A**	**B**	**C $\dot{\vee}$ (A ∧ B)**
0	0	0	0	0	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	0	1	1
1	0	0	1	0	0
1	0	1	1	0	1
1	1	0	1	1	1
1	1	1	1	1	0



□

Since this new function π_f is now reversible, it is possible to construct a unitary transformation $U_f : |x, y\rangle \to |x, y \dot{\vee} f(x)\rangle$ that implements the function, depicted as follows:

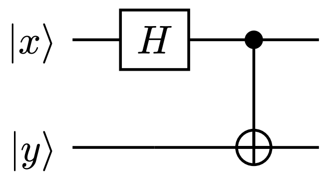


Note that, in general, this method of constructing reversible counterparts of non-reversible functions is highly inefficient, and there are ad-hoc methods that use less bits. This is not important, however, since the interest is to show that there is a method that always works, and therefore that each function that a classical computer can compute can be just as well computed by a quantum computer.

3.2. Quantum circuits for known states

3.2.1. Bell states

The following circuit generate the four different Bell states:



Applying the Hadamard gate to the first qubit:

$$|\Psi_1\rangle = (H \otimes I)(|x\rangle \otimes |y\rangle) = (H\ |x\rangle \otimes I\ |y\rangle) = \frac{\sqrt{2}}{2}((|0\rangle + (-1)^x\ |1\rangle) \otimes |y\rangle) =$$

$$= \frac{\sqrt{2}}{2}(|0y\rangle + (-1)^x\ |1y\rangle)$$

Applying a CNOT gate with first qubit as control and second as target:

$$|\Psi_2\rangle = \text{CNOT}_{1,2}\frac{\sqrt{2}}{2}(|0y\rangle + (-1)^x\ |1y\rangle) = \frac{\sqrt{2}}{2}\big(\text{CNOT}_{1,2}\ |0y\rangle + (-1)^x\text{CNOT}_{1,2}\ |1y\rangle\big) =$$

$$= \frac{\sqrt{2}}{2}(|0y\rangle + (-1)^x\ |1(y \veebar 0)\rangle)$$

Going over all the possible values of x and y it is possible to construct the entire Bell basis:

$$\text{with } x = 0, y = 0 \Longrightarrow \frac{\sqrt{2}}{2}(|00\rangle + (-1)^0\ |1(0 \veebar 0)\rangle) = \frac{\sqrt{2}}{2}(|00\rangle + |11\rangle) = |\Phi^+\rangle$$

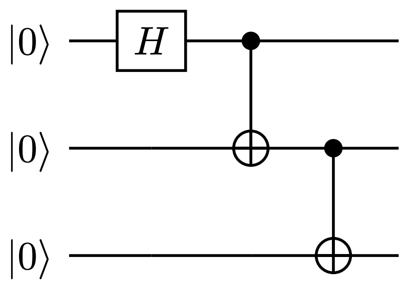
$$\text{with } x = 0, y = 1 \Longrightarrow \frac{\sqrt{2}}{2}(|01\rangle + (-1)^0\ |1(1 \veebar 0)\rangle) = \frac{\sqrt{2}}{2}(|01\rangle + |10\rangle) = |\Psi^+\rangle$$

$$\text{with } x = 1, y = 0 \Longrightarrow \frac{\sqrt{2}}{2}(|00\rangle + (-1)^1\ |1(0 \veebar 0)\rangle) = \frac{\sqrt{2}}{2}(|00\rangle - |11\rangle) = |\Phi^-\rangle$$

$$\text{with } x = 1, y = 1 \Longrightarrow \frac{\sqrt{2}}{2}(|01\rangle + (-1)^1\ |1(1 \veebar 0)\rangle) = \frac{\sqrt{2}}{2}(|01\rangle - |10\rangle) = |\Psi^-\rangle$$

3.2.2. GHZ state

The following circuit has 3 qubits prepared in state $|0\rangle$ and ends in the GHZ state:



Applying a Hadamard gate to the first qubit:

$$|\Psi_1\rangle = H\ |\Psi_0\rangle = (H \otimes I \otimes I)\ |000\rangle = (H\ |0\rangle \otimes I\ |0\rangle \otimes I\ |0\rangle) = \frac{\sqrt{2}}{2}(|0\rangle + |1\rangle) \otimes |00\rangle = \frac{\sqrt{2}}{2}(|000\rangle + |100\rangle)$$

Applying a CNOT gate with first qubit as control and second qubit as target:

$$|\Psi_2\rangle = \big(\text{CNOT}_{1,2} \otimes I\big)\ |\Psi_1\rangle = \big(\text{CNOT}_{1,2} \otimes I\big)\frac{\sqrt{2}}{2}(|000\rangle + |100\rangle) =$$

$$= \frac{\sqrt{2}}{2}(\text{CNOT}_{1,2} \otimes I)((|00\rangle + |10\rangle) \otimes |0\rangle) = \frac{\sqrt{2}}{2}\big(\text{CNOT}_{1,2}(|00\rangle + |10\rangle) \otimes I\ |0\rangle\big) =$$

$$= \frac{\sqrt{2}}{2}\big((\text{CNOT}_{1,2}\ |00\rangle + \text{CNOT}_{1,2}\ |10\rangle) \otimes |0\rangle\big) = \frac{\sqrt{2}}{2}((|00\rangle + |11\rangle) \otimes |0\rangle) = \frac{\sqrt{2}}{2}(|000\rangle + |110\rangle)$$

Applying a CNOT gate with second qubit as control and third qubit as target:

$$|\Psi_3\rangle = \big(I \otimes \text{CNOT}_{2,3}\big)\ |\Psi_2\rangle = \frac{\sqrt{2}}{2}\big(I \otimes \text{CNOT}_{2,3}\big)(|0\rangle \otimes |00\rangle + |1\rangle \otimes |10\rangle) =$$

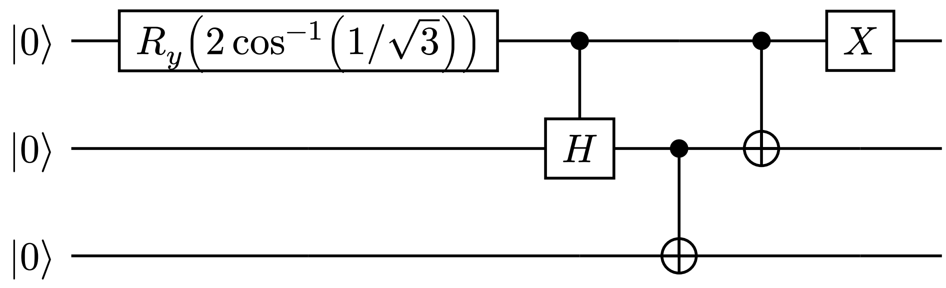
$$= \frac{\sqrt{2}}{2}\big(I \otimes \text{CNOT}_{2,3}\big)(|0\rangle \otimes |00\rangle) + \frac{\sqrt{2}}{2}\big(I \otimes \text{CNOT}_{2,3}\big)(|1\rangle \otimes |10\rangle) =$$

$$= \frac{\sqrt{2}}{2}\big(I\ |0\rangle \otimes \text{CNOT}_{2,3}\ |00\rangle\big) + \frac{\sqrt{2}}{2}\big(I\ |1\rangle \otimes \text{CNOT}_{2,3}\ |10\rangle\big) =$$

$$= \frac{\sqrt{2}}{2}(|0\rangle \otimes |00\rangle + |1\rangle \otimes |11\rangle) = \frac{\sqrt{2}}{2}(|000\rangle + |111\rangle)$$

3.2.3. W state

The following circuit has 3 qubits prepared in state $|0\rangle$ and ends in the GHZ state:



The rotation matrix is given by:

$$R_y(2\cos^{-1}\left(1/\sqrt{3}\right) = \begin{pmatrix} \cos\left(\cos^{-1}\left(1/\sqrt{3}\right)\right) & \sin\left(\cos^{-1}\left(1/\sqrt{3}\right)\right) \\ -\sin\left(\cos^{-1}\left(1/\sqrt{3}\right)\right) & \cos\left(\cos^{-1}\left(1/\sqrt{3}\right)\right) \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \sqrt{\frac{2}{3}} \\ -\sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

Applying a rotation on the y axis on the first qubit:

$$|\Psi_1\rangle = \left(\begin{pmatrix} \frac{1}{\sqrt{3}} & \sqrt{\frac{2}{3}} \\ -\sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \otimes I \otimes I\right) |000\rangle = \begin{pmatrix} \frac{1}{\sqrt{3}} & \sqrt{\frac{2}{3}} \\ -\sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} \otimes I\ |0\rangle \otimes I\ |0\rangle = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \sqrt{\frac{2}{3}} \end{pmatrix} \otimes |00\rangle =$$

$$= \frac{1}{\sqrt{3}}\begin{pmatrix}1\\ \sqrt{2}\end{pmatrix} \otimes |00\rangle = \frac{1}{\sqrt{3}}\left(\begin{pmatrix}1\\0\end{pmatrix} + \sqrt{2}\begin{pmatrix}0\\1\end{pmatrix}\right) \otimes |00\rangle = \left(\frac{1}{\sqrt{3}}\begin{pmatrix}1\\0\end{pmatrix} + \sqrt{\frac{2}{3}}\begin{pmatrix}0\\1\end{pmatrix}\right) \otimes |00\rangle =$$

$$= \left(\frac{1}{\sqrt{3}}\ |0\rangle + \sqrt{\frac{2}{3}}\ |1\rangle\right) \otimes |00\rangle = \frac{1}{\sqrt{3}}\ |000\rangle + \sqrt{\frac{2}{3}}\ |100\rangle$$

The controlled Hadamard is given by:

$$\text{CH} = (|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes Q) \otimes I = |0\rangle\langle 0| \otimes I \otimes I + |1\rangle\langle 1| \otimes Q \otimes I$$

Applying a controlled Hadamard with the first qubit as control and the second as target:

$$|\Psi_2\rangle = \text{CH}\ |\Psi_1\rangle = (|0\rangle\langle 0| \otimes I \otimes I + |1\rangle\langle 1| \otimes Q \otimes I)\left(\frac{1}{\sqrt{3}}\ |000\rangle + \sqrt{\frac{2}{3}}\ |100\rangle\right) =$$

$$= (|0\rangle\langle 0| \otimes I \otimes I + |1\rangle\langle 1| \otimes Q \otimes I)\frac{1}{\sqrt{3}}\ |000\rangle + (|0\rangle\langle 0| \otimes I \otimes I + |1\rangle\langle 1| \otimes Q \otimes I)\sqrt{\frac{2}{3}}\ |100\rangle =$$

$$= \frac{1}{\sqrt{3}}(|0\rangle\langle 0| \otimes I \otimes I)\ |000\rangle + \cancel{\frac{1}{\sqrt{3}}(|1\rangle\langle 1| \otimes Q \otimes I)\ |000\rangle} + \cancel{\sqrt{\frac{2}{3}}(|0\rangle\langle 0| \otimes I \otimes I)\ |100\rangle} +$$

$$\sqrt{\frac{2}{3}}(|1\rangle\langle 1| \otimes Q \otimes I)\ |100\rangle = \frac{1}{\sqrt{3}}(|0\rangle \otimes I\ |0\rangle \otimes I\ |0\rangle) + \sqrt{\frac{2}{3}}(|1\rangle \otimes Q\ |0\rangle \otimes I\ |0\rangle) =$$

$$= \frac{1}{\sqrt{3}}(|0\rangle \otimes |0\rangle \otimes |0\rangle) + \sqrt{\frac{2}{3}}\left(|1\rangle \otimes \frac{\sqrt{2}}{2}(|0\rangle + |1\rangle) \otimes |0\rangle\right) =$$

$$= \frac{1}{\sqrt{3}}\ |000\rangle + \sqrt{\frac{2}{3}}\left(\frac{\sqrt{2}}{2}(|100\rangle + |110\rangle)\right) = \frac{1}{\sqrt{3}}\ |000\rangle + \frac{1}{\sqrt{3}}\ |100\rangle + \frac{1}{\sqrt{3}}\ |110\rangle$$

Applying a CNOT gate with second qubit as control and third qubit as target:

$$
\begin{aligned}
\left|\Psi_3\right\rangle &= \left(I \otimes \mathrm{CNOT}_{2,3}\right) \left|\Psi_2\right\rangle = \left(I \otimes \mathrm{CNOT}_{2,3}\right)\left(\frac{1}{\sqrt{3}} \left|000\right\rangle + \frac{1}{\sqrt{3}} \left|100\right\rangle + \frac{1}{\sqrt{3}} \left|110\right\rangle\right) = \\
&= \left(I \otimes \mathrm{CNOT}_{2,3}\right)\left(\frac{1}{\sqrt{3}} \left|000\right\rangle\right) + \left(I \otimes \mathrm{CNOT}_{2,3}\right)\left(\frac{1}{\sqrt{3}} \left|100\right\rangle\right) + \left(I \otimes \mathrm{CNOT}_{2,3}\right)\left(\frac{1}{\sqrt{3}} \left|110\right\rangle\right) = \\
&= \frac{1}{\sqrt{3}} \left|000\right\rangle + \frac{1}{\sqrt{3}} \left|100\right\rangle + \frac{1}{\sqrt{3}} \left|111\right\rangle
\end{aligned}
$$

Applying a CNOT gate with first qubit as control and second qubit as target:

$$
\begin{aligned}
\left|\Psi_4\right\rangle &= \left(\mathrm{CNOT}_{1,2} \otimes I\right) \left|\Psi_3\right\rangle = \left(\mathrm{CNOT}_{1,2} \otimes I\right)\left(\frac{1}{\sqrt{3}} \left|000\right\rangle + \frac{1}{\sqrt{3}} \left|100\right\rangle + \frac{1}{\sqrt{3}} \left|111\right\rangle\right) = \\
&= \left(\mathrm{CNOT}_{1,2} \otimes I\right)\left(\frac{1}{\sqrt{3}} \left|000\right\rangle\right) + \left(\mathrm{CNOT}_{1,2} \otimes I\right)\left(\frac{1}{\sqrt{3}} \left|100\right\rangle\right) + \left(\mathrm{CNOT}_{1,2} \otimes I\right)\left(\frac{1}{\sqrt{3}} \left|111\right\rangle\right) = \\
&= \frac{1}{\sqrt{3}} \left|000\right\rangle + \frac{1}{\sqrt{3}} \left|110\right\rangle + \frac{1}{\sqrt{3}} \left|101\right\rangle
\end{aligned}
$$

Applying an X gate on the first qubit:

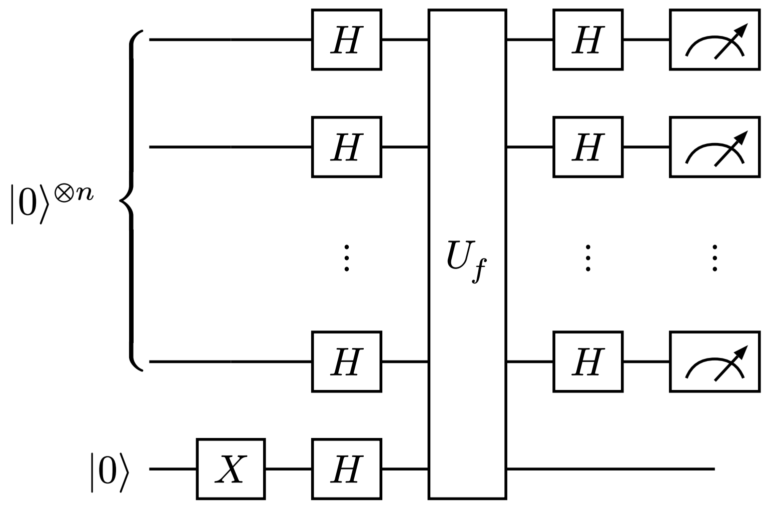
$$
\begin{aligned}
\left|\Psi_5\right\rangle &= (X \otimes I \otimes I) \left|\Psi_4\right\rangle = (X \otimes I \otimes I)\left(\frac{1}{\sqrt{3}} \left|000\right\rangle + \frac{1}{\sqrt{3}} \left|110\right\rangle + \frac{1}{\sqrt{3}} \left|101\right\rangle\right) = \\
&= (X \otimes I \otimes I)\left(\frac{1}{\sqrt{3}} \left|000\right\rangle\right) + (X \otimes I \otimes I)\left(\frac{1}{\sqrt{3}} \left|110\right\rangle\right) + (X \otimes I \otimes I)\left(\frac{1}{\sqrt{3}} \left|101\right\rangle\right) = \\
&= \frac{1}{\sqrt{3}} \left|100\right\rangle + \frac{1}{\sqrt{3}} \left|010\right\rangle + \frac{1}{\sqrt{3}} \left|001\right\rangle
\end{aligned}
$$

3.3. Deutsch-Josza Algorithm

Consider a function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ that has in input a binary string for length n and returns a binary value. This function in known in advance to be either *constant*, meaning that it has the same output for any input, or *balanced*, meaning that it returns each output an equal amount of times. In this case, f is constant if its output is always 0 or always 1 and is balanced if its output is 0 50% of the times and 1 50% of the times. The task is to determine in which of the two categories it falls.

A classical algorithm that solves this problem would necessarely resort to a "brute force" approach. In particular, in the most unfavorable case, a classical algorithm would need to test the function on half of all the 2^n possible strings, that is, 2^{n-1} function calls. This is because, assuming that the function has always returned the same output on all the previous trials, if the $2^{n-1} + 1$-th output is the same as before the function is necessarely constant, otherwise is balanced. This means that the computational complexity of the algorithm is $O(2^n)$.

It is possible to construct a quantum algorithm that solves this problem in $O(1)$ time. This algorithm is called **Deutsch-Josza Algorithm**, and is depicted as a quantum circuit as follows:



The circuit starts with $n+1$ qubits initialized in the state $|0\rangle$. The last extra qubit, called **ancillary qubit**, is necessary for the quantum computation to be reversible. The starting state of the entire system can be then written as $|000...0\rangle$.

The last qubit is changed from $|0\rangle$ to $|1\rangle$ applying an X gate, leaving the other qubits unchanged:

$$|\Psi_0\rangle = (I \otimes I \otimes ... \otimes I \otimes X)\ |000...0\rangle = (I\ |0\rangle) \otimes (I\ |0\rangle) \otimes ... \otimes (X\ |0\rangle) = |000...01\rangle$$

Then, an Hadamard gate is applied to all the qubits:

$$\begin{aligned} |\Psi_1\rangle &= (H \otimes H \otimes ... \otimes H)\ |\Psi_0\rangle = (H\ |0\rangle) \otimes (H\ |0\rangle) \otimes ... \otimes (H\ |1\rangle) = \\ &= \left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right) \otimes \left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right) \otimes ... \otimes \left(\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\right) = \\ &= \frac{1}{\sqrt{2^{n+1}}} \sum_{x \in \{0,1\}^n} |x\rangle(|0\rangle - |1\rangle) \end{aligned}$$

It is now possible to apply $U_f : |x, y\rangle \to |x, y \veebar f(x)\rangle$ to the state:

$$\begin{aligned} |\Psi_2\rangle &= U_f\ |\Psi_1\rangle = U_f \frac{1}{\sqrt{2^{n+1}}} \sum_{x \in \{0,1\}^n} |x\rangle(|0\rangle - |1\rangle) = \frac{1}{\sqrt{2^{n+1}}} \sum_{x \in \{0,1\}^n} U_f\ |x\rangle(|0\rangle - |1\rangle) = \\ &= \frac{1}{\sqrt{2^{n+1}}} \sum_{x \in \{0,1\}^n} U_f\ |x\rangle\ |0\rangle - U_f\ |x\rangle\ |1\rangle = \frac{1}{\sqrt{2^{n+1}}} \sum_{x \in \{0,1\}^n} |x\rangle\ |0 \veebar f(x)\rangle - |x\rangle\ |1 \veebar f(x)\rangle \end{aligned}$$

Consider $|x\rangle\ |0 \veebar f(x)\rangle - |x\rangle\ |1 \veebar f(x)\rangle$. Since 0 and 1 are single bits and opposite in value, $|0 \veebar f(x)\rangle$ and $|1 \veebar f(x)\rangle$ will always be one the negation of the other. It is therefore possible to rewrite the expression as $|x\rangle\ |f(x)\rangle - |x\rangle\ |\overline{f(x)}\rangle)$. The expression can be simplified even further by observing what happens when the function $f(x)$ is substituted explicitly with its possible outputs:

$$|x\rangle\ |f(x)\rangle - |x\rangle|\overline{f(x)}\rangle) = \begin{cases} |x\rangle\ |0\rangle - |x\rangle\ |\overline{0}\rangle = |x\rangle\ |0\rangle - |x\rangle\ |1\rangle = |x\rangle(|0\rangle - |1\rangle) \\ |x\rangle\ |1\rangle - |x\rangle\ |\overline{1}\rangle = |x\rangle\ |1\rangle - |x\rangle\ |0\rangle = |x\rangle(|1\rangle - |0\rangle) \end{cases} = (-1)^{f(x)}\ |x\rangle(|0\rangle - |1\rangle)$$

Substituting in the previous state gives:

$$|\Psi_2\rangle = \frac{1}{\sqrt{2^{n+1}}} \sum_{x \in \{0,1\}^n} (-1)^{f(x)}\ |x\rangle(|0\rangle - |1\rangle)$$

At this point, the ancillary bit is no longer necessary and can be ignored, and the following remains:

$$|\Psi_2\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} (-1)^{f(x)}\ |x\rangle$$

Applying the Hadamard gate (again) gives:

$$|\Psi_3\rangle = H\,|\Psi_2\rangle = H\frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}(-1)^{f(x)}\,|x\rangle = \frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}(-1)^{f(x)}H\,|x\rangle$$

Recall that $H\,|0\rangle = \frac{\sqrt{2}}{2}(|0\rangle + |1\rangle)$, whereas $H\,|1\rangle = \frac{\sqrt{2}}{2}(|0\rangle - |1\rangle)$. It is then possible to write the result of applying a Hadamard gate to a single unknown bit x_i as $H\,|x_i\rangle = \frac{\sqrt{2}}{2}(|0\rangle + (-1)^{x_i}\,|1\rangle)$. This result can be generalized:

$$\begin{aligned} H\,|x_1x_2...x_n\rangle &= H\,|x_1\rangle \otimes H\,|x_2\rangle \otimes ... \otimes H\,|x_n\rangle = \\ &= \frac{1}{\sqrt{2}}(|0\rangle + (-1)^{x_1}\,|1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + (-1)^{x_2}\,|1\rangle) \otimes ... \otimes \frac{1}{\sqrt{2}}(|0\rangle + (-1)^{x_n}\,|1\rangle) = \\ &= \frac{1}{\sqrt{2^n}}\sum_{j\in\{0,1\}^n}(-1)^{\langle x,j\rangle}\,|j\rangle \end{aligned}$$

Where $\langle\rangle$ denotes the inner product. Substituting it back:

$$\begin{aligned} |\Psi_3\rangle &= \frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}(-1)^{f(x)}\frac{1}{\sqrt{2^n}}\sum_{j\in\{0,1\}^n}(-1)^{\langle x,j\rangle}\,|j\rangle = \frac{1}{2^n}\sum_{x\in\{0,1\}^n}(-1)^{f(x)}\sum_{j\in\{0,1\}^n}(-1)^{\langle x,j\rangle}\,|j\rangle = \\ &= \frac{1}{2^n}\sum_{x\in\{0,1\}^n}\sum_{j\in\{0,1\}^n}(-1)^{f(x)+\langle x,j\rangle}\,|j\rangle \end{aligned}$$

Applying a measurement process with respect to the state $|000...0\rangle$:

$$\frac{1}{2^n}\sum_{x\in\{0,1\}^n}\sum_{j\in\{0\}^n}(-1)^{f(x)+\langle x,0\rangle}\,|0\rangle = \frac{1}{2^n}\sum_{x\in\{0,1\}^n}(-1)^{f(x)+x\cdot 0}\,|0\rangle = \frac{1}{2^n}\sum_{x\in\{0,1\}^n}(-1)^{f(x)}\,|0\rangle$$

The probability of obtaining this state is therefore:

$$P_0 = \left|\frac{1}{2^n}\sum_{x\in\{0,1\}^n}(-1)^{f(x)}\right|^2$$

Now consider the case in which $f(x)$ is constant, meaning that $f(x) = 0$ for all x or $f(x) = 1$ for all x. Consider the first case:

$$\left|\frac{1}{2^n}\sum_{x\in\{0\}^n}(-1)^0\right|^2 = \left|\frac{1}{2^n}\sum_{x\in\{0\}^n}1\right|^2 = \left|\frac{1}{\cancel{2^n}}\cancel{2^n}\right|^2 = |1|^2 = 1^2 = 1$$

As for the second case:

$$\left|\frac{1}{2^n}\sum_{x\in\{1\}^n}(-1)^1\right|^2 = \left|\frac{1}{2^n} - \sum_{x\in\{1\}^n}1\right|^2 = \left|\frac{-1}{\cancel{2^n}}\cancel{2^n}\right|^2 = |-1|^2 = 1^2 = 1$$

This means that the probability of observing the state $|000...0\rangle$ is 1 (complete certainty) when the function is constant.

Consider instead the case in which $f(x)$ is balanced. Since $f(x) = 0$ on one half of the inputs and $f(x) = 1$ on the other half of the inputs, the sum cancels:

$$\left|\frac{1}{2^n}\sum_{x\in\{0,1\}^n}(-1)^{f(x)}\right|^2 = \left|\frac{1}{2^n}((-1)^0 + (-1)^0 + ... + (-1)^1 + (-1)^1)\right|^2 = \left|\frac{1}{2^n}0\right|^2 = |0|^2 = 0^2 = 0$$

This means that the probability of observing the state $|000...0\rangle$ is 0 (complete impossibility) when the function is balanced.

The algorithm then solves the problem in $O(1)$ time, because the function and each gate is invoked exactly once. Note how this problem is, from a practical standpoint, useless, since there are no real-world applications for solving it. Nevertheless, it is an instructive example on how a quantum computer would solve a problem exponentially faster than any classical computer could.

3.4. Bernstein-Vazirani Algorithm

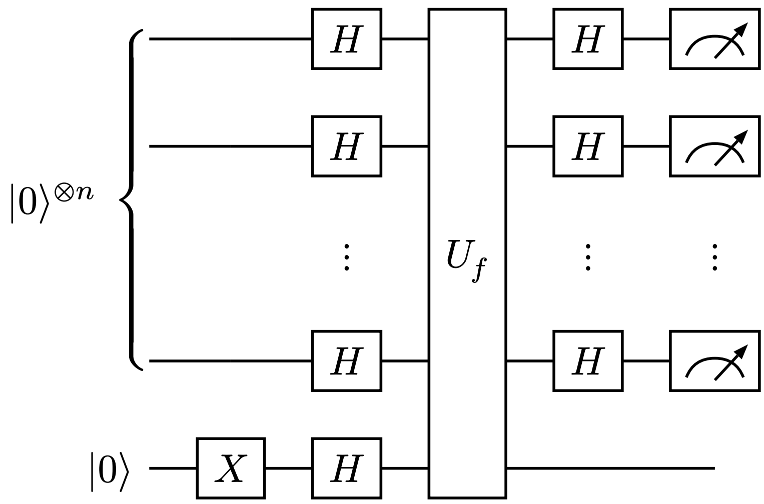
Suppose one is given a binary function $f : \{0,1\}^n \to \{0,1\}$ defined as $f(x) = \langle s, x \rangle$, where s is an unknown n-bit binary string and $\langle\rangle$ denotes the dot product. Assuming to know the value of $f(x)$ for all n-bit binary strings x, the task is to find s.

The most efficient way to solve the problem in a classical framework would be to pick n strings out of the posssible 2^n strings and evaluate the function on those strings. This leads to a system of linear equations having n unknowns. Of course, the most reasonable choice of strings are those having 0 in each position but one:

$$\begin{cases} f(1000...0) = \langle s, 1000...0 \rangle = s_1 \cdot 1 + s_2 \cdot 0 + ... + s_n \cdot 0 = s_1 \\ f(0100...0) = \langle s, 0100...0 \rangle = s_1 \cdot 0 + s_2 \cdot 1 + ... + s_n \cdot 0 = s_2 \\ \vdots \\ f(0000...1) = \langle s, 0000...1 \rangle = s_1 \cdot 0 + s_2 \cdot 0 + ... + s_n \cdot 1 = s_n \end{cases}$$

This means that the computational complexity of a classical algorithm for this problem is $O(n)$.

A quantum algorithm known as **Bernstein-Vazirani algorithm**, whose quantum circuit is presented below, can solve the problem faster. Notice how the circuit is identical to the one presented for the Deutsch-Josza Algorithm, except for the function encoded in the U_f gate:



The algorithm starts with $n+1$ qubits in state $|0\rangle$, where each of the first n qubits represents one of the bits of the string and the last qubit is an ancillary qubit. The last qubit is changed from $|0\rangle$ to $|1\rangle$, leaving the other qubits unchanged:

$$|\Psi_0\rangle = (I \otimes I \otimes ... \otimes I \otimes X)\ |000...0\rangle = (I\ |0\rangle) \otimes (I\ |0\rangle) \otimes ... \otimes (X\ |0\rangle) = |000...01\rangle$$

A Hadamard gate is the applied to all the qubits:

$$\begin{aligned} |\Psi_1\rangle &= (H \otimes H \otimes ... \otimes H \otimes H)\ |\Psi_0\rangle = (H\ |0\rangle) \otimes (H\ |0\rangle) \otimes ... \otimes (H\ |1\rangle) = \\ &= \left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right) \otimes \left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right) \otimes ... \otimes \left(\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\right) = \\ &= \frac{1}{\sqrt{2^{n+1}}} \sum_{x \in \{0,1\}^n} |x\rangle(|0\rangle - |1\rangle) \end{aligned}$$

It is now possible to apply $U_f : |x, y\rangle \to |x, y \veebar f(x)\rangle$ to the state:

$$|\Psi_2\rangle = U_f\ |\Psi_1\rangle = U_f \frac{1}{\sqrt{2^{n+1}}} \sum_{x\in\{0,1\}^n} |x\rangle(|0\rangle - |1\rangle) = \frac{1}{\sqrt{2^{n+1}}} \sum_{x\in\{0,1\}^n} U_f\ |x\rangle(|0\rangle - |1\rangle) =$$

$$= \frac{1}{\sqrt{2^{n+1}}} \sum_{x\in\{0,1\}^n} U_f\ |x\rangle\ |0\rangle - U_f\ |x\rangle\ |1\rangle = \frac{1}{\sqrt{2^{n+1}}} \sum_{x\in\{0,1\}^n} |x\rangle\ |0 \veebar f(x)\rangle - |x\rangle\ |1 \veebar f(x)\rangle =$$

$$= \frac{1}{\sqrt{2^{n+1}}} \sum_{x\in\{0,1\}^n} (-1)^{f(x)}\ |x\rangle(|0\rangle - |1\rangle) = \frac{1}{\sqrt{2^{n+1}}} \sum_{x\in\{0,1\}^n} (-1)^{\langle x,s\rangle}\ |x\rangle(|0\rangle - |1\rangle)$$

Discarding the last qubit and applying Hadamard again:

$$|\Psi_3\rangle = H\ |\Psi_2\rangle = H \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} (-1)^{\langle x,s\rangle}\ |x\rangle = \frac{1}{2^n} \sum_{x\in\{0,1\}^n} \sum_{j\in\{0,1\}^n} (-1)^{\langle x,s\rangle \veebar \langle x,j\rangle}\ |j\rangle =$$

$$= \frac{1}{2^n} \sum_{x\in\{0,1\}^n} \sum_{j\in\{0,1\}^n} (-1)^{\langle x,(s\veebar j)\rangle}\ |j\rangle$$

The amplitude of the state $|s\rangle$ is:

$$\frac{1}{2^n} \sum_{x\in\{0,1\}^n} (-1)^{\langle x,(s\veebar s)\rangle} = \frac{1}{2^n} \sum_{x\in\{0,1\}^n} (-1)^{\langle x,000\ldots0\rangle} = \frac{1}{2^n} \sum_{x\in\{0,1\}^n} (-1)^0 = \frac{1}{2^n} \sum_{x\in\{0,1\}^n} 1 = \frac{1}{\cancel{2^n}}\cancel{2^n} = 1$$

Which means that, when measuring with respect to the standard basis, the state $|s\rangle$ will be obtained with certainty. Since the algorithm has performed a single function call to solve the problem, its time bound is $O(1)$.

3.5. Grover Algorithm

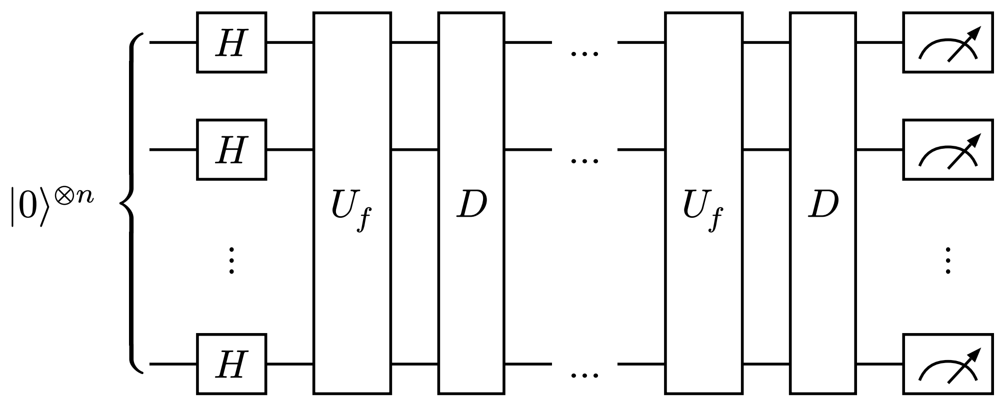
Suppose one is given 2^n objects, each uniquely identified by a binary string of n bits. Suppose that there's interest in finding, out of all these 2^n objects, a specific object identified by the string x_0. This problem could model, for example, the search of an element in a database whose IDs are unsorted.

The task of inspecting a generic string x to determine if it is equal to x_0 could be formalized mathematically by a function such as:

$$f(x) = \begin{cases} 1 \text{ if } x = x_0 \\ 0 \text{ if } x \neq x_0 \end{cases}$$

Solving this problem in a classical computation framework would simply entail applying this function to all of the possible 2^n strings that can be constructed with n bits. Therefore, the classical time bound for the problem is $O(2^n)$.

A quantum algorithm known as **Grover algorithm**, whose quantum circuit is presented below, can solve the problem faster:



The algorithm starts with n qubits (each representing one of the bits of the string) prepared in the $|0\rangle$ state. Then, each qubit undergoes a Hadamard transformation; for the sake of clarity, the resulting state is denoted as $|\eta\rangle$:

$$|\Psi_0\rangle = (H \otimes H \otimes ... \otimes H)(|0\rangle \otimes |0\rangle \otimes ... \otimes |0\rangle) = \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} |x\rangle = |\eta\rangle$$

Then, the quantum oracle that encodes f is applied. Given a generic n-qubit state $|x\rangle$, the oracle can be defined as follows:

$$U_f\ |x\rangle = \begin{cases} -|x\rangle & \text{if } x = x_0 \text{ or equivalently if } f(x) = 1 \\ |x\rangle & \text{if } x \neq x_0 \text{ or equivalently if } f(x) = 0 \end{cases} = (-1)^{f(x)}\ |x\rangle$$

Applying the oracle gives:

$$\begin{aligned} |\Psi_1\rangle &= U_f\ |\Psi_0\rangle = U_f\ |\eta\rangle = U_f \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} |x\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} U_f\ |x\rangle = \\ &= \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} (-1)^{f(x)}\ |x\rangle = |\eta\rangle - \frac{2}{\sqrt{2^n}}\ |x_0\rangle \end{aligned}$$

The resulting state is a superposition of all strings, each having the same amplitude but the $|x_0\rangle$ state had its sign flipped. Note that, at this point, nothing much has changed: if one were to sample the current state, the probability of finding any state is exactly the same, because the modulus square of a negative amplitude is still positive.

Next, the **diffusion operator** $D = 2\ |\eta\rangle\langle\eta| - I$ is applied, giving:

$$\begin{aligned} |\Psi_2\rangle &= D\ |\Psi_1\rangle = (2\ |\eta\rangle\langle\eta| - I)\ |\Psi_1\rangle = (2\ |\eta\rangle\langle\eta| - I)\left(|\eta\rangle - \frac{2}{\sqrt{2^n}}\ |x_0\rangle\right) = \\ &= 2\ |\eta\rangle\langle\eta|\eta\rangle - \frac{4}{\sqrt{2^n}}\ |\eta\rangle\langle\eta|x_0\rangle - I\ |\eta\rangle + I\frac{2}{\sqrt{2^n}}\ |x_0\rangle = |\eta\rangle - \frac{4}{\sqrt{2^n}}\ |\eta\rangle\langle\eta|x_0\rangle + \frac{2}{\sqrt{2^n}}\ |x_0\rangle \end{aligned}$$

Note that, since all base states are orthonormal:

$$\langle\eta|x_0\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} \langle x|x_0\rangle = \frac{1}{\sqrt{2^n}}(0 + 0 + ... + 1 + ... + 0) = \frac{1}{\sqrt{2^n}}$$

Where the single 1 is the contribution given by $|x_0\rangle$ itself. Substituting the expression in the previous one gives:

$$|\Psi_2\rangle = |\eta\rangle - \frac{4}{\sqrt{2^n}}\ |\eta\rangle\left(\frac{1}{\sqrt{2^n}}\right) + \frac{2}{\sqrt{2^n}}\ |x_0\rangle = |\eta\rangle - \frac{4}{2^n}\ |\eta\rangle + \frac{2}{\sqrt{2^n}}\ |x_0\rangle = \left(1 - \frac{4}{2^n}\right)\ |\eta\rangle + \frac{2}{\sqrt{2^n}}\ |x_0\rangle$$

Explicitly expanding $|\eta\rangle$ can give a clearer understanding of the result:

$$\begin{aligned} |\Psi_2\rangle &= \left(1 - \frac{4}{2^n}\right)\ |\eta\rangle + \frac{2}{\sqrt{2^n}}\ |x_0\rangle = \left(1 - \frac{4}{2^n}\right)\left(\frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} |x\rangle\right) + \frac{2}{\sqrt{2^n}}\ |x_0\rangle = \\ &= \left(1 - \frac{4}{2^n}\right)\left(\frac{1}{\sqrt{2^n}}\left(|x_0\rangle + \sum_{x \in \{0,1\}^n - \{x_0\}} |x\rangle\right)\right) + \frac{2}{\sqrt{2^n}}\ |x_0\rangle = \\ &= \left(1 - \frac{4}{2^n}\right)\left(\frac{1}{\sqrt{2^n}}\ |x_0\rangle + \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n - \{x_0\}} |x\rangle\right) + \frac{2}{\sqrt{2^n}}\ |x_0\rangle = \\ &= \frac{1}{\sqrt{2^n}}\ |x_0\rangle + \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n - \{x_0\}} |x\rangle - \frac{4}{2^n}\frac{1}{\sqrt{2^n}}\ |x_0\rangle - \frac{4}{2^n}\frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n - \{x_0\}} |x\rangle + \frac{2}{\sqrt{2^n}}\ |x_0\rangle = \\ &= \left(\frac{1}{\sqrt{2^n}} - \frac{4}{2^n}\frac{1}{\sqrt{2^n}}\right) \sum_{x \in \{0,1\}^n - \{x_0\}} |x\rangle + \left(\frac{1}{\sqrt{2^n}} - \frac{4}{2^n}\frac{1}{\sqrt{2^n}} + \frac{2}{\sqrt{2^n}}\right)\ |x_0\rangle = \\ &= \frac{2^n - 4}{2^n\sqrt{2^n}} \sum_{x \in \{0,1\}^n - \{x_0\}} |x\rangle + \frac{3 \cdot 2^n - 4}{2^n\sqrt{2^n}}\ |x_0\rangle \approx \frac{2^n - 4}{2^n\sqrt{2^n}} \sum_{x \in \{0,1\}^n - \{x_0\}} |x\rangle + 3\left(\frac{2^n - 4}{2^n\sqrt{2^n}}\right)\ |x_0\rangle \end{aligned}$$

Which means that the amplitude of the state $|x_0\rangle$ is roughly three times the amplitude of all the other states. This means that now, if the state is sampled, there is an increased probability of obtaining $|x_0\rangle$ than of obtaining any other state, even though this is not necessarely certain.

The idea is to apply repeatedly the unitary matrix and the diffusion operator so that the probability of obtaining $|x_0\rangle$ becomes arbitrarely large. With enough *amplitude amplifications*, the chance of obtaining $|x_0\rangle$ is almost certain. For example, to have a 90% probability of obtaining $|x_0\rangle$, roughly $\lceil \frac{\pi}{4}\sqrt{2^n} \rceil$ amplitude amplifications are necessary.

This means that the advantage over classical computation given by the Grover Algorithm is of a quadratic factor, giving a time bound of $O\left(\sqrt{2^n}\right)$. It has also been proven that no algorithm, whether classical or quantum, can achieve a time bound lower than this.

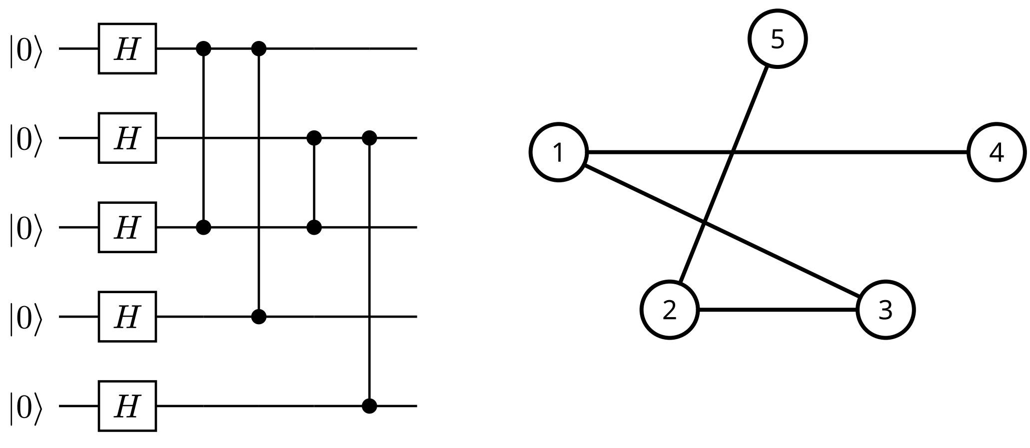
3.6. Representing graphs with quantum circuits

An interesting use case for quantum circuits is the encoding of graphs. Consider an undirected graph $G = (V, E)$, with $|V|$ nodes and $|E|$ edges. A quantum computer that encodes the graph has all qubits prepared in the state $|0\rangle$: the i-th qubit represents the i-th node of the graph (given an ordering). The circuit has one Hadamard gate applied to each qubit and has as many CZ gates as the number of edges. If the graph has an edge (i, j), then there is a CZ gate with controls in the i-th and j-th qubit.

Exercise 3.6.1: How would an undirected graph such as this be encoded into a quantum circuit?

$$G = (V, E) = (\{1, 2, 3, 4, 5\}, \{(1, 3), (1, 4), (2, 3), (2, 5)\})$$

Solution:



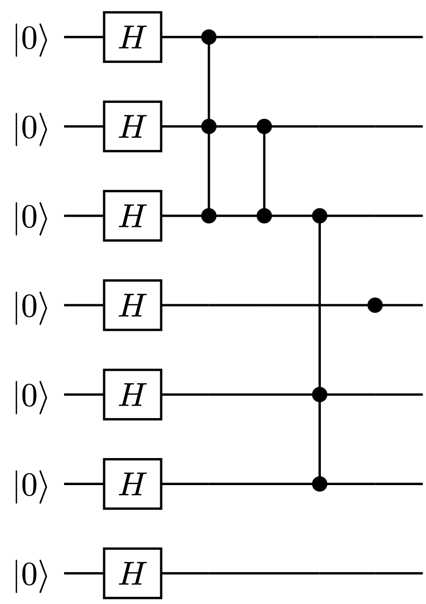
□

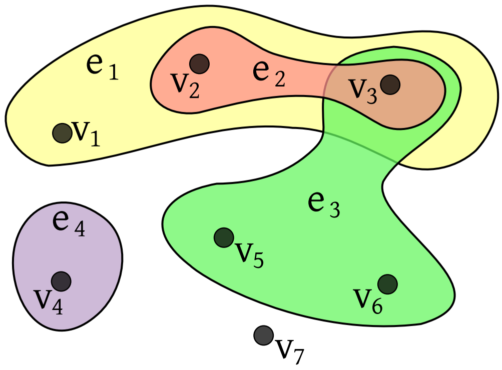
In particular, quantum computers can efficiently encode **hypergraphs**, graphs where an edge can connect more than one node at the same time. An hypergraph is encoded in the exact same way: the CZ gates can have more than one control, one for each node that constitutes the edge.

Exercise 3.6.2: How would an undirected hypergraph such as this be encoded into a quantum circuit?

$$G = (V, E) = (\{1, 2, 3, 4, 5, 6, 7\}, \{(1, 2, 3), (2, 3), (3, 5, 6), (4)\})$$

Solution:



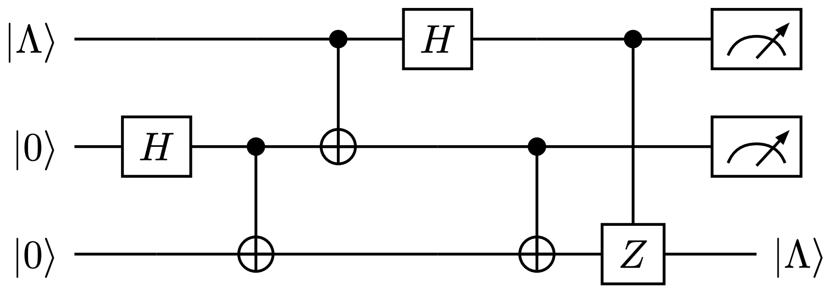


[5]

□

3.7. Quantum teleportation

The following circuit implements **quantum teleportation** of a single qubit in an unknown state $|\Lambda\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$:



Applying an Hadamard gate to the second qubit:

$$\begin{aligned}|\Psi_1\rangle &= (I \otimes H \otimes I)((\alpha\,|0\rangle + \beta\,|1\rangle) \otimes |0\rangle \otimes |0\rangle) = (I(\alpha\,|0\rangle + \beta\,|1\rangle)) \otimes (H\,|0\rangle) \otimes (I\,|0\rangle) = \\ &= (\alpha\,|0\rangle + \beta\,|1\rangle) \otimes \frac{\sqrt{2}}{2}(|0\rangle + |1\rangle) \otimes |0\rangle = \frac{\sqrt{2}}{2}(\alpha\,|0\rangle + \beta\,|1\rangle) \otimes (|00\rangle + |10\rangle) = \\ &= \frac{\sqrt{2}}{2}(\alpha\,|000\rangle + \alpha\,|010\rangle + \beta\,|100\rangle + \beta\,|110\rangle)\end{aligned}$$

Applying a CX (CNOT) gate to the second and third qubit:

$$\begin{aligned}|\Psi_2\rangle &= \left(I \otimes \mathrm{CX}_{2,3}\right)|\Psi_1\rangle = \left(I \otimes \mathrm{CX}_{2,3}\right)\frac{\sqrt{2}}{2}(\alpha\,|000\rangle + \alpha\,|010\rangle + \beta\,|100\rangle + \beta\,|110\rangle) = \\ &= \frac{\sqrt{2}}{2}\left(\left(I \otimes \mathrm{CX}_{2,3}\right)\alpha\,|000\rangle + \left(I \otimes \mathrm{CX}_{2,3}\right)\alpha\,|010\rangle + \left(I \otimes \mathrm{CX}_{2,3}\right)\beta\,|100\rangle + \left(I \otimes \mathrm{CX}_{2,3}\right)\beta\,|110\rangle\right) = \\ &= \frac{\sqrt{2}}{2}(\alpha\,|000\rangle + \alpha\,|011\rangle + \beta\,|100\rangle + \beta\,|111\rangle)\end{aligned}$$

Applying a CX gate to the first and second qubit:

[5]By Hypergraph.svg: Kilom691derivative work: Pgdx (talk) - Hypergraph.svg, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=10687664

$$|\Psi_3\rangle = \left(I \otimes \mathrm{CX}_{1,2}\right) |\Psi_2\rangle = \left(I \otimes \mathrm{CX}_{1,2}\right)\frac{\sqrt{2}}{2}(\alpha\ |000\rangle + \alpha\ |011\rangle + \beta\ |100\rangle + \beta\ |111\rangle) =$$
$$= \frac{\sqrt{2}}{2}(\left(I \otimes \mathrm{CX}_{1,2}\right)\alpha\ |000\rangle + \left(I \otimes \mathrm{CX}_{1,2}\right)\alpha\ |011\rangle + \left(I \otimes \mathrm{CX}_{1,2}\right)\beta\ |100\rangle + \left(I \otimes \mathrm{CX}_{1,2}\right)\beta\ |111\rangle) =$$
$$= \frac{\sqrt{2}}{2}(\alpha\ |000\rangle + \alpha\ |011\rangle + \beta\ |110\rangle + \beta\ |101\rangle)$$

Applying an Hadamard gate to the first qubit:

$$|\Psi_4\rangle = (H \otimes I \otimes I)\ |\Psi_3\rangle = (H \otimes I \otimes I)\frac{\sqrt{2}}{2}(\alpha\ |000\rangle + \alpha\ |011\rangle + \beta\ |110\rangle + \beta\ |101\rangle) =$$
$$= \frac{\sqrt{2}}{2}((H \otimes I \otimes I)\alpha\ |000\rangle + (H \otimes I \otimes I)\alpha\ |011\rangle + (H \otimes I \otimes I)\beta\ |110\rangle + (H \otimes I \otimes I)\beta\ |101\rangle) =$$
$$= \frac{\sqrt{2}}{2}\left(\frac{\sqrt{2}}{2}(\alpha(|0\rangle + |1\rangle) \otimes |00\rangle + \alpha(|0\rangle + |1\rangle) \otimes |11\rangle + \beta(|0\rangle - |1\rangle) \otimes |10\rangle + \beta(|0\rangle - |1\rangle) \otimes |01\rangle)\right) =$$
$$= \frac{1}{2}(\alpha\ |000\rangle + \alpha\ |100\rangle + \alpha\ |011\rangle + \alpha\ |111\rangle + \beta\ |010\rangle - \beta\ |110\rangle + \beta\ |001\rangle - \beta\ |101\rangle)$$

Applying a CX (CNOT) gate to the second and third qubit:

$$|\Psi_5\rangle = \left(I \otimes \mathrm{CX}_{2,3}\right) |\Psi_4\rangle =$$
$$= \left(I \otimes \mathrm{CX}_{2,3}\right)\frac{1}{2}(\alpha\ |000\rangle + \alpha\ |100\rangle + \alpha\ |011\rangle + \alpha\ |111\rangle + \beta\ |010\rangle - \beta\ |110\rangle + \beta\ |001\rangle - \beta\ |101\rangle) =$$
$$= \frac{1}{2}(\alpha\ |000\rangle + \alpha\ |100\rangle + \alpha\ |010\rangle + \alpha\ |110\rangle + \beta\ |011\rangle - \beta\ |111\rangle + \beta\ |001\rangle - \beta\ |101\rangle)$$

Applying a CZ gate to the first and third qubit:

$$|\Psi_6\rangle = \left(I \otimes \mathrm{CZ}_{1,3}\right) |\Psi_5\rangle =$$
$$= \left(I \otimes \mathrm{CZ}_{1,3}\right)\frac{1}{2}(\alpha\ |000\rangle + \alpha\ |100\rangle + \alpha\ |010\rangle + \alpha\ |110\rangle + \beta\ |011\rangle - \beta\ |111\rangle + \beta\ |001\rangle - \beta\ |101\rangle) =$$
$$= \frac{1}{2}(\alpha\ |000\rangle + \alpha\ |100\rangle + \alpha\ |010\rangle + \alpha\ |110\rangle + \beta\ |011\rangle + \beta\ |111\rangle + \beta\ |001\rangle + \beta\ |101\rangle)$$
$$= \frac{1}{2}\alpha(|000\rangle + |100\rangle + |010\rangle + |110\rangle) + \frac{1}{2}\beta(|011\rangle + |111\rangle + |001\rangle + |101\rangle)$$

Measuring the first and second qubit can result in four outcomes: $|00\rangle, |01\rangle, |10\rangle, |11\rangle$. Their projectors are as follows:

$$\begin{bmatrix} P_{00} = |000\rangle\langle 000| + |001\rangle\langle 001| & P_{01} = |010\rangle\langle 010| + |011\rangle\langle 011| \\ P_{10} = |100\rangle\langle 100| + |101\rangle\langle 101| & P_{11} = |110\rangle\langle 110| + |111\rangle\langle 111| \end{bmatrix}$$

Applying these projectors to the final state $|\Psi_6\rangle$:

$$P_{00}\ |\Psi_6\rangle = (|000\rangle\langle 000| + |001\rangle\langle 001|)\Big(\frac{1}{2}\alpha(|000\rangle + |100\rangle + |010\rangle + |110\rangle) + \frac{1}{2}\beta(|011\rangle + |111\rangle + |001\rangle + |101\rangle)\Big) = \frac{1}{2}\alpha\ |000\rangle + \frac{1}{2}\beta\ |001\rangle$$

$$P_{01}\ |\Psi_6\rangle = (|010\rangle\langle 010| + |011\rangle\langle 011|)\Big(\frac{1}{2}\alpha(|000\rangle + |100\rangle + |010\rangle + |110\rangle) + \frac{1}{2}\beta(|011\rangle + |111\rangle + |001\rangle + |101\rangle)\Big) = \frac{1}{2}\alpha\ |010\rangle + \frac{1}{2}\beta\ |011\rangle$$

$$P_{10}\ |\Psi_6\rangle = (|100\rangle\langle 100| + |101\rangle\langle 101|)\Big(\frac{1}{2}\alpha(|000\rangle + |100\rangle + |010\rangle + |110\rangle) + \frac{1}{2}\beta(|011\rangle + |111\rangle + |001\rangle + |101\rangle)\Big) = \frac{1}{2}\alpha\ |100\rangle + \frac{1}{2}\beta\ |101\rangle$$

$$P_{11}\ |\Psi_6\rangle = (|110\rangle\langle 110| + |111\rangle\langle 111|)\Big(\frac{1}{2}\alpha(|000\rangle + |100\rangle + |010\rangle + |110\rangle) + \frac{1}{2}\beta(|011\rangle + |111\rangle + |001\rangle + |101\rangle)\Big) = \frac{1}{2}\alpha\ |110\rangle + \frac{1}{2}\beta\ |111\rangle$$

All of these (non-normalized) states have associated the same probability:

$$\langle\Psi_6|\ P_{00}\ |\Psi_6\rangle = \langle\Psi_6|\ P_{01}\ |\Psi_6\rangle = \langle\Psi_6|\ P_{10}\ |\Psi_6\rangle = \langle\Psi_6|\ P_{11}\ |\Psi_6\rangle = \Big(\frac{1}{2}\alpha(\langle 000| + \langle 100| + \langle 010| + \langle 110|) + \frac{1}{2}\beta(\langle 011| + \langle 111| + \langle 001| + \langle 101|)\Big)\Big(\frac{1}{2}\alpha\ |000\rangle + \frac{1}{2}\beta\ |001\rangle\Big) = \Big(\frac{1}{2}\alpha\Big)^2 + \Big(\frac{1}{2}\beta\Big)^2 = \frac{1}{4}\alpha^2 + \frac{1}{4}\beta^2 = \frac{1}{4}(\alpha^2 + \beta^2) = \frac{1}{4}\cdot 1 = \frac{1}{4}$$

The resulting four possible states are therefore:

$$|\lambda_1\rangle = \frac{P_{00}\ |\Psi_6\rangle}{\sqrt{\langle\Psi_6|\ P_{00}\ |\Psi_6\rangle}} = \sqrt{4}\Big(\frac{1}{2}\alpha\ |000\rangle + \frac{1}{2}\beta\ |001\rangle\Big) = \alpha\ |000\rangle + \beta\ |001\rangle = |00\rangle \otimes (\alpha\ |0\rangle + \beta\ |1\rangle)$$

$$|\lambda_2\rangle = \frac{P_{01}\ |\Psi_6\rangle}{\sqrt{\langle\Psi_6|\ P_{01}\ |\Psi_6\rangle}} = \sqrt{4}\Big(\frac{1}{2}\alpha\ |010\rangle + \frac{1}{2}\beta\ |011\rangle\Big) = \alpha\ |010\rangle + \beta\ |011\rangle = |01\rangle \otimes (\alpha\ |0\rangle + \beta\ |1\rangle)$$

$$|\lambda_3\rangle = \frac{P_{10}\ |\Psi_6\rangle}{\sqrt{\langle\Psi_6|\ P_{10}\ |\Psi_6\rangle}} = \sqrt{4}\Big(\frac{1}{2}\alpha\ |100\rangle + \frac{1}{2}\beta\ |101\rangle\Big) = \alpha\ |100\rangle + \beta\ |101\rangle = |10\rangle \otimes (\alpha\ |0\rangle + \beta\ |1\rangle)$$

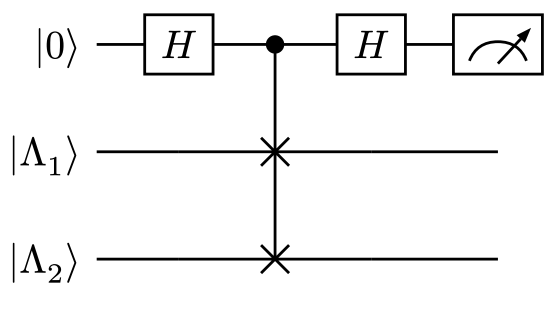
$$|\lambda_4\rangle = \frac{P_{11}\ |\Psi_6\rangle}{\sqrt{\langle\Psi_6|\ P_{11}\ |\Psi_6\rangle}} = \sqrt{4}\Big(\frac{1}{2}\alpha\ |110\rangle + \frac{1}{2}\beta\ |111\rangle\Big) = \alpha\ |110\rangle + \beta\ |111\rangle = |11\rangle \otimes (\alpha\ |0\rangle + \beta\ |1\rangle)$$

Any of these states is reached with the same probability, and in all four cases the third qubit is in the state $\alpha\ |0\rangle + \beta\ |1\rangle$, which is exactly the state $|\Lambda\rangle$ with which the circuit started. Notice how the starting state is now lost, since the first qubit had to be measured and therefore losing its superposition.

3.8. Swap test

Let Λ_1 and Λ_2 be two quantum states. Suppose that there's interest in knowing how much the two differ. Of course, this cannot be done by measuring, since measurement destroys the information contained in the two states. However, there is a way to know how much two states differ without knowing what these two states actually are.

This can be achieved through a quantum algorithm called **swap test**, whose circuit is presented below:



The algorithm starts with three qubits: an ancillary qubit prepared in the state $|0\rangle$ and two qubits prepared in the states $|\Lambda_1\rangle$ and $|\Lambda_2\rangle$ respectively. Applying an Hadamard gate to the first qubit:

$$|\Psi_0\rangle = (H \otimes I \otimes I)(|0\rangle \otimes |\Lambda_1\rangle \otimes |\Lambda_2\rangle) = H\ |0\rangle \otimes I\ |\Lambda_1\rangle \otimes I\ |\Lambda_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |\Lambda_1\rangle \otimes |\Lambda_2\rangle$$

Applying the controlled swap gate (first is control, second and third are the target):

$$\begin{aligned}
|\Psi_1\rangle = \mathrm{CSWAP}_{1,2,3}\ |\Psi_0\rangle &= \mathrm{CSWAP}_{1,2,3}\left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |\Lambda_1\rangle \otimes |\Lambda_2\rangle\right) = \\
&= \frac{1}{\sqrt{2}}\mathrm{CSWAP}_{1,2,3}(|0\rangle \otimes |\Lambda_1\rangle \otimes |\Lambda_2\rangle + |1\rangle \otimes |\Lambda_1\rangle \otimes |\Lambda_2\rangle) = \\
&= \frac{1}{\sqrt{2}}(|0\rangle \otimes |\Lambda_1\rangle \otimes |\Lambda_2\rangle + |1\rangle \otimes |\Lambda_2\rangle \otimes |\Lambda_1\rangle)
\end{aligned}$$

Applying an Hadamard gate to the first qubit once again:

$$\begin{aligned}
|\Psi_2\rangle &= H\ |\Psi_1\rangle = (H \otimes I \otimes I)\left(\frac{1}{\sqrt{2}}(|0\rangle \otimes |\Lambda_1\rangle \otimes |\Lambda_2\rangle + |1\rangle \otimes |\Lambda_2\rangle \otimes |\Lambda_1\rangle)\right) = \\
&= \frac{1}{\sqrt{2}}(H\ |0\rangle \otimes I\ |\Lambda_1\rangle \otimes I\ |\Lambda_2\rangle + H\ |1\rangle \otimes I\ |\Lambda_2\rangle \otimes I\ |\Lambda_1\rangle) = \\
&= \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |\Lambda_1\rangle \otimes |\Lambda_2\rangle + \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \otimes |\Lambda_2\rangle \otimes |\Lambda_1\rangle\right) = \\
&= \frac{1}{2}(|0\rangle \otimes |\Lambda_1\rangle \otimes |\Lambda_2\rangle + |1\rangle \otimes |\Lambda_1\rangle \otimes |\Lambda_2\rangle + |0\rangle \otimes |\Lambda_2\rangle \otimes |\Lambda_1\rangle - |1\rangle \otimes |\Lambda_2\rangle \otimes |\Lambda_1\rangle) = \\
&= \frac{1}{2}\ |0\rangle \otimes (|\Lambda_1\rangle \otimes |\Lambda_2\rangle + |\Lambda_2\rangle \otimes |\Lambda_1\rangle) + \frac{1}{2}\ |1\rangle \otimes (|\Lambda_1\rangle \otimes |\Lambda_2\rangle - |\Lambda_2\rangle \otimes |\Lambda_1\rangle) = \\
&= \frac{1}{2}\ |0\rangle \otimes (|\Lambda_1\Lambda_2\rangle + |\Lambda_2\Lambda_1\rangle) + \frac{1}{2}\ |1\rangle \otimes (|\Lambda_1\Lambda_2\rangle - |\Lambda_2\Lambda_1\rangle)
\end{aligned}$$

A measurement is performed on the first qubit. The associated projector is given by $P_0 = |0\rangle\langle 0| \otimes I \otimes I$. Applying it to the state gives:

$$\begin{aligned}
P_0\ |\Psi_2\rangle &= (|0\rangle\langle 0| \otimes I \otimes I)\left(\frac{1}{2}\ |0\rangle \otimes (|\Lambda_1\Lambda_2\rangle + |\Lambda_2\Lambda_1\rangle) + \frac{1}{2}\ |1\rangle \otimes (|\Lambda_1\Lambda_2\rangle - |\Lambda_2\Lambda_1\rangle)\right) = \\
&= \frac{1}{2}\ |0\rangle\langle 0|0\rangle \otimes (|\Lambda_1\Lambda_2\rangle + |\Lambda_2\Lambda_1\rangle) + \frac{1}{2}\ |0\rangle\langle 0|1\rangle \otimes (|\Lambda_1\Lambda_2\rangle - |\Lambda_2\Lambda_1\rangle) = \\
&= \frac{1}{2}\ |0\rangle \otimes (|\Lambda_1\Lambda_2\rangle + |\Lambda_2\Lambda_1\rangle)
\end{aligned}$$

Which means that the probability of measuring the first qubit to be in state $|0\rangle$ is:

$$\begin{aligned}
\langle\Psi_2|\ P_0\ |\Psi_2\rangle &= \left(\frac{1}{2}\langle 0| \otimes (\langle\Lambda_1\Lambda_2| + \langle\Lambda_2\Lambda_1|) + \frac{1}{2}\langle 1| \otimes (\langle\Lambda_1\Lambda_2| - \langle\Lambda_2\Lambda_1|)\right)\left(\frac{1}{2}\ |0\rangle \otimes (|\Lambda_1\Lambda_2\rangle + |\Lambda_2\Lambda_1\rangle)\right) = \\
&= \frac{1}{4}((\langle 0| \otimes (\langle\Lambda_1\Lambda_2| + \langle\Lambda_2\Lambda_1|) + \langle 1| \otimes (\langle\Lambda_1\Lambda_2| - \langle\Lambda_2\Lambda_1|))(|0\rangle \otimes (|\Lambda_1\Lambda_2\rangle + |\Lambda_2\Lambda_1\rangle))) \\
&= \frac{1}{4}((\langle 0| \otimes (\langle\Lambda_1\Lambda_2| + \langle\Lambda_2\Lambda_1|))(|0\rangle \otimes (|\Lambda_1\Lambda_2\rangle + |\Lambda_2\Lambda_1\rangle))) = \\
&= \frac{1}{4}((\langle 0|0\rangle \otimes (\langle\Lambda_1\Lambda_2| + \langle\Lambda_2\Lambda_1|)(|\Lambda_1\Lambda_2\rangle + |\Lambda_2\Lambda_1\rangle))) = \\
&= \frac{1}{4}(\langle\Lambda_1\Lambda_2|\Lambda_1\Lambda_2\rangle + \langle\Lambda_2\Lambda_1|\Lambda_1\Lambda_2\rangle + \langle\Lambda_1\Lambda_2|\Lambda_2\Lambda_1\rangle + \langle\Lambda_2\Lambda_1|\Lambda_2\Lambda_1\rangle) = \\
&= \frac{1}{4}(1 + (\langle\Lambda_2| \otimes \langle\Lambda_1|)(|\Lambda_1\rangle \otimes |\Lambda_2\rangle) + (\langle\Lambda_1| \otimes \langle\Lambda_2|)(|\Lambda_2\rangle \otimes |\Lambda_1\rangle) + 1) = \\
&= \frac{1}{4}(2 + \langle\Lambda_2|\Lambda_1\rangle\langle\Lambda_1|\Lambda_2\rangle + \langle\Lambda_1|\Lambda_2\rangle\langle\Lambda_2|\Lambda_1\rangle) = \\
&= \frac{1}{4}(2 + 2\langle\Lambda_2|\Lambda_1\rangle\langle\Lambda_1|\Lambda_2\rangle) = \frac{1}{2} + \frac{1}{2}(\langle\Lambda_1|\Lambda_2\rangle)^*\langle\Lambda_1|\Lambda_2\rangle = \frac{1}{2} + \frac{1}{2}|\langle\Lambda_1|\Lambda_2\rangle|^2
\end{aligned}$$

It is then necessary to run the circuit repeatedly on the exact same input state, or to have many identical copies of those states and run the circuit as many times in parallel. This way, the ancillary qubit will be found roughly $\frac{1}{2} + \frac{1}{2}|\langle\Lambda_1|\Lambda_2\rangle|^2$ times in state $|0\rangle$. When this sample mean is collected, it is possible to solve for $|\langle\Lambda_1|\Lambda_2\rangle|$:

$$\frac{1}{2} + \frac{1}{2}|\langle\Lambda_1|\Lambda_2\rangle|^2 \approx \frac{\text{N. of times } |0\rangle \text{ is found}}{\text{N. of observations}} \Rightarrow |\langle\Lambda_1|\Lambda_2\rangle| \approx \sqrt{2\left(\frac{\text{N. of times } |0\rangle \text{ is found}}{\text{N. of observations}}\right) - 1}$$

For example, if $|0\rangle$ is found roughly 50% of the time, then $|\langle\Lambda_1|\Lambda_2\rangle| = 0$, which means that the two states are orthogonal. If $|0\rangle$ is found almost 100% of the time, then $|\langle\Lambda_1|\Lambda_2\rangle| = 1$, which means that the two states are equal.

The precision of the test increases with the number of times the circuit is run. If the desired accuracy is ε, it can be achieved by running the circuit $O\left(\frac{1}{\varepsilon^2}\right)$ times.

4. Quantum Theory

4.1. Basics

Any quantum computer architecture that presents itself as usable must respect all of these criteria, called **DiVincenzo Criteria**[6]:

1. Possesses well isolated qubits, qubits shouldn't drift away;
2. Qubits must be initialized to a starting state that is fully under control;
3. Implements a universal set of operations;
4. Taking quantum decoherence into account: the operation time of quantum logic gates should be significantly less than the time frame in which qubits are stable;
5. There must be a way to sample the status of the qubit (readout);
6. Interconvention between qubits and flying qubits;
7. Existence of flying qubits;
8. Scalability: a technology that is not just theoretical but also usable in real applications.

Hardaware implementations of quantum computers include:

- **Superconducting quantum computing**;
- **Spin qubit quantum computing** (also known as **semiconductor quantum computing**);
- **Linear optical quantum computing** (also known as **photonic quantum computing**);
- **Trapped-ion quantum computing**;
- **Neutral atom quantum computing**.

Classical computers are based on the Von Neumann architecture, whereas quantum computers have many architecture model. Some architectures are better for some uses, whereas some are better for other uses. Most common ones are:

- **Gate model**, where gates are chained with each other in the same way as classical logic gates are combined into circuits. Supports criteria 1, 2, 4, 5;
- **Adiabatic**, arranging qubits and then applying thermodinamical processes. Supports criteria 1 and 4;
- **Measurement based**, virtualization of the gate model performing operation to condition each state. Support criteria 3;
- **Topological**, at the moment only theoretical.

Quantum noise is still problematic, but can be mitigated with **quantum error correction** introducing redundancy.

Similar to how the ISO-OSI model was formulated for classical computing, an analogous layered architecture was formulated for quantum computing. From top to bottom:

1. **Application layer**, where only algorithm live, hardware-independent;
2. **Representation layer**, where qubits are abstracted to logical qubits, hardware-independent;
3. **Quantum error correction layer**, to introduce redundancy;
4. **Virtual layer**, exploiting physical properties so that qubits are stable;
5. **Physical layer**, raw atoms and molecules.

Even though it is possible to consider qubits as the single atoms or molecules, a more reasonable approach is to go up one level of abstraction an talk about **logical qubits**, that also comprehend rendundancy qubits for error correction.

Quantum systems are different from classical systems. The evolution of a classical system can be completely determined from its starting conditions, that is, a classical system is **deterministic**. Observing a classical system at a certain time and predicting the state in which the system will find itself at that same time are, as a matter of fact, indistinguishable.

Quantum systems are not entirely deterministic. When a quantum system is not observed, it evolves in a deterministic way (according to, say, the Schroedinger equation), but when it is observed the result is only partially predictable. This is because, when observed, the system must be found in any of the possible states it can be,

[6] Only the first five criteria are present in the original formulation; the remaining three were introduced later.

but until then it could be in any of those. The probability of finding the system in a certain state depends on the initial conditions.

Quantum mechanics rests on six postulates:

1. **Superposition principle**. At any given time t_0, the state of a physical system $|\Phi(t_0)\rangle$ is described by specifying the vector `ket` as an appropriately normalized element of an Hilber space H, also called **state space**.
2. **Observable quantities**. Energy, angular momentum, position, ecc… are not described by functions. Instead, they are described by operators that act on elements of H. The matrix representation of operators is required to be Hermitian (square and has real eigenvalues). Operators, in general, do not commute, therefore the order of application matters.
3. **Spectrum of measurements**. Every possible value of an observable quantity is quantized, and it is an eigenvalue of the (matrix representation of the) operator associated to such observable.
4. **Probabilistic measurement for a non-degenerate discrete spectra of an operator**. Each possible eigenvalue has a probability to be sampled. Measuring an operator A over state $|\Psi(t_0)\rangle$ has a probability of obtaining the value a_i equal to $P(a_i)$, that goes with $|\langle\mu_i, a_i\rangle|^2$, with μ_i being the eigenvector associated to a_i. The vector $|\mu_i\rangle$ is given by an operator called **projection**, that extracts a component of a vector:
5. **Irreversibility of measurements**. The measurement of an observable A on the state $|\Psi\rangle$, equivalent to applying said operator to $|\Psi\rangle$, after the measurement process the new state is given by:

$$\frac{P_i \; |\Psi\rangle}{\sqrt{\langle | \; P_i \; | \; \Psi\rangle}}$$

   Which means that measuring a state influences the system giving a new system, states are not reversible.
6. **Time evolution**. The evolution in time of the states $|\Psi\rangle$ are governed by the **Schroedinger equation**:

$$i\frac{h}{2\pi}\frac{d}{dt} \; |\Psi(t)\rangle = H(t) \; |\Psi(t)\rangle$$

   Where $H(t)$ is the **Hermitian operator**, an operator associated to the energy of the system.

Postulates 5 and 6 seem to be contradictory, but they are not. Until a measurement is performed, a state is governed smoothly by and equation, whereas when a measurement happens the state is influenced.

4.2. Complexity

A **Turing machine** is a fundamental theoretical model of computation. It can be informally conceived as a moving head with an internal state that can move along a tape of infinite lenght, divided into cells. The machine can perform one operation at a time, reading the symbol on the current cell, replacing it with another symbol (or with the symbol itself) and moving one cell to the left or to the right.

A Turing machine M is formally defined as the tuple:

$$M = (Q, A, b \in A, \Sigma = A \cup \{L, R\}, \delta : Q \times A \to Q \times \Sigma, q_0 \in Q, F \subseteq Q)$$

Where:

- Q is the finite control set of states;
- A is the alphabet of the tape (the symbols that can be written on it);
- b is a special symbol called *blank*;
- Σ is the symbol output alphabet;
- δ is a function that, given a state and a tape symbol, outputs a state and an output symbol;
- q_0 is a special state, called *starting state*;
- F are special states, called *final states*;

Each Turing machine can be encoded into a binary string. That is, to each tuple as defined above is possible to associate a binary string that is able to represent the machine, without any loss of information. For a Turing machine M, its binary encoding is denoted as $\langle M\rangle$.

Any string S can be expressed in different languages. The most generic way to express S is as $\langle M \rangle w$, where w is an input string and $\langle M \rangle$ is a Turing machine that accepts w as input and has S as output. This equivalent description of S with respect to $\langle M \rangle$ and w is $d(S)$.

The lenght of $d(S)$ is denoted as $l(s)$. Note that both $\langle M \rangle$ and w are not unique, therefore there are countably infinitely many combinations of Turing machines and inputs outputting S. A Turing machine-input combination constitutes a **program**: $P = \langle M \rangle$

Being countable, there must be (at least) one program that is *minimal*, that is, constituted by the smallest number of characters. The length of one of those minimal programs is called **Kolmogorov complexity** of the string S, denoted as $K(S)$:

$$K(S) = \min\{l(P) \mid M(P) = S\}$$

The Kolmogorov complexity of a string can be conceived as the minimum number of characters necessary to encode a string in the most generic language possible.

The Turing machine here described is, to be more precise, a **deterministic Turing machine**, because the transition relation is a function: each time the head reads a symbol on the tape, it performs a single action. It is also possible to construct a **non deterministic Turing machine**, where the transition relation is not a function: each time the head reads a symbol on the tape, it performs one or more actions. Of course, it is not possible to construct a non deterministic Turing machine in practice, but it is still possible to employ it as a theoretical model.

Other Turing machines extensions include **probabilistic Turing machines** and **bounded probabilistic Turing machines**

Computational complexity is defined by a language and a machine capable of recognizing the language. In this context, a *machine* is any classical or quantum device that executes a single algorithm of which it is possible to compute the number of steps needed to complete its operation (**time complexity**) or the number of bits needed to store information (**space complexity**). A *language* is simply any sets of strings on an alphabet. A machine *recognises* a language if it is able to stop in a finite number of steps with an affermative answer for all strings in the language.

A set of languages recognised by a particular kind of machine within given resource bounds in terms of transition relation is called **complexity class**. For each algorithm it is possible to have a complexity class with respect to time and to space; the two might not be the same.

Note that, while Kolmogorov complexity is incomputable, complexity class is not. That is, there is an algorithm that, given in input another algorithm, is capable of (always) determining its complexity class, whereis there is no algorithm that, given in input a string, is capable of (always) determining its Kolmogorov complexity.

All previously stated computation models are still based on classical computations. A computational model for quantum computation is given by the **quantum Turing machine**:

$$M = \left(H_Q, H_A, b \in H_A, \Sigma = H_A \cup \{L, R\}, \delta : H_Q \rightarrow H_Q, q_0 \in H_Q, F \subseteq H_Q\right)$$

Where:

- H_Q is an Hilbert space containing the states;
- H_A is an Hilbert space containing the alphabet of the tape;
- b is the null vector of H_Q;
- Σ is the set that contains vectors of H_Q;
- δ is an automorphism from H_Q to itself;
- q_0 is a special state, called *starting state*;
- F are special states, called *final states*;

The **quantum speedup**, that is, the improvement in algorithm speed that a quantum computer has with respect to classical computers, is not due to the raw power of the machine. It is instead due to the fact that complexity classes of quantum algorithms are not arranged in the same way as classical algorithms: