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 $z = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$:

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

1.2. Hilbert spaces

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

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

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

The square root of the inner product of a vector x with itself is called the **norm** of the vector:

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

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.

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

Every element $x \in V \oplus W$ can be written as $v \oplus w$, for some $v \in V$ and some $w \in W$. 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:

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

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 v and w is the vector $v \otimes w$ constructed as:

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

The tensor product between vectors satisfies the following relations:

- $(v_1 + v_2) \otimes v_3 = v_1 \otimes (v_2 + v_3) = v_1 \otimes v_2 + v_1 \otimes v_3$
- $(av_1) \otimes v_2 = v_1 \otimes (av_2) = a(v_1 \otimes v_2)$

> **Exercise 1.2.1**: What is the tensor product of the two following vectors?
>
> $$v = \begin{pmatrix} 1 \\ 7 \end{pmatrix} \qquad\qquad w = \begin{pmatrix} 3 \\ 10 \end{pmatrix}$$

Solution:

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

$\square$

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

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

By definition, $\dim(A \otimes B) = \dim(A)\dim(B)$. 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 $A = \{v_1, v_2, ..., v_n\}$ and $B = \{w_1, w_2, ..., w_m\}$ 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:

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

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

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

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

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 $A = \{v_1, v_2, ..., v_n\}$ and $B = \{w_1, w_2, ..., w_m\}$ respectively. If A and B are both orthonormal, then the basis:

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

Of $V \otimes W$ is also orthonormal.

The **outer product** of two vectors v and w, denoted as $\rangle v, w \langle$, is given by:

$$\rangle v, w \langle = vw^\dagger = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}^\dagger = \begin{pmatrix} v_1 w_1^* & v_1 w_2^* & \cdots & v_1 w_n^* \\ v_2 w_1^* & v_2 w_2^* & \cdots & v_2 w_n^* \\ \vdots & \vdots & \ddots & \vdots \\ v_m w_1^* & v_m w_2^* & \cdots & v_m w_n^* \end{pmatrix}$$

Where m is the dimension of v and n is the dimension of w.

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:

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

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:

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

$\square$

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

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 $v \in V$ associated to λ, $Ov = \lambda v$. Consider $\lambda \langle x, x \rangle$:

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle Ov, v \rangle = \langle v, O^\dagger v \rangle = \langle v, Ov \rangle = \langle v, \lambda v \rangle = \lambda^* \langle v, v \rangle$$

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

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

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

Theorem 1.3.2: If A and B are two unitary matrices, then $(AB)^\dagger AB = I$. In other words, the product of two (product-conforming) unitary matrices is a unitary matrix or, equivalently, the set of unitary matrices of a certain dimension is closed under multiplication.

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

$$(AB)^\dagger AB = B^\dagger A^\dagger AB = B^\dagger (A^\dagger A) B = B^\dagger I B = B^\dagger B = I$$

$\square$

Theorem 1.3.3: Let U be a matrix. The following definitions are equivalent (if one is true, the others are as well):

- U is a unitary matrix;
- The columns of U form an orthonormal basis of $\mathbb{C}^n$ with respect to the inner product;
- The rows of U form an orthonormal basis of $\mathbb{C}^n$ with respect to the inner product;
- U is L^2-norm invariant: given any vector $v \in \mathbb{C}^n$, $||Uv|| = ||v||$.

Lemma 1.3.1: Unitary matrices are inner product-invariant. In other words, given a unitary matrix U and two vectors $u, v \in \mathbb{C}^n$, $\langle Uu, Uv \rangle = \langle u, v \rangle$.

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 operators define a unique orthogonal subspace decomposition called **eigenspace decomposition**. Moreover, for every such decomposition, there exists a Hermitian operator whose eigenspace decomposition is this decomposition.

> **Theorem 1.3.4**: Let $O : V \to V$ be an Hermitian operator in matrix form, having eigenvalues $\lambda_1, ..., \lambda_n$. Let S_{λ_i} be the subspace generated by the i-th eigenvalue, and let $\{S_{\lambda_i}\}$ be the set that contains those subspaces. $\{S_{\lambda_i}\}$ is an orthogonal set and:
>
> $$S_{\lambda_1} \oplus S_{\lambda_2} \oplus ... \oplus S_{\lambda_n} = V$$
>
> This decomposition is unique.

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 $\langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = 0$.

$\square$

1.4. Bra-ket notation

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

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

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 $|\Psi\rangle \in$ H then $X(\Psi) \in \mathbb{C}$:

$$X(\alpha_1 |\Psi_1\rangle + \alpha_2 |\Psi_2\rangle) = X(\alpha_1 |\Psi_1\rangle) + X(\alpha_2 |\Psi_2\rangle)$$

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**:

$$\langle|\Psi_1\rangle, |\Psi_2\rangle\rangle = |\Psi_1\rangle^\dagger |\Psi_2\rangle = \langle\Psi_1|\Psi_2\rangle \qquad\qquad \rangle |\Psi_1\rangle, |\Psi_2\rangle\langle = |\Psi_1\rangle|\Psi_2\rangle^\dagger = |\Psi_1\rangle\langle\Psi_2|$$

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.

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.
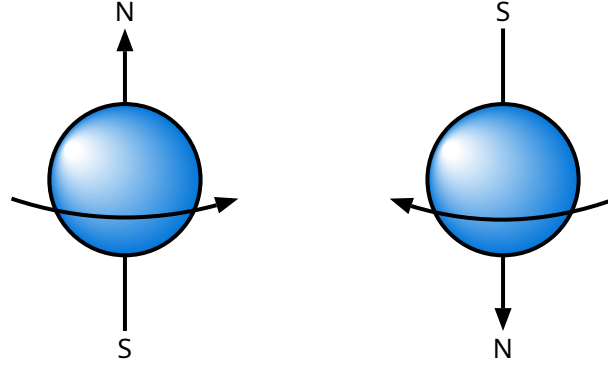

Figure 1: The spin is an intrinsic property of fundamental particles. Electrons have a value of spin that is either equal to $1/2$ or $-1/2$, also referred to as "up" and "down" respectively.

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:

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

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 $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, 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:

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

$$\square$$

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 $\{|+\rangle, |-\rangle\}$:

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

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 $\{|+\rangle, |-\rangle\}$:

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

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

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

$$\square$$

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

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

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 $\{|\circlearrowleft\rangle, |\circlearrowright\rangle\}$:

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

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

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

$\square$

According to Born's rule, the probability of finding $|\Psi\rangle$ in the state $|\varphi_1\rangle$ when measured is given by $|\alpha|^2$, whereas 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:

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

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:

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

$$=\frac{1}{2}(|\circlearrowleft\rangle+i\mid\circlearrowright\rangle)-\frac{1}{2}(|\circlearrowleft\rangle-i\mid\circlearrowright\rangle)=\frac{1}{2}\cancel{|\circlearrowleft\rangle}+\frac{i}{2}\mid\circlearrowright\rangle-\frac{1}{2}\cancel{|\circlearrowleft\rangle}+\frac{i}{2}\mid\circlearrowright\rangle=$$

$$=i\mid\circlearrowright\rangle$$

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

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 $\{|0\rangle,|1\rangle\}$?

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

Solution: Note how:

$$|\Psi_1\rangle=\frac{\sqrt{2}}{2}(|+\rangle+|-\rangle)=\frac{\sqrt{2}}{2}\left(\frac{\sqrt{2}}{2}(|0\rangle+|1\rangle)+\frac{\sqrt{2}}{2}(|0\rangle-|1\rangle)\right)=\frac{1}{2}(|0\rangle+|1\rangle)+\frac{1}{2}(|0\rangle-|1\rangle)=$$

$$=\frac{1}{2}\mid0\rangle+\frac{1}{2}\cancel{|1\rangle}+\frac{1}{2}\mid0\rangle-\frac{1}{2}\cancel{|1\rangle}=|0\rangle$$

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 $|\Psi_1\rangle=1\mid0\rangle+0\mid1\rangle$, which is a trivial combination. On the other hand:

$$|\Psi_2\rangle = \frac{\sqrt{3}}{2}\,|+\rangle - \frac{1}{2}\,|-\rangle = \frac{\sqrt{3}}{2}\left(\frac{\sqrt{2}}{2}(|0\rangle + |1\rangle)\right) - \frac{1}{2}\left(\frac{\sqrt{2}}{2}(|0\rangle - |1\rangle)\right) =$$

$$= \frac{\sqrt{6}}{4}(|0\rangle + |1\rangle) - \frac{\sqrt{2}}{4}(|0\rangle - |1\rangle) = \frac{\sqrt{6}}{4}\,|0\rangle + \frac{\sqrt{6}}{4}\,|1\rangle - \frac{\sqrt{2}}{4}\,|0\rangle + \frac{\sqrt{2}}{4}\,|1\rangle =$$

$$= \frac{\sqrt{6} - \sqrt{2}}{4}\,|0\rangle + \frac{\sqrt{6} + \sqrt{2}}{4}\,|1\rangle$$

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**:
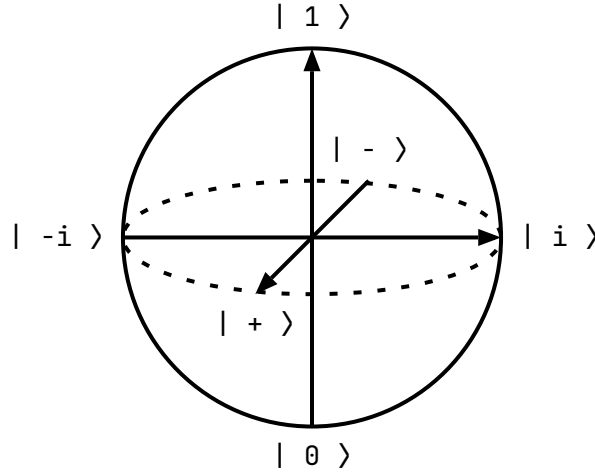
2.2. Multiple qubits systems

Consider two Hilbert spaces H_1 and H_2, each describing the state space of a qubit. The Hilbert space describing the state of the two qubits is given by the tensor product of the two Hilbert spaces. If the two qubits have $\{|\varphi_1\rangle_1, |\varphi_2\rangle_1\}$ and $\{|\varphi_1\rangle_2, |\varphi_2\rangle_2\}$ respectively as bases, the basis of $H_1 \otimes H_2$ 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:

$$\lambda_{1,1}|\varphi_1\rangle_1 \otimes |\varphi_1\rangle_2 + \lambda_{1,2}|\varphi_1\rangle_1 \otimes |\varphi_2\rangle_2 + \lambda_{2,1}|\varphi_2\rangle_1 \otimes |\varphi_1\rangle_2 + \lambda_{2,2}|\varphi_2\rangle_1 \otimes |\varphi_2\rangle_2$$

In general, the subscript from the basis is dropped, since the ordering of the states mirrors the orders of the qubits. Also, it is convenient to use the shortand $|vw\rangle$ for $|v\rangle \otimes |w\rangle$, allowing the basis of the two-qubit system to be written more compactly:

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

Note that, due to Lemma 1.2.1 and Lemma 1.2.2, tensor product preserves both the unity 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.

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 $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. 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 justaposing 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 2.2.1: 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:

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

$\square$

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**.

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.

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 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**, and indeed the majority of states of a n-qubit system are entangled states. For example, Bell states are entangled states.

> **Exercise 2.2.2**: Consider this two 2-qubit states. Are they entangled states?
>
> $$|\varphi_1\rangle = \frac{1}{8}|00\rangle + \frac{\sqrt{7}}{8}|01\rangle + \frac{\sqrt{7}}{8}|10\rangle + \frac{7}{8}|11\rangle \qquad\qquad |\varphi_2\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

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. $\qquad\qquad\square$

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 $H_1 \otimes H_2 \otimes ... \otimes H_n$ is said to be *separable* or *unentangled* if:

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

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).

2.3. Qubit measurement

In the first section, the issue of measuring the state of a single-qubit system was introduced. Measuring the state of a n-qubit system follows the same idea, but the set of possible measurements and measurement outcomes is significantly larger. It also serves the purpose of describing in greater detail the formal process of measuring in general.

First, recall from Theorem 1.2.1 that any vector space can be decomposed into a direct sum of one or more of its orthogonal subspaces. This means that the state space of an n-qubit system, having dimension 2^n, can be decomposed into $k \leq 2^n$ of its orthogonal subspaces. The number k corresponds to the maximum number of possible measurement outcomes for a state measured with that particular device. This number varies from device to device, even between devices measuring the same system.

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 2.3.1**: 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. $\square$

When a measuring device with associated direct sum decomposition $V = S_1 \oplus ... \oplus S_k$ 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.

More formally, any state $|\Psi\rangle$ in V has a unique direct sum decomposition $|\Psi\rangle = a_1\,|\varphi_1\rangle \oplus ... \oplus a_k\,|\varphi_k\rangle$, where each $|\varphi_i\rangle$ is a unit vector in S_i and a_i is a real and non-negative number. When $|\Psi\rangle$ is measured, the state $|\varphi_i\rangle$ is obtained with probability equal to $|a_i|^2$. This is not a fact that can be deduced from theory, but is instead an axiom of quantum mechanics.

> **Exercise 2.3.2**: Rephrase the measurement of the 2-qubit state $|\Psi\rangle = \frac{1}{8}\,|00\rangle + \frac{\sqrt{7}}{8}\,|01\rangle + \frac{\sqrt{7}}{8}\,|10\rangle + \frac{7}{8}\,|11\rangle$ under this formalism.

Solution: 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:

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

To denote more explicitly what happens when a measurement is performed on the first qubit, the state can be rewritten as $|\Psi\rangle = c_1\,|\Psi_1\rangle + c_2\,|\Psi_2\rangle$, or more explicitly:

$$|\Psi\rangle = \sqrt{\left|\frac{1}{8}\right|^2 + \left|\frac{\sqrt{7}}{8}\right|^2}\left(\frac{\frac{1}{8}\,|00\rangle + \frac{\sqrt{7}}{8}\,|01\rangle}{\sqrt{\left|\frac{1}{8}\right|^2 + \left|\frac{\sqrt{7}}{8}\right|^2}}\right) + \sqrt{\left|\frac{\sqrt{7}}{8}\right|^2 + \left|\frac{7}{8}\right|^2}\left(\frac{\frac{\sqrt{7}}{8}\,|10\rangle + \frac{7}{8}\,|11\rangle}{\sqrt{\left|\frac{\sqrt{7}}{8}\right|^2 + \left|\frac{7}{8}\right|^2}}\right)$$

In this way, the state is correctly normalized and the probabilities for the two possible outcomes for the first qubit can be computed as:

$$P_1 = \left|\frac{1}{8}\right|^2 + \left|\frac{\sqrt{7}}{8}\right|^2 = \frac{1}{64} + \frac{7}{64} = \frac{1}{8} \qquad\qquad P_2 = \left|\frac{\sqrt{7}}{8}\right|^2 + \left|\frac{7}{8}\right|^2 = \frac{7}{64} + \frac{49}{64} = \frac{7}{8}$$

$\square$

Recall that the outer product of two kets results in a matrix. Since any matrix is a representation of a linear operator, this must mean that the outer product of two kets results in an operator. More operators can be constructed by summing outer products. An operator is also represented by denoting how each vector of the basis is mapped to its result.

> **Exercise 2.3.3**: Consider the operator $X = |10\rangle\langle00| + |00\rangle\langle10| + |11\rangle\langle11| + |01\rangle\langle01|$. What's the corresponding matrix?

Solution: X exchanges the ket $|10\rangle$ with the ket $|00\rangle$ and vice versa, while leaving $|11\rangle$ and $|01\rangle$ unchanged. Keep in mind that the operator is constructed with respect to the standard basis.

$$X = |10\rangle\langle00| + |00\rangle\langle10| + |11\rangle\langle11| + |01\rangle\langle01| = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}^\dagger + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}^\dagger + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}^\dagger + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}^\dagger = $$

$$= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}(1\ 0\ 0\ 0) + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}(0\ 0\ 1\ 0) + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}(0\ 0\ 0\ 1) + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}(0\ 1\ 0\ 0) = $$

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

This can also be represented as:

$$|10\rangle \rightarrow |00\rangle \qquad |00\rangle \rightarrow |10\rangle \qquad |11\rangle \rightarrow |11\rangle \qquad |01\rangle \rightarrow |01\rangle$$

$\square$

Results in Exercise 2.3.3 can be generalized as follows. In an n-qubit system, an operator $O : V \to V$ for the basis $\{|\varphi_i\rangle\}$ on the state space V can be written as:

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

Where N is the dimension of V. The matrix for O with respect to the basis $\{|\varphi_i\rangle\}$ has the value p_{ij} in the (i, j)-th cell.

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 vector is equal to:

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

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 $s_1 + s_2$, with $s_1 \in S$ and $s_2 \in S^\perp$.

For any S of the sort it is possible to construct a linear operator $P_S : S \to V$, 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.

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 s_i belonging

the i-th subspace S_i. Framed this way, a measuring device with direct sum decomposition $V = S_1 \oplus ... \oplus S_k$ acting on a state $|\Psi\rangle$ results in the state

$$|\phi_i\rangle = \frac{P_i \, |\Psi\rangle}{|P_i \, |\Psi\rangle|}$$

With probability $|P_i \, |\Psi\rangle|^2$. 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:

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

Indeed, since the base vectors are all orthogonal, applying P_S to a generic state $|\Psi\rangle$ nulls all contributions given by basis absent in S. This also means that a projector is well-defined, since applying a projector two times results in no difference. Note that applying a projector to a state might not return a normalized vector.

Exercise 2.3.4: Consider the two-qubit state $|\Psi\rangle = \frac{1}{8} |00\rangle + \frac{\sqrt{7}}{8} |01\rangle + \frac{\sqrt{7}}{8} |10\rangle + \frac{7}{8} |11\rangle$ and the subspace spanned by $\{|00\rangle, |01\rangle\}$. What is the projector with respect to the basis? What is its effect on $|\Psi\rangle$?

Solution: The projector with respect to S is given by $P_S = |00\rangle\langle00| + |01\rangle\langle01|$. Applying to $|\Psi\rangle$ gives:

$$P_S \, |\Psi\rangle = (|00\rangle\langle00| + |01\rangle\langle01|) \left(\frac{1}{8} |00\rangle + \frac{\sqrt{7}}{8} |01\rangle + \frac{\sqrt{7}}{8} |10\rangle + \frac{7}{8} |11\rangle \right) = \frac{1}{8} |00\rangle + \frac{\sqrt{7}}{8} |01\rangle$$

Note how this state is not normalized. Note also how applying P_S to the result does not change the state:

$$P_S(P_S \, |\Psi\rangle) = (|00\rangle\langle00| + |01\rangle\langle01|) \left(\frac{1}{8} |00\rangle + \frac{\sqrt{7}}{8} |01\rangle \right) = \frac{1}{8} |00\rangle\langle00|00\rangle + \frac{\sqrt{7}}{8} |00\rangle\langle00|01\rangle +$$

$$\frac{1}{8} |01\rangle\langle01|00\rangle + \frac{\sqrt{7}}{8} |01\rangle\langle01|01\rangle = \frac{1}{8} |00\rangle + \frac{\sqrt{7}}{8} |01\rangle = P_S \, |\Psi\rangle$$

$\square$

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 2.3.5: Consider the state $|\Psi\rangle = a_{00} |00\rangle + a_{01} |01\rangle + a_{10} |10\rangle + a_{11} |11\rangle$ and a decomposition $S_{00} \oplus S_{01} \oplus S_{10} \oplus S_{11}$. Consider the projector P_{01} and describe the assigned measurement.

Solution: Applying the projection P_{01} to $|\Psi\rangle$ gives:

$$P_{01} \, |\Psi\rangle = |01\rangle\langle01| \, (a_{00} |00\rangle + a_{01} |01\rangle + a_{10} |10\rangle + a_{11} |11\rangle) =$$
$$= a_{00} |01\rangle\langle01|00\rangle + a_{01} |01\rangle\langle01|01\rangle + a_{10} |01\rangle\langle01|10\rangle + a_{11} |01\rangle\langle01|11\rangle = a_{01} |01\rangle$$

The state $\frac{P_{01} \, |\Psi\rangle}{|P_{01} \, |\Psi\rangle|}$ is measured with probability:

$$|P_{01} \, |\Psi\rangle|^2 = \langle\Psi|P_{01}|\Psi\rangle = \langle\Psi| \, (|01\rangle\langle01|) \, |\Psi\rangle = \langle\Psi|01\rangle\langle01|\Psi\rangle = a_{01}^\dagger a_{01} = |a_{01}|^2$$

$\square$

Recall that, for any Hermitian operator, Theorem 1.3.4 describes a unique decomposition into its eigenspaces. Given this correspondence, Hermitian operators can be used to describe measurements.

Let P_i be the projectors onto the subspaces S_i, and let $\{\lambda_1, ..., \lambda_k\}$ be a set of distinct real values. Then, the Hermitian operator $O = \sum_{i=1}^{k} \lambda_i P_i$ has $S_1 \oplus ... \oplus S_k$ as its direct sum decomposition. Thus, when describing a measurement, instead of directly specifying the associated subspace decomposition, it is sufficient to specify a Hermitian operator whose eigenspace decomposition is that decomposition.

Any Hermitian operator 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[3].

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. The Hermitian operator is just a way (frequently used in quantum mechanics) to write in compact form the projectors associated to a measuring apparatus.

2.4. Qubit manipulations

Manipulating the state of a qubit is performed through a **quantum transformation**. 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 respected, before introducing some examples of quantm 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:

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

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 1.3.1 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.

The tensor product $U_1 \otimes U_2$ is a unitary transformation of the space $X_1 \otimes X_2$ if U_1 and U_2 are unitary transformations of X_1 and X_2 respectively.

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.

---

[3]In quantum physics, these labels are often chosen to represent a shared property, such as the energy, of the eigenstates in the corresponding eigenspace.

The term "gate" is used to suggest a similarity with the classical logical gates, but does not necessarily 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.

$$|\Psi\rangle - \boxed{U_1} - \boxed{U_2} - \ ... \ - \boxed{U_n} - |\Psi'\rangle \qquad\qquad |\Psi'\rangle = U_n...U_2U_1\,|\Psi\rangle$$

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 $\boxed{\nearrow}$. 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**[4], are also obiquitous:

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

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

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:

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

Which produces an even superposition of $|0\rangle$ and $|1\rangle$ from either of the standard basis elements:

$$H\,|0\rangle = |+\rangle \qquad\qquad H\,|0\rangle = |-\rangle \qquad\qquad H\,|+\rangle = |0\rangle \qquad\qquad H\,|-\rangle = |1\rangle$$

Extending quantum transformations that act on single qubits from an n qubit system is trivial. As a matter of fact, applying a transformation U to the i-th qubit from an n qubit system is equivalent to applying $I \otimes ... \otimes I \otimes U \otimes I \otimes ... \otimes I$ to the entire system, that is applying U to the i-th qubit and applying the identity operator to each of the other qubits of the system.

In general, the simplest quantum transformations that acts on more than a single qubit at a time are those that can be reduced to the application of a transformation to each qubit separately. For example, given the

---

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

transformation $U \otimes V$ that acts on two qubits at the same time can be broken down into first applying $U \otimes I$ and then applying $I \otimes V$.

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:

$$C_{\text{not}} = |00\rangle\langle00| + |01\rangle\langle01| + |11\rangle\langle10| + |10\rangle\langle11| = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

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

Solution:

$$(|00\rangle\langle00| + |01\rangle\langle01| + |11\rangle\langle10| + |10\rangle\langle11|)\frac{\sqrt{2}}{2}(|00\rangle + |10\rangle) = \frac{\sqrt{2}}{2}(|00\rangle + |11\rangle)$$

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 2.4.2**: 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:

$$(|00\rangle\langle00| + |01\rangle\langle01| + |11\rangle\langle10| + |10\rangle\langle11|)\frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle) = \frac{1}{2}(|00\rangle - |01\rangle - |10\rangle + |11\rangle)$$
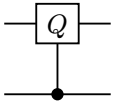
Converting it back:

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

$$= \frac{1}{2}\left(\sqrt{2}\,|-\rangle \otimes \sqrt{2}\,|-\rangle\right) = \frac{1}{2}(2\,|--\rangle) = |--\rangle$$

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:

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

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:

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

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:

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

2.5. Quantum 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 2.5.1**: 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 $\pi : \mathbf{Z}_{2^n} \to \mathbf{Z}_{2^n}$ 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:

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

Now consider a non-reversible classical function $f : \mathbf{Z}_{2^n} \to \mathbf{Z}_{2^m}$ with n input and m output bits. This function can be modified in a standard way to create a reversible function $\pi_f : \mathbf{Z}_{2^{n+m}} \to \mathbf{Z}_{2^{n+m}}$ 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[5] 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.

> **Exercise 2.5.2**: Construct a reversible version of the logical AND.

Solution:

Inputs			Outputs		
A	**B**	**C**	**A**	**B**	**C ∨ (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

□

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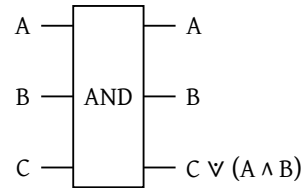
[5]A much more common notation for the logical XOR is $\oplus$, but this notation is here avoided because it conflicts with the direct sum symbol.

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

$$|x\rangle \quad \boxed{} \quad |x\rangle$$
$$U_f$$
$$|y\rangle \quad \boxed{} \quad |y \veebar f(x)\rangle$$
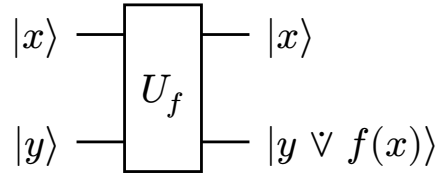
Note that, in general, this method of constructing reversible counterparts of non-reversible functions is highly inefficient, and there are ad-hoc methods that use less bits. This is not important, however, since the interest is to show that there is a method that always works, and therefore that each function that a classical computer can compute can be just as well computed by a quantum computer.

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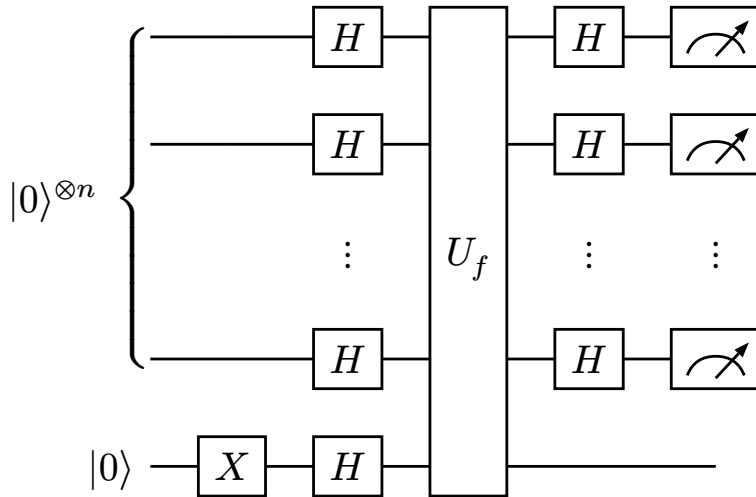
The circuit starts with $n + 1$ qubits initialized in the state $|0\rangle$. The last extra qubit, called **ancillary qubit**, is necessary for the quantum computation to be reversible. The starting state of the entire system can be then written as $|000...0\rangle$.

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:

$$|\Psi_1\rangle = (H \otimes H \otimes ... \otimes H)\, |\Psi_0\rangle = (H\,|0\rangle) \otimes (H\,|0\rangle) \otimes ... \otimes (H\,|1\rangle) =$$

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

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

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

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

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

Consider $|x\rangle\,|0 \vee f(x)\rangle - |x\rangle\,|1 \vee f(x)\rangle$. Since 0 and 1 are single bits and opposite in value, $|0 \vee f(x)\rangle$ and $|1 \vee f(x)\rangle$ 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:

$$H\,|x_1 x_2 ... x_n\rangle = H\,|x_1\rangle \otimes H\,|x_2\rangle \otimes ... \otimes H\,|x_n\rangle =$$

$$= \frac{1}{\sqrt{2}}(|0\rangle + (-1)^{x_1}\,|1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + (-1)^{x_2}\,|1\rangle) \otimes ... \otimes \frac{1}{\sqrt{2}}(|0\rangle + (-1)^{x_n}\,|1\rangle) =$$

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

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

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

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

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:

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

As for the second case:

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

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.

2.5.2. 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:
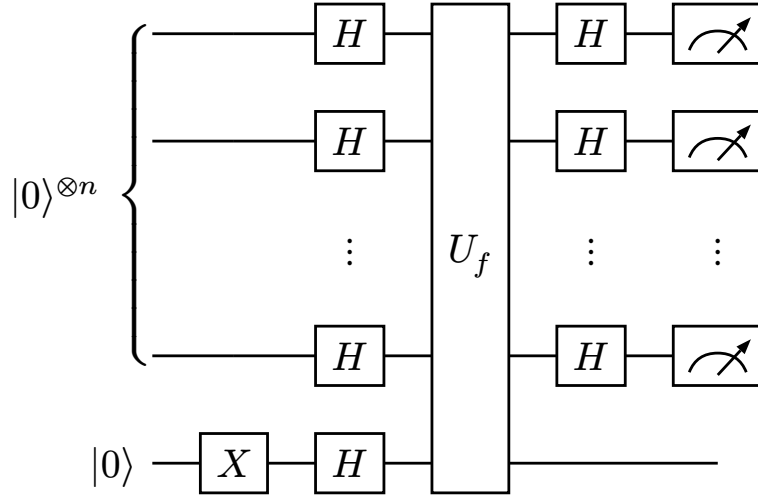
The algorithm starts with $n + 1$ qubits in state $|0\rangle$, where each of the first n qubits represents one of the bits of the string and the last qubit is an ancillary qubit. The last qubit is changed from $|0\rangle$ to $|1\rangle$, leaving the other qubits unchanged:

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

A Hadamard gate is the applied to all the qubits:

$$|\Psi_1\rangle = (H \otimes H \otimes ... \otimes H \otimes H) |\Psi_0\rangle = (H\,|0\rangle) \otimes (H\,|0\rangle) \otimes ... \otimes (H\,|1\rangle) =$$

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

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

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

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

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

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

Discarding the last qubit and applying Hadamard again:

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

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

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

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

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 $O(1)$.

2.5.3. 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:
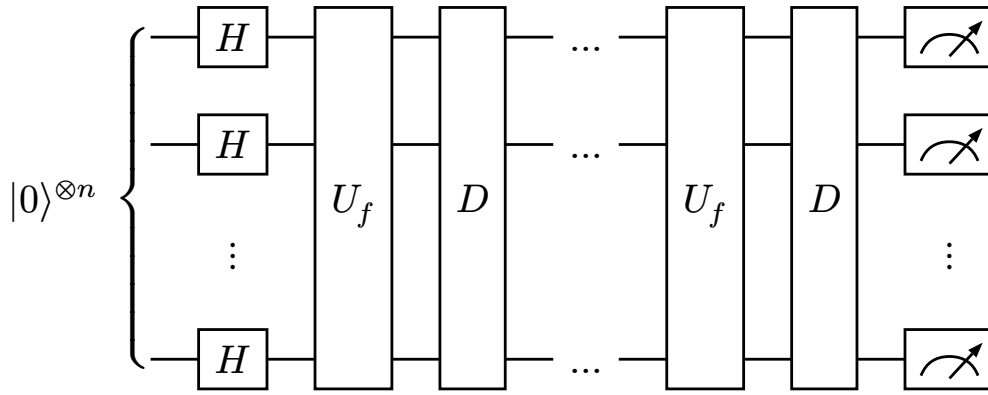
The algorithm starts with n qubits (each representing one of the bits of the string) prepared in the $|0\rangle$ state. Then, each qubit undergoes a Hadamard transformation; for the sake of clarity, the resulting state is denoted as $|\eta\rangle$:

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

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

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

Applying the oracle gives:

$$|\Psi_1\rangle = U_f |\Psi_0\rangle = U_f |\eta\rangle = U_f \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} |x\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} U_f |x\rangle =$$

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

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:

$$|\Psi_2\rangle = D\,|\Psi_1\rangle = (2\,|\eta\rangle\langle\eta| - I)\,|\Psi_1\rangle = (2\,|\eta\rangle\langle\eta| - I)\left(|\eta\rangle - \frac{2}{\sqrt{2^n}}\,|x_0\rangle\right) =$$

$$= 2\,|\eta\rangle\langle\eta|\eta\rangle - \frac{4}{\sqrt{2^n}}\,|\eta\rangle\langle\eta|x_0\rangle - I\,|\eta\rangle + I\frac{2}{\sqrt{2^n}}\,|x_0\rangle = |\eta\rangle - \frac{4}{\sqrt{2^n}}\,|\eta\rangle\langle\eta|x_0\rangle + \frac{2}{\sqrt{2^n}}\,|x_0\rangle$$

Note that, since all base states are orthonormal:

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

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:

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

$$= \left(1 - \frac{4}{2^n}\right)\left(\frac{1}{\sqrt{2^n}}\left(|x_0\rangle + \sum_{x\in\{0,1\}^n - \{x_0\}}|x\rangle\right)\right) + \frac{2}{\sqrt{2^n}}\,|x_0\rangle =$$

$$= \left(1 - \frac{4}{2^n}\right)\left(\frac{1}{\sqrt{2^n}}\,|x_0\rangle + \frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n - \{x_0\}}|x\rangle\right) + \frac{2}{\sqrt{2^n}}\,|x_0\rangle =$$

$$= \frac{1}{\sqrt{2^n}}\,|x_0\rangle + \frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n - \{x_0\}}|x\rangle - \frac{4}{2^n}\frac{1}{\sqrt{2^n}}\,|x_0\rangle - \frac{4}{2^n}\frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n - \{x_0\}}|x\rangle + \frac{2}{\sqrt{2^n}}\,|x_0\rangle =$$

$$= \left(\frac{1}{\sqrt{2^n}} - \frac{4}{2^n}\frac{1}{\sqrt{2^n}}\right)\sum_{x\in\{0,1\}^n - \{x_0\}}|x\rangle + \left(\frac{1}{\sqrt{2^n}} - \frac{4}{2^n}\frac{1}{\sqrt{2^n}} + \frac{2}{\sqrt{2^n}}\right)|x_0\rangle =$$

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

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. Quantum Theory

3.1. 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.

Many real implementations of quantum computers include:

- **Superconducting qubits**;
- **Semiconductors**;
- **Photonic qubits**;
- **Trapped ions**;
- **Neutral atoms**.

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.

3.2. 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 affirmative 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**:

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

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: