Appunti di

Aspetti matematici nella computazione quantistica

dalle lezioni di Paola Boito e Dario Trevisan

A cura di Francesco Baldino

Anno accademico 2022/2023

Contents

1 | Remarks on Hilbert Spaces

We'll start with some remarks on Hilbert Spaces that will allow us to introduce the notation that we will use throughout the course.

Definition 1.0.1 – Definite positive scalar product

A definite positive scalar product on a complex vector space $\mathbb{V}$ is a map $\langle\cdot|\cdot\rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{C}$ with the following properties

- $\langle\varphi|\psi\rangle = \overline{\langle\psi|\varphi\rangle}$
- $\langle\varphi|\varphi\rangle \geq 0, \forall\varphi \in \mathbb{V}$
- $\langle\varphi|\varphi\rangle = 0 \iff \varphi = 0$
- $\langle\varphi|a\psi_1 + b\psi_2\rangle = a\langle\varphi|\psi_1\rangle + b\langle\varphi|\psi_2\rangle \, \forall a, b \in \mathbb{C}, \forall\varphi, \psi_1\psi_2 \in \mathbb{V}$

Remark 1.0.1. A scalar product on $\mathbb{V}$ defines a norm on $\mathbb{V}$ defined as

$$\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$$

Definition 1.0.2 – Hilbert space

An Hilbert Space is a vector space $\mathbb{H}$ over $\mathbb{C}$ endoned with a definite positive scalar product such that $\mathbb{H}$ is complete with respecto to the given norm

We also have the following useful properties:

1. $\forall a \in \mathbb{C}$, $\forall\psi, \varphi \in \mathbb{H}$ it holds $\langle a\varphi|\psi\rangle = \overline{a}\langle\varphi|\psi\rangle$ and $\|a\varphi\| = |a| \cdot \|\varphi\|$
2. Let $\psi \in \mathbb{H}$, then $\langle\psi|\varphi\rangle = 0 \quad \forall\varphi \in \mathbb{H} \iff \psi = 0$

   Proof.
   $[\Rightarrow]$ it must also hold $\langle\psi|\psi\rangle = 0$ so $\psi = 0$
   $[\Leftarrow]$ it follows from linearity on the second argument □

3. *(Complex parallelogram identity)*
   $\forall\varphi, \psi \in \mathbb{H}$ it holds that

$$\langle\psi|\varphi\rangle = \frac{1}{4}\left(\|\psi + \varphi\|^2 - \|\psi - \varphi\|^2 + i\|\psi - i\varphi\|^2 - i\|\psi + i\varphi\|^2\right)$$

Definition 1.0.3

- A vector $\psi \in \mathbb{H}$ is unitary (normed) if $\|\psi\| = 1$
- $\varphi, \psi \in \mathbb{H}$ are orthogonal if $\langle\varphi|\psi\rangle = 0$
- Given $\psi \in \mathbb{H}$ the orthogonal subspace to ψ is

$$\mathbb{H}_{\psi^\perp} = \{\psi \in \mathbb{H} \,|\, \langle\varphi|\psi\rangle = 0\}$$

Note that given $\psi \in \mathbb{H}$, $\psi \neq 0$, then $\forall \varphi \in \mathbb{H}$, $\varphi - \frac{\langle\psi|\varphi\rangle}{\|\psi\|^2}\psi \in \mathbb{H}_{\psi^\perp}$

Let $\mathbb{H}$ be a Hilbert space and I a set of indices, then

- The vectors $\{\psi_i\}_{i\in I} \subset \mathbb{H}$ are linearly independent if for every finite subset $\{i_1, \ldots, i_n\} \subset I$ it holds that

$$a_1\psi_{i_1} + \cdots + a_n\psi_{i_n} = 0 \Rightarrow a_1 = \ldots a_n = 0$$

- $\mathbb{H}$ is finite-dimensional if any set of linearly independent vectors is finite
- An orthonormal basis *(ONB)* for $\mathbb{H}$ is a set of linearly independant vector $\{e_j\}_{j\in I} \subset \mathbb{H}$ such that $\langle e_i|e_j\rangle = \delta_{ij}$ and any $\psi \in \mathbb{H}$ can be written as

$$\psi = \sum_j a_j e_j$$

for some $a_j \in \mathbb{C}$

We will only work with separable Hilbert spaces, that is spaces where every basis has a countable number of elements.

Let $\psi, \varphi \in \mathbb{H}$, let $\{e_j\}$ be an ONB and let $\psi = \sum \psi_j e_j$ and $\varphi = \sum \varphi_j e_j$. Then it holds that $\langle\varphi|\psi\rangle = \sum \overline{\varphi_j}\psi_j$ and $\|\psi\| = \sqrt{\sum |\psi|^2}$.

In particular if $\langle\varphi|\psi\rangle = 0$ then $\|\psi + \varphi\|^2 = \|\psi\|^2 + \|\varphi\|^2$ because we get

$$\|\psi + \varphi\|^2 = \langle\psi + \varphi|\psi + \varphi\rangle = \|\psi\|^2 + \|\varphi\|^2 + \underbrace{\langle\psi|\varphi\rangle}_{=0} + \underbrace{\langle\varphi|\psi\rangle}_{=0}$$

Theorem 1.0.4 – Reisz representation theorem

All linear continuous maps $\mathbb{H} \to \mathbb{C}$ can be written as

$$\begin{aligned} \langle\psi|\cdot\rangle : \ \mathbb{H} &\longrightarrow \mathbb{C} \\ |\varphi\rangle &\mapsto \langle\psi|\varphi\rangle \end{aligned}$$

for some $\psi \in \mathbb{H}$

Definition 1.0.5 – Dual space

The space of linear continuous maps from $\mathbb{H}$ to $\mathbb{C}$is called the dual space

$$\mathbb{H}^* = \{f : \mathbb{H} \to \mathbb{C} \,|\, f \text{ is linear continuous}\}$$

As a concequence of the last theorem, $\mathbb{H}^*$ is in bijection with $\mathbb{H}$

We'll now introduce the *bra/ket* notation that we will use throghout the course:

Definition 1.0.6 – Dirac notation

- A ket vector is $|\psi\rangle \in \mathbb{H}$
- A bra vector is $\langle\psi| \in \mathbb{H}^*$

If $A : \mathbb{H} \to \mathbb{H}$ is a linear map then $|A\psi\rangle = A\,|\psi\rangle$

Note that this somewhat contradicts the notation we used beforehand, where we called vectors in $\mathbb{H}$ simply ψ and not $|\psi\rangle$. This notation will actually be really useful later on when the symbol inside the $|\cdot\rangle$ will be a way to identify some specific state (eg: if we put inside the $|\cdot\rangle$ a numeric value or a binary string, it will represent the corrisponding element of a fixed base of the space $\mathbb{H}$).

From now on vectors will always be represented with the ket notation, and the symbol inside the ket will either be a mute variable or a representation of something else.

Given an ONB $\{|e_j\rangle\}$ for $\mathbb{H}$ we can write any $|\psi\rangle$ as $|\psi\rangle = \sum_j \langle e_j|\psi\rangle\,|e_j\rangle$. If we then apply a linear map A we get

$$\begin{aligned} A\,|\psi\rangle &= \sum_j A\,|e_j\rangle\,\langle e_j|\psi\rangle \\ &\parallel \\ |A\psi\rangle &= \sum_j |e_j\rangle\,\langle e_j|A\psi\rangle = \sum_j |e_j\rangle \left\langle e_j \middle| A \sum_k |e_k\rangle\,\langle e_k|\psi\rangle \right\rangle \end{aligned}$$

thus we can write $A = \sum_{j,k} |e_j\rangle\,\langle e_j|e_k\rangle\,\langle e_k|$ as an operator acting on $|\psi\rangle$

Example 1.0.7 (The finite-dimensional case). We will often consider the finite-dimensional case where $\mathbb{H} \cong \mathbb{C}^n$ (that is $\dim\mathbb{H} = n$) which has an explicit representation given by $\{|e_j\rangle\}_{j=1}^n$ being the canonical basis (which is an ONB).

In this case we can easily represent elements $|\psi\rangle \in \mathbb{H}$, $\langle\psi| \in \mathbb{H}^*$ and products $|\psi\rangle\langle\varphi|$.

Because we can write $|\psi\rangle = \sum_{j=1}^n \psi_j\,|e_j\rangle$ there's a corrispondence between $|\psi\rangle$ and the complex column vector $(\psi_1, \ldots, \psi_n)^\intercal$.

Because we can write $\langle\psi| = \sum_{j=1}^n \langle e_j|\,\overline{\psi_j}$ there's a corrispondence between $\langle\psi|$ and the complex row vector $(\overline{\psi_1}, \ldots, \overline{\psi_n})$.

Then we can represent $|\psi\rangle\langle\varphi|$ as

$$|\psi\rangle\langle\varphi| = \begin{pmatrix}\psi_1\\ \vdots \\ \psi_n\end{pmatrix}\begin{pmatrix}\overline{\varphi_1} & \dots & \overline{\varphi_n}\end{pmatrix} = \begin{pmatrix}\psi_1\overline{\varphi_1} & \dots & \psi_1\overline{\varphi_n}\\ \vdots & & \vdots \\ \psi_n\overline{\varphi_1} & \dots & \psi_n\overline{\varphi_n}\end{pmatrix}$$

Example 1.0.8. Let's first consider the case $n = 2$

We have $\mathbb{H} \cong \mathbb{C}^2$. We usually represent an ONB for $\mathbb{H}$ as $\{|0\rangle, |1\rangle\}$ and the isomorphism is given by the map $|0\rangle \mapsto \begin{pmatrix}1\\0\end{pmatrix}$ and $|1\rangle \mapsto \begin{pmatrix}0\\1\end{pmatrix}$

Elements in $\mathbb{H}$ can be written as $|\psi\rangle = a|0\rangle + b|1\rangle \in \mathbb{H}$. By linearity, the isomorphism sends $|\psi\rangle$ to $\begin{pmatrix}a\\b\end{pmatrix}$.

Similarly, elements in $\mathbb{H}^*$ can be written $\langle\psi| = a\langle 0| + b\langle 1| \in \mathbb{H}^*$ which would be mapped to $(\overline{a}, \overline{b}) \in (\mathbb{C}^2)^*$

We've seen that given $|\psi\rangle, |\varphi\rangle \in \mathbb{H}$ we can construct the operator $|\psi\rangle\langle\varphi| : \mathbb{H} \to \mathbb{H}$. Let's see some examples. Using the elements of the ONB we get

- $|0\rangle\langle 0| \mapsto \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}$ which is the projection on the first coordinate
- $|1\rangle\langle 1| \mapsto \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0 & 1\end{pmatrix} = \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}$ which is the projection on the second coordinate
- $|0\rangle\langle 1| \mapsto \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0 & 1\end{pmatrix} = \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}$
- $|1\rangle\langle 0| \mapsto \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} = \begin{pmatrix}0 & 0\\1 & 0\end{pmatrix}$

By linearity, given $|\psi\rangle = a|0\rangle + b|1\rangle$ and $|\varphi\rangle = c|0\rangle + d|1\rangle$, we get

$$|\psi\rangle\langle\varphi| \mapsto \begin{pmatrix}a\\b\end{pmatrix}\begin{pmatrix}\overline{c} & \overline{d}\end{pmatrix} = \begin{pmatrix}a\overline{c} & b\overline{c}\\ a\overline{d} & b\overline{d}\end{pmatrix}$$

Let $A : \mathbb{H} \to \mathbb{H}$ be a linear operator. We will assume that A is bounded, that is

$$\|A\| = \sup\{\|A\psi\| \mid |\psi\rangle \in \mathbb{H}, \|\psi\| = 1\} < \infty$$

Definition 1.0.9 – Adjoint

The adjoint of A is an operator $A^* : \mathbb{H} \to \mathbb{H}$ such that $\forall |\psi\rangle, |\varphi\rangle \in \mathbb{H}$ it holds

$$\langle A^*\psi|\varphi\rangle = \langle\psi|A\varphi\rangle$$

We say that A is self-adjoint if $A = A^*$

The following properties of the adjoint hold:

- $(A^*)^* = A$, because $\forall \left|\psi\right\rangle, \left|\varphi\right\rangle \in \mathbb{H}$ we have
$$\langle (A^*)^*\psi|\varphi\rangle = \langle \psi|A^*\varphi\rangle = \overline{\langle A^*\varphi|\psi\rangle} = \overline{\langle \varphi|A\psi\rangle} = \langle A\psi|\varphi\rangle$$
- $\forall c \in \mathbb{C}\ (cA)^* = \overline{c}(A^*)$, because
$$\langle \psi|cA\varphi\rangle = c\left\langle \psi|A\varphi\right\rangle = c\left\langle A*\psi|\varphi\right\rangle = \langle \overline{c}A^*\psi|\varphi\rangle$$
- $(AB)^* = B^*A^*$
- $(A^*)_{ij} = \overline{A_{ij}}$, because
$$(A^*)_{ij} = \langle e_i|A^*e_j\rangle = \langle Ae_i|e_j\rangle = \overline{\langle e_j|Ae_i\rangle} = \overline{A_{ij}}$$
- $\langle A\psi| = \left\langle\psi\right| A^*$, where
$$\begin{aligned} \langle A\psi| : \ \mathbb{H} &\longrightarrow \mathbb{C} \\ |\varphi\rangle &\longmapsto \langle A\psi|\varphi\rangle \end{aligned}$$
$$\begin{aligned} \left\langle\psi\right| A^* : \ \mathbb{H} &\longrightarrow \mathbb{C} \\ |\varphi\rangle &\longmapsto \langle \psi|A^*\varphi\rangle \end{aligned}$$
- $(|\psi\rangle\langle\varphi|)^* = |\varphi\rangle\langle\psi|$, because $\forall \left|\xi\right\rangle, \left|\eta\right\rangle \in \mathbb{H}$ it holds
$$\begin{aligned} \langle(|\psi\rangle\langle\varphi|)^*\eta|\xi\rangle &= \langle\eta|(|\psi\rangle\langle\varphi|)\xi\rangle = \\ &= \langle\eta|\psi\rangle\left\langle\varphi|\xi\right\rangle = \\ &= \langle\varphi|\xi\rangle\left\langle\eta|\psi\right\rangle = \\ &= \overline{\langle\xi|\varphi\rangle\langle\psi|\eta\rangle} = \\ &= \overline{\langle\xi|\varphi\rangle\left\langle\psi|\eta\right\rangle} = \\ &= \overline{\langle\xi|(|\varphi\rangle\langle\psi|)\eta\rangle} = \\ &= \langle(|\varphi\rangle\langle\psi|)\eta|\xi\rangle \end{aligned}$$

Definition 1.0.10 – Unitary operator

We say that $U : \mathbb{H} \to \mathbb{H}$ is unitary if $\forall \left|\psi\right\rangle, \left|\varphi\right\rangle \in \mathbb{H}\ \langle U\psi|U\varphi\rangle = \langle\psi|\varphi\rangle$

Proposition 1.0.11

The following properties are equivalent:

1. U is unitary
2. $U^*U = \mathrm{Id}$
3. $\forall \left|\psi\right\rangle \in \mathbb{H}\ \|U\psi\| = \|\psi\|$

Proof.
$[1 \Rightarrow 2]$ It holds that $\forall \left|\psi\right\rangle \left|\varphi\right\rangle \in \mathbb{H}$, $\langle\psi|\varphi\rangle = \langle U\psi|U\varphi\rangle = \langle U^*U\psi|\varphi\rangle$, so it must be $U^*U = \mathrm{Id}$

$[1 \Rightarrow 3]$ $\|U\psi\| = \sqrt{\langle U\psi|U\varphi\rangle} = \sqrt{\langle\psi|\psi\rangle} = \|\psi\|$

The left of the proof is left as an exercise to the reader □

Remark 1.0.2. In the finite-dimension case, if $\{e_j\}_{j\in I}$ is an ONB and U is unitary, then $\{Ue_j\}_{j\in I}$ is also ONB

Definition 1.0.12 – Eigenvalues and eigenvectors

Let A be an operator on $\mathbb{H}$. We say that a non-zero vector $\left|\psi\right\rangle \in \mathbb{H}$ is an eigenvector of A with respect to to the eigenvalue $\lambda \in \mathbb{C}$ if $A\left|\psi\right\rangle = \lambda\left|\psi\right\rangle$

We call $\sigma(A) = \{\lambda \in \mathbb{C} \,|\, (A - \lambda\mathrm{Id}) \text{ is not invertible}\}$ the spectrum of A

Remark 1.0.3. If λ is an eigenvalue for A, then $\lambda \in \sigma(A)$. In the finite-dimensional case, the viceversa is also true

Let $A = A^*$ and λ be an eigenvalue of A with eigenvector $\left|\psi\right\rangle$. Then we get

$$\begin{array}{c} \langle\psi|A\psi\rangle = \langle\psi|\lambda\psi\rangle = \lambda\langle\psi|\psi\rangle = \lambda\|\psi\| \\ \| \\ \langle A^*\psi|\psi\rangle = \langle A\psi|\psi\rangle = \langle\lambda\psi|\psi\rangle = \overline{\lambda}\|\psi\| \end{array}$$

from which we deduce that $\lambda \in \mathbb{R}$

Let U be a unitary operator and λ an eigenvalue of U with eigenvector $\left|\psi\right\rangle$. Then we get $\|\psi\| = \|U\psi\| = \|\lambda\psi\| = |\lambda|\|\psi\|$ from which we deduce that $|\lambda| = 1$

Let A be a compact self-adjointed operator on $\mathbb{H}$. Then A can be diagonalized with respect to an ONB of $\mathbb{H}$. If $\{\lambda_j\}_{j\in I}$ are the eigenvalues of A, then $A = \sum_{j,\alpha} \lambda_j \left|e_{j,\alpha}\right\rangle \left\langle e_{j,\alpha}\right|$ for some ONB $\{\left|e_{j,\alpha}\right\rangle\}$ (where the vectors of the basis are indicized on j and some other index α because A might have some eigenvalues of multiplicity bigger than 1)

Definition 1.0.13 – Projection operator

A projection is an operator $P : \mathbb{H} \to \mathbb{H}$ such that $P^2 = P$. Let $\mathbb{K}$ be a subspace of $\mathbb{H}$, then if $P(\mathbb{H}) \subset \mathbb{K}$ we say that P is a projection onto $\mathbb{K}$

We say that P is an orthogonal projection if $P^2 = P$ and $P^* = P$

Remark 1.0.4. If P is an orthogonal projection, then there exists an orthonormal set $\{\left|\psi_j\right\rangle\}_{j\in I}$ such that $P = \sum_j |\psi_j\rangle\langle\psi_j|$. This is a consequence of the property of the possible eigenvalues of a projection (0 and 1) and the diagonal representation of self-adjointed operators

Definition 1.0.14 – Commutator

Given two operators $A, B : \mathbb{H} \to \mathbb{H}$, we call the commutator of A and B the operator $[A, B] = AB - BA$.
A and B commute if and only if $[A, B] = 0$

Note that given $A = A^*$ and $B = B^*$ then $(AB)^* = AB$ if and only if A and B commute

Definition 1.0.15

Given an operator A we say that A is

- positive, if $\forall \left|\psi\right\rangle \in \mathbb{H} \ \langle\psi|A\psi\rangle \geq 0$
- strictly positive, if it's positive and $\langle\psi|A\psi\rangle = 0 \iff \left|\psi\right\rangle = 0$

Definition 1.0.16 – Trace

We call trace operator the map $A \mapsto \text{tr}\,(A) = \sum_j \langle e_j|Ae_j\rangle$ for some $\{e_j\}_{j\in I}$ ONB.
Note that this definition is independent on the choiche of the ONB.

Moreover, $\forall A, B : \mathbb{H} \to \mathbb{H}$ it holds $\text{tr}\,(AB) = \text{tr}\,(BA)$

2 | Quantum Mechanics

The theory of quantum mechanics was initialy developed at the beginning of the XX century by physicists Heisenberg, Pauli, Weyl and Schrödinger.

For what concerns us, quantum mechanics gives a set of rules to compute probabilities associated to events related to "measures", which practically can be something along the line of "the electron will be located in some volume $D \subset \mathbb{R}^3$" or "the light ray will be polarized along some direction $\vec{v}$".

More precisely, we will not bother with the details of actual quantum mechanics in a physical world, and we will only focus on the possible behaviours of a quantum computer, which we will formalize and take for granted as a set of postulates . These postulates will have each their own physical interpretation, but they can be seen as just a set of rules that we can use and must obey when building quantum algorithms and quantum circuits.

The approach we will use to formalize quantum mechanics is the frequentist approach: every single event will be random, but by repeating the same experiment over and over, the fraction of success will approach the probability of success of the single random event.

We will see that many of the algorithms we will study are only probabilistic, which means that the algorithm will only have a (hopefully high) probability of returning the correct result. It might be the case that some algorithms will have to be ran multiple times in order to get the correct result with sufficiently high probability. Still, most of the times this will be only a linear increase of the cost, which will make those algorithms still faster than the corresponding classical algorithm used to solve the same problem.

2.1 Postulates

First we'll introduce some instruments from probability theory:

Definition 2.1.1

Given a discrete space of events $\Omega = \{1, \dots, n\}$ and a probability density $(p_j)_{j=1}^n$ such that $p_j \in [0, 1]$ and $\sum_{j=1}^n p_j = 1$, we define the expected value of some function $f : \Omega \to \mathbb{R}$ as

$$\mathbb{E}_p[f] = \sum_{j=1}^n p_j f(j)$$

We're now going to replace many of these concepts, and introduce four postulates that will be the base of our formalization of quantum mechanics.

First, we note that f can be described as a vector $(f_j)_{j=1}^n$ where $f_j = f(j)$. Then, we can replace the expected value with

$$\mathbb{E}_p[f] = \sum_{j=1}^n p_j f_j = \left[(\sqrt{p_j})_{j=1}^n\right]^\intercal \cdot F \cdot (\sqrt{p_j})_{j=1}^n$$

for $F = \text{diag}(f_1, \dots, f_n)$

Definition 2.1.2 – Postulate 1

A quantum system is described as a Hilbert space $\mathbb{H}$. Given $\mathbb{H}$:

- an observable is any self-adjointed $A : \mathbb{H} \to \mathbb{H}$, and represents a physically measurable quantity of a quantym system
- a pure state is any $|\psi\rangle \in \mathbb{H}$ such that $\|\psi\| = 1$

When we say that the physical quantum system is in state $|\psi\rangle$ we mean that for every observable A, its quantum expected value $\langle A\rangle_{|\psi\rangle}$ is $\langle A\rangle_{|\psi\rangle} = \langle\psi|A\psi\rangle$. In this case we call $|\psi\rangle$ the state vector.

We will identify the physical system with the Hilbert space $\mathbb{H}$

The role of $|\psi\rangle$ corresponds in a way to the role of (the square root of) the fixed probability distribution p, and saying that the system is in a state $|\psi\rangle$ corresponds to fixing a certain probability over the space of events. We will see later that this correspondence between states and probability distributions does not work as nicely as we would like (in particular for the interference that we will describe later on).

A simple deduction is that $|\langle A\rangle_{|\psi\rangle}| \leq \|A\|$. Note also that by the spectral theorem we also have that any observable A can be diagonalized with an ONB $(|e_{j,\alpha}\rangle)_{j,\alpha}$, that is $A = \sum_{j,\alpha} \lambda_j |e_{j,\alpha}\rangle \langle e_{j,\alpha}|$. We can then rewrite the expected value of the observable as

$$\langle A\rangle_{|\psi\rangle} = \sum_{j,\alpha} \lambda_j \langle\psi|e_{j,\alpha}\rangle \langle e_{j,\alpha}|\psi\rangle = \sum_{j,\alpha} \lambda_j |\langle e_{j,\alpha}|\psi\rangle|^2$$

If we conseder as an observable the identity operator (which we will write as $\mathbb{1}$ but is also often referred to as Id or $\mathbf{1}$) we get

$$\langle \mathbb{1}\rangle_{|\psi\rangle} = \sum_\alpha 1|\langle e_\alpha|\psi\rangle|^2 = \|\psi\|^2 = 1$$

Note also that for A, B observables and $\lambda \in \mathbb{R}$, $\langle A + \lambda B\rangle_{|\psi\rangle} = \langle A\rangle_{|\psi\rangle} + \lambda \langle B\rangle_{|\psi\rangle}$

Definition 2.1.3 – Postulate 2

Suppose that a system is in state $|\psi\rangle \in \mathbb{H}$ and let A be an observable. The possible outcomes of measuring the observable A are all the elements of $\sigma(A)$.

For any possible result $\lambda \in \sigma(A)$ the probability of measuring λ is

$$\mathbb{P}_\psi(\lambda) = \|P_\lambda \psi\|^2$$

where the operator $P_\lambda : \mathbb{H} \to \mathbb{H}$ denotes the orthogonal projection on $\text{Eig}_A(\lambda)$ where

$$\text{Eig}_A(\lambda) = \{|\varphi\rangle \in \mathbb{H} |\ |A\varphi\rangle = \lambda |\varphi\rangle\}$$

Remark 2.1.1. Note that P_λ is also an observable. We can rewrite the previous definition as

$$\begin{aligned}
\mathbb{P}_\psi(\lambda) = \|P_\lambda \psi\|^2 &= \langle P_\lambda \psi | P_\lambda \psi \rangle = && (P_\lambda^* = P_\lambda) \\
&= \langle \psi | P_\lambda P_\lambda \psi \rangle = && (P_\lambda^2 = P_\lambda) \\
&= \langle \psi | P_\lambda \psi \rangle = \\
&= \langle P_\lambda \rangle_{|\psi\rangle}
\end{aligned}$$

We can now see that two different states $|\psi\rangle, |\varphi\rangle$ can be "physically" the same, in that they can't be physically distinguished by measurements. Two systems are physically indistinguishable if $|\varphi\rangle = e^{i\alpha} |\psi\rangle$ for some $\alpha \in \mathbb{R}$, called phase. The reason why they are indistinguishable is that there is no observable A that can tell them apart, as two states that differ only by a phase give the same expected value:

Proof.
[$\Rightarrow$] $\langle A \rangle_{|\psi\rangle} = \langle \psi | A\psi \rangle = e^{-i\alpha} \langle \varphi | A\varphi \rangle e^{i\alpha} = \langle \varphi | A\varphi \rangle = \langle A \rangle_{|\varphi\rangle}$
[$\Leftarrow$] left as an exercise to the reader □

Definition 2.1.4 – Ray

We call a ray the set $R_\psi = \{e^{i\alpha} |\psi\rangle \mid \alpha \in \mathbb{R}\}$, that is the states physically indistinguishible from $|\psi\rangle$, or more abstractly, the state $|\psi\rangle$ up to a phase

Definition 2.1.5 – Superpositon

Let $|\psi\rangle, |\varphi\rangle$ be states and $a, b \in \mathbb{C}$ such that $a|\psi\rangle + b|\varphi\rangle$ is also a state (ie such that $\|a|\psi\rangle + b|\varphi\rangle\| = 1$).
Such a state is called a quantum superposition of $|\psi\rangle$ and $|\varphi\rangle$

An important example of quantum superposition is the case where $\langle \psi | \varphi \rangle = 0$, where we get that a sufficient condition to obtain a superposition is $|a|^2 + |b|^2 = 1$.

For example $\frac{1}{\sqrt{2}} |\psi\rangle + \frac{1}{\sqrt{2}} |\varphi\rangle$ is an example of a superposition. If the correspondence between states and probabilities where to behave nicely, given that this state is a "convex combination" (keep in mind that we redefined the expected value to use the square root of the probabilities) of the states $|\psi\rangle$ and $|\varphi\rangle$, one would expect that this state would influence the behaviour of observables in a convex combination way. Indeed, for classical probabilities, it holds:

$$\mathbb{E}_{\alpha p_1+(1-\alpha)p_2}[f] = \alpha\mathbb{E}_{p_1}[f] + (1-\alpha)\mathbb{E}_{p_2}[f]$$

Sadly, it isn't always as straightforward. Consider the case of a more generic superposition $\sqrt{\alpha}\,|\psi\rangle + \sqrt{1-\alpha}\,|\varphi\rangle$. Then for an observable A we get

$$\begin{aligned}\langle A\rangle_{\sqrt{\alpha}|\psi\rangle+\sqrt{1-\alpha}|\varphi\rangle} &= \langle\sqrt{\alpha}\psi + \sqrt{1-\alpha}\varphi|A(\sqrt{\alpha}\psi + \sqrt{1-\alpha}\varphi)\rangle = \\ &= \alpha\,\langle\psi|A\psi\rangle + (1-\alpha)\,\langle\varphi|A\varphi\rangle + \sqrt{\alpha(1-\alpha)}(\langle\psi|A\varphi\rangle + \overline{\langle\psi|A\varphi\rangle}) = \\ &= \alpha\,\langle\psi|A\psi\rangle + (1-\alpha)\,\langle\varphi|A\varphi\rangle + \underbrace{2\sqrt{\alpha(1-\alpha)}\text{Re}(\langle\psi|A\varphi\rangle)}_{\text{interference term}}\end{aligned}$$

This goes to show that a quantum superpositon state isn't just a state composed of a set of states where each has a certain probability of occurring, but something more complex. One can also notice that the interference term appeared because we required no additional hypothesis on $|\psi\rangle$, $|\varphi\rangle$ and A. In the case where we're able to decompose the current state as a superposition states of an ONB that diagonalizes A (which is also useful to calculate the probability of each possible outcome, since this procedure uses a projection on eigenspaces) then the interference term disappears, because if $|\psi\rangle$ and $|\varphi\rangle$ are orthogonal eigenstates of A then $|\psi\rangle$ and $|A\varphi\rangle$ remain orthogonal and the interference term vanishes.

Proposition 2.1.6

Given a system $\mathbb{H}$ prepared in state $|\psi\rangle \in \mathbb{H}$ and given another state vector $|\varphi\rangle \in \mathbb{H}$, we can check if the system is in state $|\varphi\rangle$ and the probability of this occurring is $|\langle\varphi|\psi\rangle|^2$

Proof. The observable we measure when querying if the system is in state $|\varphi\rangle$ is the orthogonal projection $A = |\varphi\rangle\langle\varphi|$, with eigenvalue 1 and 0 respectively corresponding to the system being and not being in state $|\varphi\rangle$. Note that being $\text{Eig}_A(1) = \text{Span}(|\varphi\rangle)$, P_1 from the definition of Postulate 2 coincides with A, and we get

$$\begin{aligned}\mathbb{P}_\psi(1) = & & \text{by definition}\\ = \|P_1(\psi)\|^2 = & & \text{by } P_1 = A\\ = \|A\psi\|^2 = & &\\ = \|\,|\varphi\rangle\,\langle\varphi|\psi\rangle\,\|^2 = & &\\ = |\,\langle\varphi|\psi\rangle\,|^2\underbrace{\|\psi\|^2}_{=1} = & &\\ = |\,\langle\varphi|\psi\rangle\,|^2 & &\end{aligned}$$

□

We've seen the quantum version of the expectation. We'll introduce now the quantum version of the standard deviation

Definition 2.1.7 – Uncertainty

The uncertainty of an observable A in a state vector $|\psi\rangle$ is defined as

$$\Delta_\psi(A) = \sqrt{\left\langle (A - \langle A\rangle_{|\psi\rangle} \, \mathsf{Id})^2 \right\rangle_{|\psi\rangle}}$$

The uncertainty can be rewritten as

$$\begin{aligned}\Delta_\psi(A) &= \sqrt{\left\langle (A - \langle A\rangle_{|\psi\rangle} \, \mathsf{Id})^2 \right\rangle_{|\psi\rangle}} = \\ &= \sqrt{\left\langle \psi \middle| (A - \langle A\rangle_{|\psi\rangle} \, \mathsf{Id})^2 \psi \right\rangle} = \\ &= \sqrt{\left\langle (A - \langle A\rangle_{|\psi\rangle} \, \mathsf{Id})\psi \middle| (A - \langle A\rangle_{|\psi\rangle} \, \mathsf{Id})\psi \right\rangle} = \\ &= \|(A - \langle A\rangle_{|\psi\rangle} \, \mathsf{Id})\psi\|\end{aligned}$$

An observable is sharp on a state $|\psi\rangle$ id $\Delta_\psi(A) = 0$

Proposition 2.1.8

A is sharp on $|\psi\rangle$ if and only if $|\psi\rangle$ is an eigenvector of A

Proof.

$$\begin{aligned}\Delta_\psi(A) = 0 &\iff \|(A - \langle A\rangle_{|\psi\rangle} \, \mathsf{Id})\psi\| = 0 \\ &\iff A\,|\psi\rangle = \langle A\rangle_{|\psi\rangle}\,|\psi\rangle\end{aligned}$$

□

Definition 2.1.9

A, B observables are compatible if $[A, B] = 0$. Compatible observables can be "measured at the same time"

The meaning of "can be measured at the same time" is given by the following result

Proposition 2.1.10 – Heisenberg's uncertainty principle

If A and B are not compatible, then $\forall\,|\psi\rangle$ state it holds

$$\langle i[A, B]\rangle_{|\psi\rangle} \leq 2\Delta_\psi(A)\Delta_\psi(B)$$

In other words, we get a lower bound on the product of the standard deviation of measuring A and B, which means we can't measure them both precisely at the same time.

Proof. Observe that $\forall K : \mathbb{H} \to \mathbb{H}$ operator we can get K^*K which is a non-negative operator, because $\langle \psi | K^* K \psi \rangle = \langle K\psi | K\psi \rangle = \|K\psi\|^2 \geq 0$.

Pick $K = A - iB$ which gives $K^* = A + iB$. Then we have

$$\begin{aligned} K^*K &= (A+iB)(A-iB) = \\ &= A^2 - iAB + iBA + B^2 = \\ &= A^2 + B^2 - i[A,B] \end{aligned}$$

and we get

$$0 \leq \langle K^*K \rangle_{|\psi\rangle} = \left\langle A^2 \right\rangle_{|\psi\rangle} + \left\langle B^2 \right\rangle_{|\psi\rangle} - \langle i[A,B] \rangle_{|\psi\rangle}$$

which implies

$$\langle i[A,B] \rangle_{|\psi\rangle} \leq \left\langle A^2 \right\rangle_{|\psi\rangle} + \left\langle B^2 \right\rangle_{|\psi\rangle}$$

Suppose we substitute A with $A - \mathsf{Id} \langle A \rangle_{|\psi\rangle}$. The commutator would become

$$[A - \mathsf{Id} \langle A \rangle_{|\psi\rangle}, B] = [A,B] - \langle A \rangle_{|\psi\rangle} \underbrace{[\mathsf{Id}, B]}_{=0}$$

so the left-hand side does not change, while on the right-handside, $\left\langle A^2 \right\rangle_{|\psi\rangle}$ would become $\Delta_\psi(A)^2$. By applying the same substitution with B we get

$$\langle i[A,B] \rangle_{|\psi\rangle} \leq \Delta_\psi(A)^2 + \Delta_\psi(B)^2$$

We can parametrize this inequality by replacing A with λA and B with $\frac{B}{\lambda}$ for some variable $\lambda > 0$, which yields the inequality

$$\langle i[A,B] \rangle_{|\psi\rangle} \leq \lambda^2 \Delta_\psi(A)^2 + \frac{1}{\lambda^2} \Delta_\psi(B)^2$$

If we minimize the right-hand side with respect to λ we find that the minimum is reached for $\lambda^2 = \frac{\Delta_\psi(B)}{\Delta_\psi(A)}$, which when substituted gives the thesis □

Note that in the premise of Heisenberg's uncertainty principle we required that A and B were not compatible, though we didn't use it in the proof. Indeed, the inequality also holds if A and B are compatible, but it becomes much less interesting. Consider the following result:

Theorem 2.1.11

If $[A,B] = 0$ then there exists $(e_j)_j$ ONB such that A and B are both diagonal

In the light of this result we gat that if A and B are compatible then they can be simultaneously diagonalized. Suppose $|\psi\rangle$ was an eigenstate of both A and B. Then the inequality becomes an equality, and more precisely

$$0 = \langle i[A,B] \rangle_{|\psi\rangle} \leq 2\Delta_\psi(A)\Delta_\psi(B) = 0 \cdot 0$$

where the first equality holds because $[A,B] = 0$

Example 2.1.12. Consider Heisenberg's inequality on the space $\mathbb{H} = L^2(\mathbb{R})$. We can define a position operator Q such that $(Q\psi)(x) = x \cdot \psi(x)$, and a momentum operator P such that $(P\psi)(x) = -i\frac{d}{dx}\psi(x)$.

It is possible to prove that if $\psi \in \mathcal{C}_c^1(R)$ then

$$\begin{aligned}[Q, P]\psi(x) &= QP\psi - PQ\psi = \\ &= -ix\frac{d}{dx}\psi(x) + i\frac{d}{dx}(x\psi(x)) = \\ &= i\psi(x) = \\ &= i(\mathsf{Id}\psi)(x)\end{aligned}$$

Following the process in the Heisenberg's inequality proof, we get

$$\frac{\langle\psi|\psi\rangle}{2} = \frac{1}{2} \leq \Delta_\psi(P)\Delta_\psi(Q)$$

Which proves the more commonly known version of Heisenberg's uncertainty principle, that states that it's impossible to measure precisely (in our terms, sharply) both position and momentum at the same time

Definition 2.1.13 – Postulate 3

If a system is in state $|\psi\rangle$ and we measure an observable A with outcome $\lambda \in \sigma(A)$, then the state after the measurement is described by $\frac{|P_\lambda\psi\rangle}{\|P_\lambda\psi\|}$

Definition 2.1.14 – Postulate 4

If the system is closed (ie isolated) the evolution of a state $|\psi_t\rangle$ from a time t_0 to a time t_1 is described by a unitary $U(t_0, t_1)$ such that

$$|\psi_{t_1}\rangle = U(t_0, t_1)\,|\psi_{t_0}\rangle$$

Postulate 4 gives us a restriction on the algorithms that we can build. Isolated quantum system can evolve only according to unitary operators, which means that we cannot just apply any arbitrary operator. One interesting consequence is that, because unitary operators are necessarily invertible, quantum algorithms (composed only of quantum steps and no classical steps, that is "pure" quantum algorithms) will always be invertible

Definition 2.1.15 – Hamiltonian

An observable H is the Hamiltonian of a system if $U(t_0, t_1) = e^{-i(t_1 - t_0)H}$

If we assume that H does not depend on t_0, t_1, we can write $|\psi_t\rangle = e^{-itH}\,|\psi_0\rangle$, and by deriving we get

$$\frac{\partial}{\partial t}\,|\psi_t\rangle = -iH\,|\psi_t\rangle$$

which is a different way of writing Schrödinger's equation.

2.2 Mixed states

Given $|\varphi_0\rangle, |\varphi_1\rangle$ states, we called (under certain hypotesis) $|\psi\rangle = a\,|\varphi_0\rangle + b\,|\varphi_1\rangle$ a quantum superposition, and we saw that some interference showed up in the expected value. We will now define a type of state where this interference never shows up, an behaves more like a classical probability upon possible states. For this we will introduce the density operator.

Definition 2.2.1 – Mixed states

A mixed state on a system $\mathbb{H}$ is described by any density operator $\rho : \mathbb{H} \to \mathbb{H}$ such that

- ρ is self-adjointed
- ρ is non-negative
- ρ has unit trace

We call $D(\mathbb{H}) = \{\rho : \mathbb{H} \to \mathbb{H} \,|\, \rho^* = \rho,\ \rho \geq 0,\ \mathrm{tr}\,(\rho) = 1\}$ the set of all the density operators on $\mathbb{H}$

Remark 2.2.1. The following properties on density operators hold:

- $D(\mathbb{H})$ is a convex set
- If $|\psi\rangle$ is a pure state, we can build $\rho = |\psi\rangle\langle\psi| \in D(\mathbb{H})$ which shows that every pure state is (as in, it can be represented as) a mixed state. There are some mixed states that aren't pure. It holds that $|\psi\rangle\langle\psi| \in D(\mathbb{H})$ because
  - $\langle\varphi|\rho\varphi\rangle = \langle\varphi|\psi\rangle\,\langle\psi|\varphi\rangle = \langle\psi|\varphi\rangle\,\langle\varphi|\psi\rangle = \langle\rho\varphi|\varphi\rangle$
  - $\mathrm{tr}\,(\rho) = \|\psi\|^2 = 1$
- If $|\varphi\rangle \in R_{|\psi\rangle}$, then $|\varphi\rangle\langle\varphi| = |\psi\rangle\langle\psi|$
- If U is unitary, then $\forall\rho \in D(\mathbb{H})$, $U\rho U^*$ is also a state

Theorem 2.2.2

Let $\dim\mathbb{H} = n$. Given $\rho \in D(\mathbb{H})$

1. $\exists(p_i)_i$ with $p_i \in [0,1]$, $\sum_i p_i = 1$ and $\exists(|\psi_i\rangle)_i$ ONB of $\mathbb{H}$ such that $\rho = \sum_i p_i\,|\psi_i\rangle\langle\psi_i|$
2. $\rho^2 \leq \rho$ as quadratic forms. The equality holds if and only if ρ is pure, that is $\rho = |\psi\rangle\langle\psi|$ for some $|\psi\rangle$
3. $\|\rho\| \leq 1$, and the equality holds if and only if ρ is pure

Proof. 1. By applying the spectral theorem to ρ we get $\rho = \sum_j \lambda_j \sum_\alpha |\varphi_{j,\alpha}\rangle\,\langle\varphi_{j,\alpha}|$ with $\sigma(\rho) = \{\lambda_j\}$.

Up to renumbering we can send $|\varphi_{j,\alpha}\rangle \mapsto |\psi_i\rangle$, $\lambda_j \mapsto p_i$ and get

$$1 = \mathrm{tr}\,(\rho) = \sum_j \lambda_j m_j = \sum_i p_i \qquad m_j = \text{multiplicity of } \lambda_j$$

Also $p_i \geq 0$ so $p_i \in [0, 1]$

2. Since $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ we have

$$\begin{aligned}\rho^2 &= \left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right)\left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right) = \qquad \text{because of orthonormality}\\ &= \sum_i p_i^2 |\psi_i\rangle\langle\psi_i| \leq \\ &= \sum_i p_i |\psi_i\rangle\langle\psi_i| = \rho\end{aligned}$$

Formally, $\rho - \rho^2 = \sum_i p_i(1 - p_i) |\psi_i\rangle\langle\psi_i|$ which can be zero if and only if exactly one of the $p_i = 1$ and all the others are zero, that is if and only if ρ is pure

3. Left as an exercise

□

Let's now see the four postulates we saw in the previous section, but with mixed states

Definition 2.2.3 – Postulate 1, with mixed states

If a system is in a mixed state ρ, given an observable A its expected value is $\langle A\rangle_\rho = \mathrm{tr}\,(A\rho) = \sum_i p_i \langle A\rangle_{|\psi_i\rangle}$

Note that by postulate 1, a mixed state on $|\psi_1\rangle$, $|\psi_2\rangle$ and a superposition of $|\psi_1\rangle$, $|\psi_2\rangle$ behave extremely differently, and the mixed state behaves more like a classical probability

Definition 2.2.4 – Postulate 2, with mixed states

If the system is in the mixed state ρ, the probability of observing $\lambda \in \sigma(A)$ is

$$\mathbb{P}_\rho(\lambda) = \mathrm{tr}\,(P_\lambda \rho) = \sum_i p_i \mathbb{P}_{\psi_i}(\lambda)$$

Definition 2.2.5 – Postulate 3, with mixed states

After measuring $\lambda \in \sigma(A)$, the state ρ collapses to

$$\frac{P_\lambda \rho P_\lambda}{\mathbb{P}_\rho(\lambda)} \in D(\mathbb{H})$$

Remark 2.2.2. Say that A is measured but the outcome is not given to us. We can still describe the system as

$$\tilde{\rho} = \sum_{\lambda \in \sigma(A)} \mathbb{P}_\rho(\lambda) \left(\frac{P_\lambda \rho P_\lambda}{\mathbb{P}_\rho(\lambda)} \right)$$

Definition 2.2.6 – Postulate 4, with mixed states

If the system is closed, it evolves from time t_0 to time t_1 according to an unitary operator $U(t_0, t_1)$ such that $\rho_{t_0} \mapsto \rho_{t_1} = U(t_0, t_1)\rho_{t_0} U(t_0, t_1)$

The Schrödinger equation associated to H is

$$i\partial_t \rho_t = H\rho_t - \rho_t H = [H, \rho_t]$$

The uncertainty of A over ρ is

$$\Delta_\rho(A) = \sqrt{\left\langle (A - \langle A \rangle_\rho)^2 \right\rangle_\rho}$$

If we represent $\rho \in D(\mathbb{H})$ as $\rho = \sum_{i=1}^m q_i |\varphi_i\rangle\langle\varphi_i|$, with $|\varphi_i\rangle$ state vectors (not necessarily orthogonal), and $\rho = \sum_{i=1}^m p_i |\psi_i\rangle\langle\psi_i|$ with $(|\psi_i\rangle)_{i=1}^n$ ONB, we can find a relation between the two decompositions. It holds that $m \geq n$ and exists $U \in \mathbb{C}^{m \times m}$ unitary such that

$$\sqrt{q_i}\,|\varphi_i\rangle = \sum_{j=1}^n U_{ij} \sqrt{p_j}\,|\psi_j\rangle$$

Note that despite U being of size $m \times m$ (which it has to be in order to be unitary), we only use a submatrix of size $m \times n$, so U is not unique

Hint of a proof. Start with $U_{ij} = \sqrt{\frac{q_i}{p_j}} \langle \varphi_i | \psi_j \rangle$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$ and complete to $m \times m$ unitary □

Remark 2.2.3. If $\rho, \rho' \in D(\mathbb{H})$ are such that $\forall A$ observable it holds$\langle A \rangle_\rho = \langle A \rangle_{\rho'}$, then $\rho = \rho'$

2.3 Spin

We know that electrons have an intrinsic property called spin. This is not an actual spinning motion, but it's useful to think of it like that, and more precisely as an axis of rotation.

We identify Spin with a vector in $\mathbb{R}^3$. We're interested in being able to measure the spin's component along three axis with three observables that we will call S_x, S_y and S_z. From now on, we will consider the Hilbert space $\mathbb{H} \cong \mathbb{C}^2$, that is $n = 2$.

We identify an ONB for $\mathbb{H}$ as $|{\uparrow_z}\rangle$, $|{\downarrow_z}\rangle$ (also called up and down). We will also use the notation $|0\rangle$, $|1\rangle$

Definition 2.3.1 – Pauli matrices

The Pauli matrices are

$$\sigma_x = \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

We can build the operators $S_x = \frac{1}{2}\sigma_x$, $S_y = \frac{1}{2}\sigma_y$ and $S_z = \frac{1}{2}\sigma_z$

Note that $|{\uparrow_z}\rangle$ and $|{\downarrow_z}\rangle$ are eigenvectors for the observable S_z with eigenvalues $\frac{1}{2}$ and $-\frac{1}{2}$ respectively, which explains the notation

Proposition 2.3.2

The following properties hold:

1. $\sigma_j^2 = \mathbb{1}$ and if $j \neq k$ then $\sigma_j \sigma_k = i\varepsilon_{jkl}\sigma_l$ where

   - l is the missing index if $j \neq k$ and any value (it doesn't matter) if $j = k$
   - $\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1$, $\varepsilon_{321} = \varepsilon_{213} = \varepsilon_{132} = -1$ and $\varepsilon_{jkl} = 0$ for any other combination of j, k, l

   In one formula we can write $\sigma_j \sigma_k = \delta_{jk}\mathbb{1} + i\varepsilon_{jkl}\sigma_l$

2. $[\sigma_j, \sigma_k] = \sigma_j\sigma_k - \sigma_k\sigma_j = 2i\varepsilon_{jkl}\sigma_l$
3. $\{\sigma_j, \sigma_k\} = \sigma_j\sigma_k + \sigma_k\sigma_j = 2\delta_{jk}\mathbb{1}$
4. σ_x, σ_y and σ_z are unitary

Example 2.3.3. Suppose a system is in state $|0\rangle = |{\uparrow_z}\rangle$. What would happen after we measure $\sigma_z = 2S_z$? From the postulates we expect that:

- the measurement is sharp
- the expected value is 1
- the system stays in state $|0\rangle$ after the measurement

Let's manually check.

$\langle \sigma_z \rangle_{|0\rangle} = \langle 0 | \sigma_z 0 \rangle = (1, 0) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1$, so the expected value is indeed 1.

For the uncertainty we get

$$\begin{aligned}\sigma_z - \langle\sigma_z\rangle_{|0\rangle}\mathbb{1} &= \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} - \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = \begin{pmatrix}0 & 0\\ 0 & -2\end{pmatrix}\\ \Rightarrow \left(\sigma_z - \langle\sigma_z\rangle_{|0\rangle}\mathbb{1}\right)^2 &= \begin{pmatrix}0 & 0\\ 0 & 4\end{pmatrix}\\ \Rightarrow \left\langle 0 \middle| \left(\sigma_z - \langle\sigma_z\rangle_{|0\rangle}\mathbb{1}\right)^2 0\right\rangle &= (1,0)\begin{pmatrix}0 & 0\\ 0 & 4\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix} = 0\end{aligned}$$

so the measurement is indeed sharp.
Finally, because $|0\rangle$ is an eigenstate, the system remains in that state.

Suppose we started with state $|1\rangle$ instead of state $|0\rangle$. The only thing that would change is the expected value, which would now be -1.

Let's say that starting with state $|0\rangle$, we choose to measure σ_x instead of σ_z. Then, for the expected value, we would get $\langle\sigma_x\rangle_{|0\rangle} = (1,0)\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix} = 0$.
For the uncertainty we have $\left(\sigma_x - \langle\sigma_x\rangle_{|0\rangle}\mathbb{1}\right)^2 = \sigma_x^2 = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ which implies $\left\langle 0 \middle| \left(\sigma_x - \langle\sigma_x\rangle_{|0\rangle}\mathbb{1}\right)^2 0\right\rangle = (1,0)\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix} = 1$.
The possible outcomes of this measurement are the eigenvalues of σ_x which are ± 1, each of which has a probability of occurring which we can compute via projection on the eigenspaces.
The eigenstates of σ_x are $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix} = |\uparrow_x\rangle$ and $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -1\end{pmatrix} = |\downarrow_x\rangle$ respectively for the eigenvalues 1 and -1. We can write these states with respect to the ONB $\{|0\rangle, |1\rangle\}$ as

$$|\uparrow_x\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle, \qquad |\downarrow_x\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$$

and equivalently

$$|0\rangle = \frac{1}{\sqrt{2}}|\uparrow_x\rangle + \frac{1}{\sqrt{2}}|\downarrow_x\rangle, \qquad |1\rangle = \frac{1}{\sqrt{2}}|\uparrow_x\rangle - \frac{1}{\sqrt{2}}|\downarrow_x\rangle$$

So to calculate the probability of measuring 1 from the state $|0\rangle$ we take the coefficient of the eigenstate corresponding to 1 in the expression for $|0\rangle$ and square its modulus. The same goes for the probability of measuring -1.

Suppose we decide to measure σ_z after we just measured σ_x. Let's assume the system started in state $|0\rangle$ and when measuring σ_x we got the result 1, so now the system is in the state $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$.
If we measure σ_z in the current state, we have a probability of $\frac{1}{2}$ of measuring $|0\rangle$ and a probability of $\frac{1}{2}$ of measuring $|1\rangle$, while previously the measurement was sharp!
Imagine we start in state $|0\rangle$, measure σ_z (and obtain $|0\rangle$), then measure σ_x and then again σ_z. Because after σ_x we have that σ_z is not sharp anymore, the act of measuring σ_x possibly

changed a property that we already "knew" (as in, measured).

Notice that σ_z and σ_x do not commute, so they are not compatible and cannot be measured at the same time

3 Quantum Systems

3.1 Qubits

A classical bit can be imagined as a set $\{0, 1\}$ where the possible states are 0 and 1. The equivalent of a bit in a quantum system is a qubit, of which we give now a definition

Definition 3.1.1 – Qubit

We identify a qubit as a 2-dimensional quantum system defined by a Hilbert space $\mathbb{H}$ with ONB $\{|0\rangle, |1\rangle\}$ and an observable σ_z with $|0\rangle$ and $|1\rangle$ as eigenstates, respectively with eigenvalues 1 and -1.

The possible states $|0\rangle$ and $|1\rangle$ correspond to the classical states 0 and 1, but w also have states of the form

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle\,, \qquad \text{where } \|a\|^2 + \|b\|^2 = 1,\ a, b \in \mathbb{C}$$

For this reason, a qubit contains much more information than a classical bit, but after a measurement this additional information is lost.
Measuring a qubit means measuring σ_z, yielding either 1 or -1 after which the state will collapse respectively in $|0\rangle$ or $|1\rangle$

3.1.1 Bloch sphere representation

Since $\|a\|^2 + \|b\|^2 = 1$, there exists angle α, β, θ such that

$$a = e^{i\alpha} \cos\frac{\theta}{2}, \qquad b = e^{i\beta} \sin\frac{\theta}{2}$$

which means that up to a global phase, we can write $|\psi\rangle = e^{-i\frac{\varphi}{2}} \cos\frac{\theta}{2}\,|0\rangle + e^{i\frac{\varphi}{2}} \sin\frac{\theta}{2}\,|1\rangle$ for some angles $0 \leq \theta \leq \pi$ and $0 \leq \varphi \leq 2\pi$. Note that since we're interested in states up to a global phase, we're free to write $|\psi\rangle = \cos\frac{\theta}{2}\,|0\rangle + e^{i\theta} \sin\frac{\varphi}{2}\,|1\rangle$ and "put all the phase on $|1\rangle$".

We can uniquely identify (again up to a global phase) $|\psi\rangle = |\psi\rangle_{\varphi,\theta}$ as the point $\begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix}$ in $S^3 \subset \mathbb{R}^3$. This representation is called the Bloch sphere representation

Definition 3.1.2

Given a qubit state $|\psi\rangle$ we define $\hat{n}_{|\psi\rangle} = \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix}$ where θ, φ are the parameters that represent $|\psi\rangle$ in Bloch sphere representation

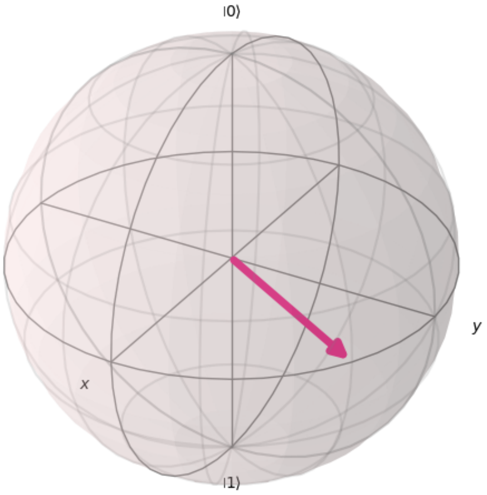


Figure 3.1: Bloch sphere representation for the state corresponding to $\theta = \frac{\pi}{2}$, $\varphi = \frac{\pi}{3}$

Note that this is not a bijection because for $\theta = 0$ every choice of φ would end up in $|0\rangle$, and for $\theta = \pi$ every choice of φ would end up in $|1\rangle$.

Fix a qubit state $|\psi\rangle$. We would like to build an observable which has $|\psi\rangle$ as an eigenvector.

Definition 3.1.3

Let $a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \in \mathbb{R}^3$. We define

$$a \cdot \sigma = \sum_{k=1}^{3} a_k \sigma_k = a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3$$

With $a \in \mathbb{R}^3$ we identify the linear combination of the Pauli matrices given by

$$a \cdot \sigma = \begin{pmatrix} a_3 & a_1 - ia_2 \\ a_1 + ia_2 & -a_3 \end{pmatrix}$$

Remark 3.1.1. It holds

$$\begin{aligned}(a\cdot\sigma)(b\cdot\sigma) =&(a_1\sigma_1+a_2\sigma_2+a_3\sigma_3)(b_1\sigma_1+b_2\sigma_2+a_3\sigma_3)=\\ =&(a_1b_1+a_2b_2+a_3b_3)\mathbb{1}+a_1b_2\sigma_1\sigma_2+a_1b_3\sigma_1\sigma_3+\\ &+a_2b_1\sigma_2\sigma_1+a_2b_3\sigma_2\sigma_3+a_3b_1\sigma_3\sigma_1+a_3b_2\sigma_3\sigma_2=\\ =&(a\cdot b)\mathbb{1}+i(a_1b_2\sigma_3-a_1b_3\sigma_2-a_2b_1\sigma_3+a_2b_3\sigma_1+a_3b_1\sigma_2a_3b_2\sigma_1)=\\ =&(a\cdot b)\mathbb{1}+i\left[(a_2b_3-a_3b_2)\sigma_1+(a_3b_1-a_1b_3)\sigma_2+(a_1b_2a_2b_1)\sigma_3\right]=\\ =&(a\cdot b)\mathbb{1}+i(a\times b)\cdot\sigma\end{aligned}$$

Notice that we get

$$\begin{aligned}\hat{n}_{|\psi\rangle}\cdot\sigma &= \begin{pmatrix}\cos\theta & \sin\theta\cos\varphi-i\sin\theta\sin\varphi\\ \sin\theta\cos\varphi+i\sin\theta\sin\varphi & -\cos\theta\end{pmatrix}=\\ &=\begin{pmatrix}\cos\theta & e^{-i\varphi}\sin\theta\\ e^{i\varphi}\sin\theta & -\cos\theta\end{pmatrix}\end{aligned}$$

The operator $\hat{n}_{|\psi\rangle}\cdot\sigma$ is exactly the operator we were looking for, as $|\psi\rangle$ is an eigenvector for $\hat{n}_{|\psi\rangle}\cdot\sigma$, because:

$$\begin{aligned}\begin{pmatrix}\cos\theta & e^{-i\varphi}\sin\theta\\ e^{i\varphi}\sin\theta & -\cos\theta\end{pmatrix}\begin{pmatrix}e^{-i\frac{\varphi}{2}}\cos\frac{\theta}{2}\\ e^{i\frac{\varphi}{2}}\sin\frac{\theta}{2}\end{pmatrix} &= \begin{pmatrix}e^{-i\frac{\varphi}{2}}\cos\theta\cos\frac{\theta}{2}+e^{-i\frac{\varphi}{2}}\sin\theta\sin\frac{\theta}{2}\\ e^{i\frac{\varphi}{2}}\sin\theta\cos\frac{\theta}{2}-e^{i\frac{\varphi}{2}}\cos\theta\sin\frac{\theta}{2}\end{pmatrix}=\\ &=\begin{pmatrix}e^{-i\frac{\varphi}{2}}(\cos\theta\cos\frac{\theta}{2}+\sin\theta\sin\frac{\theta}{2})\\ e^{i\frac{\varphi}{2}}(\sin\theta\cos\frac{\theta}{2}-\cos\theta\sin\frac{\theta}{2})\end{pmatrix}=\\ &=\begin{pmatrix}e^{-i\frac{\varphi}{2}}\cos\left(\theta-\frac{\theta}{2}\right)\\ e^{i\frac{\varphi}{2}}\sin\left(\theta-\frac{\theta}{2}\right)\end{pmatrix}=\\ &=|\psi\rangle\end{aligned}$$

Similarly, $\left|\downarrow_{\hat{n}_{|\psi\rangle}}\right\rangle=\begin{pmatrix}e^{-i\frac{\varphi}{2}}\cos\frac{\theta}{2}\\ -e^{i\frac{\varphi}{2}}\sin\frac{\theta}{2}\end{pmatrix}$ is an eigenvector with eigenvalue -1

Now that we've seen pure states in the context of qubits, we'll now discuss mixed states.

A mixed state is described by a density operator, that is an operator $\rho:\mathbb{H}\to\mathbb{H}$ that satisfies

1. $\rho=\rho^*$
2. tr $(\rho)=1$
3. $\rho\geq 0$

In dimension $n=2$, a density operator is described by a matrix $\rho=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ satisfying (1.), (2.) and (3.). As (1.) implies $a,d\in\mathbb{R}$, $b=\overline{c}$ and (2.) implies $a+d=1$, we can describe ρ with three real parameters $x_1,x_2,x_3\in\mathbb{R}$ by

$$a=\frac{1+x_3}{2},\qquad d=\frac{1-x_3}{2},\qquad b=\frac{x_1+ix_2}{2},\qquad c=\frac{x_1-ix_2}{2}$$

which yields

$$\rho = \frac{1}{2}\begin{pmatrix} 1+x_3 & x_1+ix_2 \\ x_1-ix_2 & 1-x_3 \end{pmatrix}$$

To satisfy (3.) we need to find the eigenvalues of ρ, so

$$\begin{gathered}(1+x_3-2\lambda)(1-x_3-2\lambda)-(x_1-ix_2)(x_1+ix_2)=0 \\ \Downarrow \\ (1-2\lambda)^2-x_3^2=x_1^2+x_2^2 \\ \Downarrow \\ (1-2\lambda)^2=\|x\|_2^2 \\ \Downarrow \\ \lambda=\frac{1\pm\|x\|_2^2}{2}\end{gathered}$$

so we get $\rho \geq 0 \iff \lambda \geq 0 \iff \|x\|_2 \leq 1$. This means that we can represent any mixed state ρ in dimension $n=2$ with a vector $x \in \mathbb{R}^3$ inside the unitary disk. This gives a representation corresponding to the Bloch sphere, and indeed we have the following remarks

Remark 3.1.2.

$$\begin{aligned}\rho &= \frac{1}{2}\left[\begin{pmatrix}1&0\\0&1\end{pmatrix}+x_1\begin{pmatrix}0&1\\1&0\end{pmatrix}+x_2\begin{pmatrix}0&-i\\i&0\end{pmatrix}+x_3\begin{pmatrix}1&0\\0&-1\end{pmatrix}\right]= \\ &=\frac{1}{2}\left[\mathbb{1}+x_1\sigma_x+x_2\sigma_y+x_3\sigma_z\right]= \\ &=\frac{1}{2}[\mathbb{1}+x\cdot\sigma]\end{aligned}$$

Remark 3.1.3. ρ is pure if and only if it's represented by a point on the sphere (ie $\|x\|=1$)

Proof. We know that ρ is pure if and only if $\rho^2=\rho$ so

$$\begin{aligned}\rho^2 &= \left(\frac{1}{2}[\mathbb{1}+x\cdot\sigma]\right)= \\ &=\frac{1}{4}(\mathbb{1}+(x\cdot\sigma)(x\cdot\sigma)+2x\cdot\sigma)= \qquad (a\cdot\sigma)(b\cdot\sigma)=(a\cdot b)\mathbb{1}+i(a\times b)\cdot\sigma \\ &=\frac{1}{4}(\mathbb{1}+\|x\|_2^2\mathbb{1}+2x\cdot\sigma)= \\ &=\frac{1}{2}\left(\frac{1+\|x\|_2^2}{2}\mathbb{1}+x\cdot\sigma\right)\end{aligned}$$

so $\rho^2=\rho \iff \|x\|_2=1$ □

Remark 3.1.4. For $j=1,2,3$ it holds $\operatorname{tr}(\rho_x\sigma_j)=x_j$

Proof.

$$\begin{aligned}\text{tr}\left(\rho_x\sigma_j\right) = \text{tr}\left(\frac{1}{2}(\mathbb{1} + x\cdot\sigma)\sigma_j\right) &= \\ &= \frac{1}{2}\text{tr}\left(\sigma_j + (x\cdot\sigma)\sigma_j\right) = \\ &= \frac{1}{2}\underbrace{\text{tr}\left(\sigma_j\right)}_{=0} + \frac{1}{2}\text{tr}\left((x\cdot\sigma)(e_j\cdot\sigma)\right) = \\ &= \frac{1}{2}\text{tr}\left(x\cdot e_j\mathbb{1} + i(x\times e_j)\cdot\sigma\right) = \\ &= \frac{1}{2}\underbrace{\text{tr}\left(x_j\mathbb{1}\right)}_{=2x_j} + \frac{1}{2}i\sum_{k=1}^{3}\underbrace{\text{tr}\left((x\times e_j)_k\sigma_k\right)}_{=0} = \\ &= x_j\end{aligned}$$

□

3.1.2 Operators on qubits

We now want to study unitary operators on $\mathbb{H}$

Example 3.1.4. Let A be an operator on $\mathbb{H}$, then we also have $e^A = \sum_{k=0}^{\infty}\frac{A^k}{k!}$. Suppose that $A^2 = \mathbb{1}$, then $\forall\alpha\in\mathbb{R}$ we have

$$\begin{aligned}e^{i\alpha A} &= \mathbb{1} + i\alpha A - \frac{1}{2}\alpha^2\mathbb{1} - \frac{1}{3!}i\alpha^3 A + \dots \\ &= \cos(\alpha)\mathbb{1} + i\sin(\alpha)A\end{aligned}$$

Definition 3.1.5

Let $\hat{n}\in\mathbb{R}^3$ be a unit vector and $\alpha\in R$. We define the rotation around $\hat{n}$ of an angle $\frac{\alpha}{2}$ as the operator $D_{\hat{n}}(\alpha) = e^{-i\frac{\alpha}{2}\hat{n}\cdot\sigma}$.
This is called the spin operator.

Note that much like in the example, $(\hat{n}\cdot\sigma)^2 = \underbrace{(\hat{n}\cdot\hat{n})}_{=1}\mathbb{1} + i\underbrace{(\hat{n}\times\hat{n})}_{=0}\cdot\sigma = \mathbb{1}$, and we get the following properties

Proposition 3.1.6

1. $D_{\hat{n}}(\alpha) = e^{-i\frac{\alpha}{2}\hat{n}\cdot\sigma} = \cos\left(\frac{\alpha}{2}\right)\mathbb{1} - i\sin\left(\frac{\alpha}{2}\right)\hat{n}\cdot\sigma$
2. $D_{\hat{n}}(\alpha)^* = D_{\hat{n}}(-\alpha)$
3. $D_{\hat{n}}(\alpha)D_{\hat{n}}(\alpha)^* = \mathbb{1}$, so $D_{\hat{n}}(\alpha)$ is unitary

4. $D_{\hat{n}}(\alpha)D_{\hat{n}}(\beta) = D_{\hat{n}}(\alpha+\beta)$, because

$$\begin{aligned} D_{\hat{n}}(\alpha)D_{\hat{n}}(\beta) = &\left[\cos\left(\frac{\alpha}{2}\right)\mathbb{1} - i\sin\left(\frac{\alpha}{2}\right)\hat{n}\cdot\sigma\right]\left[\cos\left(\frac{\beta}{2}\right)\mathbb{1} - i\sin\left(\frac{\beta}{2}\right)\hat{n}\cdot\sigma\right] = \\ &= \cos\left(\frac{\alpha}{2}\right)\cos\left(\frac{\beta}{2}\right)\mathbb{1} - \sin\left(\frac{\beta}{2}\right)\sin\left(\frac{\beta}{2}\right)\mathbb{1} - \\ &\quad - i\left[\cos\left(\frac{\alpha}{2}\right)\sin\left(\frac{\beta}{2}\right) + \sin\left(\frac{\alpha}{2}\right)\cos\left(\frac{\beta}{2}\right)\right]\hat{n}\cdot\sigma = \\ &= \cos\left(\frac{\alpha+\beta}{2}\right)\mathbb{1} - i\sin\left(\frac{\alpha+\beta}{2}\right)\hat{n}\cdot\sigma = \\ &= D_{\hat{n}}(\alpha+\beta) \end{aligned}$$

Lemma 3.1.7

Let U be a unitary operator on $\mathbb{H}$, then $\exists\alpha,\beta,\gamma,\delta\in\mathbb{R}$ such that the matrix of U with respect to the ONB $\{|0\rangle, |1\rangle\}$ is

$$U = e^{i\alpha}\begin{pmatrix} e^{-i\frac{\beta+\delta}{2}}\cos\left(\frac{\gamma}{2}\right) & -e^{-i\frac{\delta-\beta}{2}}\sin\left(\frac{\gamma}{2}\right) \\ e^{i\frac{\beta-\delta}{2}}\sin\left(\frac{\gamma}{2}\right) & e^{i\frac{\beta+\delta}{2}}\cos\left(\frac{\gamma}{2}\right) \end{pmatrix}$$

Hint of a proof. Write $U = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{C}^2$ and impose $UU^* = \mathbb{1}$. This is written in matrix form as

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

which is equivalent to the following system

$$\begin{cases} |a|^2 + |b|^2 = 1 \\ a\overline{c} + b\overline{d} = 0 \\ |c|^2 + |d|^2 = 1 \end{cases}$$

from which is possible to find a set of angles that gives the thesis □

Lemma 3.1.8

Let U be a unitary operator on $\mathbb{H}$, then $\exists\alpha,\beta,\gamma,\delta$ such that

$$U = e^{i\alpha}D_{\hat{z}}(\beta)D_{\hat{y}}(\gamma)D_{\hat{z}}(\delta)$$

that is we can express U as a composition of rotations using only rotations around $\hat{z}$ and $\hat{y}$

Proof. Remember that

$$D_{\hat{z}}(\beta) = \cos\left(\frac{\beta}{2}\right)\mathbb{1} - i\sin\left(\frac{\beta}{2}\right)\sigma_z = \begin{pmatrix} \cos\frac{\beta}{2} - i\sin\frac{\beta}{2} & 0 \\ 0 & \cos\frac{\beta}{2} + i\sin\frac{\beta}{2} \end{pmatrix}$$

$$D_{\hat{y}}(\gamma) = \cos\left(\frac{\gamma}{2}\right)\mathbb{1} - i\sin\left(\frac{\gamma}{2}\right)\sigma_y = \begin{pmatrix} \cos\frac{\beta}{2} & -\sin\frac{\beta}{2} \\ \sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix}$$

If you explicit the product $e^{i\alpha}D_{\hat{z}}(\beta)D_{\hat{y}}(\gamma)D_{\hat{z}}(\delta)$ with the matrices we've just written, you get the representation given by the previous lemma □

Lemma 3.1.9

Let U be a unitary operator on $\mathbb{H}$, then there exists operators A, B, C on $\mathbb{H}$ and $\alpha \in \mathbb{R}$ such that

- $ABC = \mathbb{1}$
- $U = e^{i\alpha}A\sigma_x B\sigma_x C$

Hint of a proof. We know that $U = e^{i\alpha}D_{\hat{z}}(\beta)D_{\hat{y}}(\gamma)D_{\hat{z}}(\delta)$. If we take

$$A = D_{\hat{z}}(\beta)D_{\hat{y}}\left(\frac{\gamma}{2}\right)$$

$$B = D_{\hat{y}}\left(-\frac{\gamma}{2}\right)D_{\hat{z}}\left(-\frac{\beta+\delta}{2}\right)$$

$$C = D_{\hat{z}}\left(\frac{\delta-\beta}{2}\right)$$

It's clear that $ABC = \mathbb{1}$. For the second point we get

$$\begin{aligned} A\sigma_x B\sigma_x C &= D_{\hat{z}}(\beta)D_{\hat{y}}\left(\frac{\gamma}{2}\right)\sigma_x D_{\hat{y}}\left(-\frac{\gamma}{2}\right)D_{\hat{z}}\left(-\frac{\beta+\delta}{2}\right)\sigma_x D_{\hat{z}}\left(\frac{\delta-\beta}{2}\right) = \\ &= D_{\hat{z}}(\beta)D_{\hat{y}}\left(\frac{\gamma}{2}\right)\sigma_x D_{\hat{y}}\left(-\frac{\gamma}{2}\right)\underbrace{\mathbb{1}}_{=\sigma_x\sigma_x} D_{\hat{z}}\left(-\frac{\beta+\delta}{2}\right)\sigma_x D_{\hat{z}}\left(\frac{\delta-\beta}{2}\right) = \\ &= D_{\hat{z}}(\beta)D_{\hat{y}}\left(\frac{\gamma}{2}\right)\underbrace{\sigma_x D_{\hat{y}}\left(-\frac{\gamma}{2}\right)\sigma_x}_{D_{\hat{y}}\left(\frac{\gamma}{2}\right)}\underbrace{\sigma_x D_{\hat{z}}\left(-\frac{\beta+\delta}{2}\right)\sigma_x}_{D_{\hat{z}}\left(\frac{\beta+\delta}{2}\right)} D_{\hat{z}}\left(\frac{\delta-\beta}{2}\right) = \\ &= D_{\hat{z}}(\beta)D_{\hat{y}}(\gamma)D_{\hat{z}}(\delta) \end{aligned}$$

Note that the substitutions on the last step hold specifically because $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and that we're using $D_{\hat{y}}$ and $D_{\hat{z}}$. This would need a rigorous proof which is left as an exercise to the reader, and that's the reason why this is only a hint of a proof □

Lemma 3.1.10

Let U be a unitary operator on $\mathbb{H}$, then there exist $\alpha, \xi \in \mathbb{R}$ angles and $\hat{n} \in \mathbb{R}^3$ unit vector such that $U = e^{i\alpha} D_{\hat{n}}(\xi)$

Hint of a proof. Use the representation from the first lemma and split the matrix as $U = c_0 \mathbb{1} + c_1 \sigma_x + c_2 \sigma_y + c_3 \sigma_z$ and deduce the coordinates of $\hat{n}$ and ξ from $\{c_i\}_{i=0}^3$. This is not a difficult proof, but it is a long one □

Proposition 3.1.11

Let A be an operator on $\mathbb{H}$, then there exist $z_0, z_1, z_2, z_3 \in \mathbb{C}$ such that $A = z_0 \mathbb{1} + z \cdot \sigma$ where $z = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \in \mathbb{C}^3$. If A is also unitary, then $|z_0|^2 + \|z\|_2^2 = 1$

Proof. Write $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} z_0 + z_3 & z_1 - iz_2 \\ z_1 + iz_2 & z_0 - z_3 \end{pmatrix} \in \mathbb{C}^{2 \times 2}$, from which we get

$$z_0 = \frac{a+d}{2}, \qquad z_1 = \frac{b+c}{2}, \qquad z_2 = i\frac{b-c}{2}, \qquad z_3 = \frac{a-d}{2}$$

If A is unitary, then $A = e^{i\alpha} D_{\hat{n}}(\xi)$ from which we can deduce $|z_0|^2 + \|z\|_2^2 = 1$ □

3.1.3 Hadamard operator

Definition 3.1.12 – Hadamard operator

The Hadamard operator is $H = \frac{\sigma_x + \sigma_z}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ as written in matrix form with respect to $\{|0\rangle, |1\rangle\}$

Remark 3.1.5. $H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$ and $H|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$. It also holds $H^2 = \mathbb{1}$.

The decomposition of H in rotations is $H = e^{i\frac{3}{2}\pi} D_{\hat{z}}(0) D_{\hat{y}}\left(\frac{\pi}{2}\right) D_{\hat{z}}(-\pi)$

3.2 Composite systems

If a classical bit can be represented with the set $\{0, 1\}$, where the possible states are 0 and 1, a classical system of bits can be represented with the set $\{0, 1\}^n$, where the possible states are the binary strings of length n.

In a similar fashion, we can combine different qubits (each of which is represented with the Hilbert space $\mathbb{H} \cong \mathbb{C}$) to obtain a collection of qubits, that we will call quantum system. The way we will combine Hilbert spaces to obtain a quantum system is via the tensor product. Suppose we have a qubit that we will call A, identified by a Hilbert space that we will call $\mathbb{H}^A$, and a qubit B identified by its space $\mathbb{H}^B$, then we will call the combined system of the two qubits $\mathbb{H}^{AB} = \mathbb{H}^A \otimes \mathbb{H}^B$, which we will now define formally.

Given a joint probability $\rho \in D(\mathbb{H}^{AB})$ of $\mathbb{H}^{AB}$, we would like to be able to compute the marginal probabilities on $\mathbb{H}^A$ and $\mathbb{H}^B$, $\rho^A \in D(\mathbb{H}^A)$ and $\rho^B \in D(\mathbb{H}^B)$ (called reduced states)

Definition 3.2.1

Given $(\mathbb{H}^A, \langle\cdot|\cdot\rangle^{\mathbb{H}^A})$ and $(\mathbb{H}^B, \langle\cdot|\cdot\rangle^{\mathbb{H}^B})$ two Hilbert spaces with the respective scalar products, we define $\forall\, |\varphi\rangle \in \mathbb{H}^A, |\psi\rangle \in \mathbb{H}^B$ the functional $|\varphi\rangle \otimes |\psi\rangle$ given by

$$\begin{aligned} |\varphi\rangle \otimes |\psi\rangle :\ \mathbb{H}^A \otimes \mathbb{H}^B &\longrightarrow \mathbb{C} \\ (\xi, \eta) &\longmapsto \langle\xi|\varphi\rangle \langle\psi|\eta\rangle \end{aligned}$$

The following properties hold:

- $|\varphi\rangle \otimes |\psi\rangle\,(\xi + \xi', \eta) = |\varphi\rangle \otimes |\psi\rangle\,(\xi, \eta) + |\varphi\rangle \otimes |\psi\rangle\,(\xi', \eta)$
- $|\varphi\rangle \otimes |\psi\rangle\,(a\xi, \eta) = \overline{a}\,|\varphi\rangle \otimes |\psi\rangle\,(\xi, \eta)$

so $|\varphi\rangle \otimes |\psi\rangle$ is a biantilinear functional on $\mathbb{H}^A \times \mathbb{H}^B$

Definition 3.2.2

To lighten the notation, we will sometimes use the following notations equivalently

$$|\varphi\rangle \otimes |\psi\rangle = |\varphi \otimes \psi\rangle = |\varphi, \psi\rangle = |\varphi\rangle\,|\psi\rangle = |\varphi\psi\rangle$$

Definition 3.2.3 – Tensor product

The tensor product of two Hilbert spaces $\mathbb{H}^A$ and $\mathbb{H}^B$ is defined as

$$\mathbb{H}^A \otimes \mathbb{H}^B = \{\Phi : \mathbb{H}^A \times \mathbb{H}^B \to \mathbb{C} \text{ biantilinear}\}$$

Remark 3.2.1.

1. $\mathbb{H}^A \otimes \mathbb{H}^B$ is a complex vector space
2. If $\{|e_i\rangle\}_{i=1,\dots,n_A}$ and $\{|f_j\rangle\}_{j=1,\dots,n_B}$ are ONBs for $\mathbb{H}^A$ and $\mathbb{H}^B$ respectively, then $\{|e_i\rangle \otimes |f_j\rangle\}_{i=1,\dots,n_A//j=1,\dots,n_B}$ is a basis for $\mathbb{H}^A \otimes \mathbb{H}^B$. It is also an ONB, but we haven't yet defined a scalar product over the tensor product
3. $\dim \mathbb{H}^A \otimes \mathbb{H}^B = \dim \mathbb{H}^A \cdot \dim \mathbb{H}^B$

Proof.

2. If $\xi \in \mathbb{H}^A$, $\eta \in \mathbb{H}^B$ then we can write $\xi = \sum_i \langle e_i|\xi\rangle^{\mathbb{H}^A} |e_i\rangle$ and $\eta = \sum_j = \langle f_j|\eta\rangle^{\mathbb{H}^B} |f_j\rangle$.

We can decompose a biantilinear map Ψ as

$$\Psi(\xi,\eta)=\sum_i\sum_j\overline{\xi_i\eta_j}\Psi_{ij}$$

where $\overline{\xi_i\eta_j}=\langle\xi|e_i\rangle^{\mathbb{H}^A}\langle\eta_j|f_j\rangle^{\mathbb{H}^B}=|e_i\rangle\otimes|f_j\rangle(\xi,\eta)$, so

$$\Psi(\xi,\eta)=\sum_i\sum_j\Psi_{ij}\,|e_i\rangle\otimes|f_j\rangle(\xi,\eta)$$

which means that $\{|e_i\rangle\otimes|f_j\rangle\}_{i,j}$ generate the tensor product. It remains to show that they are linearly independent.
If you pick $\xi=e_h$ and $\eta=f_k$, then

$$|e_i\rangle\otimes|f_j\rangle=\delta_{ih}\delta_{jk}$$

which, with some linear algebra, shows the linear independence

□

Definition 3.2.4

Given the tensor product $\mathbb{H}^A\otimes\mathbb{H}^B$, we define a scalar product $\langle\cdot|\cdot\rangle^{\mathbb{H}^A\otimes\mathbb{H}^B}$ on this space defined on the canonical basis as

$$\forall\varphi_1,\varphi_2\in\mathbb{H}^A,\ \psi_1,\psi_2\in\mathbb{H}^B,\qquad\langle\varphi_1\otimes\psi_1|\varphi_2\otimes\psi_2\rangle^{\mathbb{H}^A\otimes\mathbb{H}^B}=\langle\varphi_1|\varphi_2\rangle^{\mathbb{H}^A}\langle\psi_1|\psi_2\rangle^{\mathbb{H}^B}$$

and extended by linearity on the whole tensor product

Remark 3.2.2. If $\{|e_i\rangle\},\{|f_j\rangle\}$ are ONBs on $\mathbb{H}^A$ and $\mathbb{H}^B$ respectively, then $\{|e_i\rangle\otimes|f_j\rangle\}$ is indeed an ONB, and we can write the scalar product as

$$\langle\Psi|\Phi\rangle^{\mathbb{H}^A\otimes\mathbb{H}^B}=\sum_{i,j}\overline{\Psi_{ij}}\Phi_{ij}=\mathrm{tr}\left(M_\Psi^*M_\Phi\right)$$

where $M_\Psi=(\Psi_{ij})_{ij}\in\mathbb{C}^{n_A\times n_B}$.

The norm on the tensor product induced by the scalar product satisfies

$$\|\,|\varphi\rangle\otimes|\psi\rangle\,\|_{\mathbb{H}^A\otimes\mathbb{H}^B}=\|\varphi\|_{\mathbb{H}^A}\|\psi\|_{\mathbb{H}^B}$$

The definition of tensor product of two Hilbert spaces can be generalized to a tensor product of n Hilbert spaces, where

$$\begin{aligned}|\varphi_1\rangle\otimes\cdots\otimes|\varphi_n\rangle:\mathbb{H}^{A_1}\times\cdots\times\mathbb{H}^{A_n}&\to\mathbb{C}\\(\xi_1,\ldots,\xi_n)&\mapsto\prod_{i=1}^n\langle\xi_i|\varphi_i\rangle^{\mathbb{H}^{A_i}}\end{aligned}$$

Note that setwise, $(\mathbb{H}^A\otimes\mathbb{H}^B)\otimes\mathbb{H}^C\neq\mathbb{H}^A\otimes\mathbb{H}^B\otimes H^C$, but they are isomorphic as Hilbert spaces.

It's also important to notice that in general $\mathbb{H}^A \otimes \mathbb{H}^B \neq \mathbb{H}^A \otimes \mathbb{H}^B$. They, too, are isomorphic as Hilbert spaces, but in the future the order of qubits (and hence the Hilbert spaces representing them) will be important.

Example 3.2.5. Take $\mathbb{H}^A = \mathbb{H}^B = \mathbb{C}^2$ with ONBs $\{|0\rangle^{\mathbb{H}^A}, |1\rangle^{\mathbb{H}^A}\}$ and $\{|0\rangle^{\mathbb{H}^B}, |1\rangle^{\mathbb{H}^B}\}$. Then $\mathbb{H}^A \otimes \mathbb{H}^B \cong \mathbb{C}^4$ with ONB given by

$$\begin{aligned}
|0\rangle^{\mathbb{H}^A} \otimes |0\rangle^{\mathbb{H}^B} &= |00\rangle = |0\rangle \\
|0\rangle^{\mathbb{H}^A} \otimes |1\rangle^{\mathbb{H}^B} &= |01\rangle = |1\rangle \\
|1\rangle^{\mathbb{H}^A} \otimes |0\rangle^{\mathbb{H}^B} &= |10\rangle = |2\rangle \\
|1\rangle^{\mathbb{H}^A} \otimes |1\rangle^{\mathbb{H}^B} &= |11\rangle = |3\rangle
\end{aligned}$$

Any $\Psi \in \mathbb{H}^A \otimes \mathbb{H}^B$ can be written as $\Psi = \Psi_{00}|00\rangle + \Psi_{01}|01\rangle + \Psi_{10}|10\rangle + \Psi_{11}|11\rangle$ and can be representet either as a matrix in $\mathbb{C}^{2\times 2}$ or a vector in $\mathbb{C}^4$

$$\Psi \mapsto \begin{pmatrix} \Psi_{00} & \Psi_{01} \\ \Psi_{10} & \Psi_{11} \end{pmatrix} \mapsto \begin{pmatrix} \Psi_{00} \\ \Psi_{01} \\ \Psi_{10} \\ \Psi_{11} \end{pmatrix}$$

In this correspondence we have

$$|00\rangle \mapsto \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad |01\rangle \mapsto \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |10\rangle \mapsto \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad |11\rangle \mapsto \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Notice that if you take a number $x \in \{0, \dots, 3\}$ and express it in binary digits $x = (x_1 x_0)$ with le least significant digit being on the right (that is, formally $x = 2x_1 + x_0$), then in the correspondence we have $|x_1 x_0\rangle \mapsto e_{x+1}$. This is a useful fact that will allow us to use a simpler notation of $|x\rangle$ instead of writing x explicitly in binary digits

Definition 3.2.6 – Computational basis of n-fold tensor product of qubits

Given $\mathbb{H}^n = \underbrace{\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{n \text{ times}}$ with $\dim \mathbb{H}^n = 2^n$, the computational basis of $\mathbb{H}^n$ is defined as $\{|s\rangle\}_{s\in\{0,1\}^n}$, where $s = (x_{n-1} \dots x_1 x_0)$ with $x_i \in \{0, 1\}$ and $|s\rangle$ being a short representation for $|s\rangle = |x_{n-1}\rangle \otimes \cdots \otimes |x_0\rangle$. It can be proven by induction that the computational basis is an ONB

Definition 3.2.7 – Qubit system notation

We already know the meaning of $|b\rangle$ if b is a binary digit.
Given $s \in \{0, \dots, 2^n - 1\}$, and given its representation $s = (x_{n-1} \dots x_0)$ in binary digits (that

is, $s = \sum_{i=0}^{n-1} 2^i x_i$), we will use both $|s\rangle$ and $|x_{n-1} \dots x_0\rangle$ to represent the state

$$|s\rangle = |x_{n-1} \dots x_0\rangle = |x_{n-1}\rangle \otimes \cdots \otimes |x_0\rangle$$

of the computational base.
We will sometimes add a superscript to the state just to make its dimension explicit, as using the notations above can obscure the actual dimension of the system. For example,

$$|0\rangle^2 = |0\rangle \otimes |0\rangle \in \mathbb{H}^2, \qquad |0\rangle^3 = |0\rangle \otimes |0\rangle \otimes |0\rangle \in \mathbb{H}^3, \qquad |0\rangle^n = |0\rangle \otimes \cdots \otimes |0\rangle \in \mathbb{H}^n$$

The computational basis is the main basis we will use when building circuits and for measurements, but this is not the only possible basis we could use. Another possible choice is the Bell basis, that is the one given by the following Bell states

Definition 3.2.8 – Bell states

The Bell states are the following

$$|\Phi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \qquad |\Phi^-\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}$$
$$|\Psi^+\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}, \qquad |\Psi^-\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

Indeed one can check that the Bell states form an ONB for $\mathbb{H}^2$

3.2.1 States and observables for composite systems

Definition 3.2.9 – Postulate for composite systems

Given a system represented by $\mathbb{H}^A$, $\mathbb{H}^B$, the composite system is represented by $\mathbb{H}^A \otimes \mathbb{H}^B$.

Pure states in the composite system are represented by $|\psi\rangle \in \mathbb{H}^A \otimes \mathbb{H}^B$ with $\|\psi\| = 1$.
Mixed states are represented by $\rho \in D(\mathbb{H}^A \otimes \mathbb{H}^B)$, that is:

- $\rho : \mathbb{H}^A \otimes \mathbb{H}^B \to \mathbb{H}^A \otimes \mathbb{H}^B$ linear
- $\rho = \rho^*$
- $\rho \geq 0$
- $\text{tr}(\rho) = 1$

The mixed state $\rho \in D(\mathbb{H}^A \otimes \mathbb{H}^B)$ can be written as $\rho = \sum_{s,s' \in \{0,1\}^n} \rho_{s,s'} |s\rangle\langle s'|$ and represented by the density matrix $(\rho_{s,s'})_{s,s'} \in \mathbb{C}^{2^n \times 2^n}$.

We can consider $|\psi\rangle = |00\rangle \in \mathbb{H}^2$, which is a pure state. If we consider it as a mixed state, the reduced states on each qubit are $|0\rangle$ and $|0\rangle$.

If we consider $|\psi\rangle = |\Psi^+\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}$ thought, the situation on the reduced states is not so clear.

We're tempted to say that the reduced state on each qubit is the pure state $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$, but the correct interpretation is actually the mixed state $\frac{1}{2}|0\rangle\langle 0|+\frac{1}{2}|1\rangle\langle 1|$. We will now introduce some instruments that we will need to define the partial trace, which is exactly the tool that we need to define the state of a subregister given a composite system, and will explain why the answare is the mixed state $\frac{1}{2}|0\rangle\langle 0|+\frac{1}{2}|0\rangle\langle 0|$

October 26th, 2022

Let's build a particular observable on $\mathbb{H}^A\otimes\mathbb{H}^B$. Let $M_A : \mathbb{H}^A \to \mathbb{H}^A$ self-adjoint and $M_B : \mathbb{H}^B \to \mathbb{H}^B$ self-adjoint.
We can define the linear map $M_A \otimes M_B : \mathbb{H}^A \otimes \mathbb{H}^B \to \mathbb{H}^A \otimes \mathbb{H}^B$ on the canonical base by sending the element $|\psi\rangle \otimes |\varphi\rangle \to |M_A\psi\rangle \otimes |M_B\varphi\rangle$ and extending it by linearity (the verification that this is indeed a well defined linear map is left as an exercise to the reader).

This linear map is self-adjoint, hence an observable. This is because

$$\begin{aligned}
\langle \eta\otimes\xi | M_A\otimes M_B \psi\otimes\varphi\rangle^{\mathbb{H}^A\otimes\mathbb{H}^B} &= \langle \eta|M_A\psi\rangle^{\mathbb{H}^A}\ \langle \xi|M_B\varphi\rangle^{\mathbb{H}^B} = \\
&= \langle M_A\eta|\psi\rangle^{\mathbb{H}^A}\ \langle M_B\xi|\varphi\rangle^{\mathbb{H}^B} = \\
&= \langle M_A\otimes M_B\eta\otimes\xi|\psi\otimes\varphi\rangle^{\mathbb{H}^A\otimes\mathbb{H}^B}
\end{aligned}$$

Remark 3.2.3. Remember that given $\{|e_i\rangle\}$ and $\{|f_j\rangle\}$ ONBs for $\mathbb{H}^A$ and $\mathbb{H}^B$ then $\{|e_i\otimes f_j\rangle\}$ is and ONB for $\mathbb{H}^A\otimes\mathbb{H}^B$.

In this basis, the observable $M_A\otimes M_B$ is represented by a matrix in $\mathbb{C}^{(n_An_B)\times(n_An_B)}$ given by the Kronecker product of the two matrices representing M_A and M_B in the respective basis.

The two matrices representing M_A and M_B are

$$\begin{aligned}
(M^A_{ik})_{i,k=1,\dots,n_A} &= (\langle e_i|M_Ae_k\rangle) \in \mathbb{C}^{n_A\times n_A}\\
(M^B_{jl})_{j,l=1,\dots,n_B} &= (\langle e_j|M_Ae_l\rangle) \in \mathbb{C}^{n_B\times n_B}
\end{aligned}$$

and the Kronecker product is given by

$$\begin{pmatrix} M^A_{11}M^B & \dots & M^A_{1n_A}M^B \\ \vdots & & \vdots \\ M^A_{n_A1}M^B & \dots & M^A_{n_An_A}M^B \end{pmatrix} \in \mathbb{C}^{(n_An_B)\times(n_An_B)}$$

Example 3.2.10. Consider the case $\mathbb{H}^A = \mathbb{H}^B = \mathbb{C}^2$ (so $\mathbb{H}^A\otimes\mathbb{H}^B = \mathbb{C}^2\otimes\mathbb{C}^2$).
Some possible operators expressed in matrix form are

$$\sigma_x\otimes\sigma_x = \left(\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right)$$

which acts by sending

$$\begin{aligned}
|00\rangle &\mapsto \sigma_x \otimes \sigma_x |00\rangle = |11\rangle \\
|01\rangle &\mapsto \sigma_x \otimes \sigma_x |01\rangle = |10\rangle \\
|10\rangle &\mapsto \sigma_x \otimes \sigma_x |10\rangle = |01\rangle \\
|11\rangle &\mapsto \sigma_x \otimes \sigma_x |11\rangle = |00\rangle
\end{aligned}$$

Also

$$\sigma_z \otimes \sigma_z = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

which acts by sending

$$\begin{aligned}
|00\rangle &\mapsto \sigma_x \otimes \sigma_z |00\rangle = |00\rangle \\
|01\rangle &\mapsto \sigma_x \otimes \sigma_z |01\rangle = -|01\rangle \\
|10\rangle &\mapsto \sigma_x \otimes \sigma_z |10\rangle = -|10\rangle \\
|11\rangle &\mapsto \sigma_x \otimes \sigma_z |11\rangle = |11\rangle
\end{aligned}$$

In general, with a bit of an abuse of notation on the indices of the Kronecker product matrix, we have

$$\begin{aligned}
\left(M^A \otimes M^B\right)_{(i,j)(l,k)} &= \langle e_i \otimes f_j | M_A \otimes M_B e_k \otimes f_l \rangle = \\
&= \left\langle e_i \middle| M^A e_k \right\rangle \left\langle f_j \middle| M^B f_l \right\rangle = \\
&= M^A_{ik} M^B_{jl}
\end{aligned}$$

The index $(i,j)(l,k)$ actually means $((i-1)n_B + j, (k-1)n_B + l)$, but by writing the Kronecker product in block form, the double indices $(i,j)(l,k)$ end up being more useful as they specify which block, and then the coordinates inside that block.

Note that quite surprisingly, the observables $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$ commute, because

$$\begin{aligned}
(\sigma_x \otimes \sigma_x)(\sigma_z \otimes \sigma_z) |\varphi \otimes \psi\rangle &= (\sigma_x \otimes \sigma_x) |\sigma_z \varphi \otimes \sigma_z \psi\rangle = \\
&= |(\sigma_x \sigma_z \varphi) \otimes (\sigma_x \sigma_z \psi)\rangle = \\
&= (\sigma_x \sigma_z) \otimes (\sigma_x \sigma_z) |\varphi \otimes \psi\rangle
\end{aligned}$$

$$\begin{aligned}
(\sigma_z \otimes \sigma_z)(\sigma_x \otimes \sigma_x) |\varphi \otimes \psi\rangle &= (\sigma_z \otimes \sigma_z) |\sigma_x \varphi \otimes \sigma_x \psi\rangle = \\
&= |(\sigma_z \sigma_x \varphi) \otimes (\sigma_z \sigma_x \psi)\rangle = \\
&= (\sigma_z \sigma_x) \otimes (\sigma_z \sigma_x) |\varphi \otimes \psi\rangle
\end{aligned}$$

and since $\sigma_x \sigma_z = i\sigma_y$ and $\sigma_x \sigma_z = -i\sigma_y$ we get

$$\begin{aligned}
(\sigma_x \otimes \sigma_x)(\sigma_z \otimes \sigma_z) &= (-i\sigma_y) \otimes (-i\sigma_y) = \\
&= (i\sigma_y) \otimes (i\sigma_y) = \\
&= (\sigma_z \otimes \sigma_z) \otimes (\sigma_x \otimes \sigma_x)
\end{aligned}$$

Since they commute they can be simultaneously diagonalized. In fact, the basis that diagonalizes them both is Bell's basis, which is composed of eigenvectors of both. The eigenvalues are

$$
\begin{aligned}
(\sigma_x \otimes \sigma_x) \left|\Phi^{\pm}\right\rangle &= \frac{\sigma_x \otimes \sigma_x \left|00\right\rangle \pm \sigma_x \otimes \sigma_x \left|11\right\rangle}{\sqrt{2}} = \\
&= \frac{\left|11\right\rangle \pm \left|00\right\rangle}{\sqrt{2}} = \pm \left|\Phi^{\pm}\right\rangle \\
(\sigma_z \otimes \sigma_z) \left|\Phi^{\pm}\right\rangle &= \frac{\sigma_z \otimes \sigma_z \left|00\right\rangle \pm \sigma_z \otimes \sigma_z \left|11\right\rangle}{\sqrt{2}} = \\
&= \frac{\left|00\right\rangle \pm \left|11\right\rangle}{\sqrt{2}} = \left|\Phi^{\pm}\right\rangle \\
(\sigma_x \otimes \sigma_x) \left|\Psi^{\pm}\right\rangle &= \frac{\sigma_x \otimes \sigma_x \left|01\right\rangle \pm \sigma_x \otimes \sigma_x \left|10\right\rangle}{\sqrt{2}} = \\
&= \frac{\left|10\right\rangle \pm \left|01\right\rangle}{\sqrt{2}} = \pm \left|\Psi^{\pm}\right\rangle \\
(\sigma_z \otimes \sigma_z) \left|\Psi^{\pm}\right\rangle &= \frac{\sigma_z \otimes \sigma_z \left|01\right\rangle \pm \sigma_z \otimes \sigma_z \left|10\right\rangle}{\sqrt{2}} = \\
&= \frac{-\left|01\right\rangle \mp \left|10\right\rangle}{\sqrt{2}} = -\left|\Psi^{\pm}\right\rangle
\end{aligned}
$$

by measuring simultaneously $(\sigma_x \otimes \sigma_x)$ and $(\sigma_z \otimes \sigma_z)$ we can detect in which of the Bell states the system is, where the correspondence is given by the following table

$\sigma_x \otimes \sigma_x$ $\sigma_z \otimes \sigma_z$	1	-1
1	$\left|\Phi^+\right\rangle$	$\left|\Phi^-\right\rangle$
-1	$\left|\Psi^+\right\rangle$	$\left|\Psi^-\right\rangle$

By the postulates, the expectation of $M_A \otimes M_B$ on the mixed state $\rho \in D(\mathbb{H}^A \otimes \mathbb{H}^B)$ is $\operatorname{tr}((M_A \otimes M_B)\rho)$. If we represent ρ with a density matrix with respect to the ONB $\{\left|e_i \otimes f_j\right\rangle\}$, then we get

$$
\begin{aligned}
\operatorname{tr}((M_A \otimes M_B)\rho) &= \sum_{\substack{i=1\ldots n_A \\ j=1\ldots n_B}} \left\langle e_i \otimes f_j \middle| M_A \otimes M_B \rho e_i \otimes f_j \right\rangle = \\
&= \sum_{i,j} \left\langle M_A e_i \otimes M_B f_j \middle| \sum_{k,l} \left|e_k \otimes f_l\right\rangle \rho_{(k,l)(i,j)} \right\rangle = \\
&= \sum_{i,j,k,l} \left\langle M_A e_i \otimes M_B f_j \middle| e_k \otimes f_k \right\rangle \rho_{(k,l)(i,j)} = \\
&= \sum_{i,j,k,l} \left\langle M_A e_i \middle| e_k \right\rangle \left\langle M_B f_j \middle| f_l \right\rangle \rho_{(k,l)(i,j)}
\end{aligned}
$$

For example, if we take $M_B = \mathbb{1}_B$ we get

$$\begin{aligned}\mathrm{tr}\left((M_A \otimes \mathbb{1}_B)\rho\right) &= \sum_{i,j,k,l} \langle M_A e_i | e_k \rangle \, \delta_{jl} \rho_{(k,l)(i,j)} = \\ &= \sum_{i,k=1}^{n_A} \langle M_A e_i | e_k \rangle \sum_{j=1}^{n_B} \rho_{(k,j)(i,j)}\end{aligned}$$

Notice that $\mathrm{tr}\,(\rho) = \sum_{i=1}^{n_A} \sum_{j=1}^{n_B} \rho_{(i,j)(i,j)}$.

We can now define the partial trace

Definition 3.2.11 – Partial trace

Let $M : \mathbb{H}^A \otimes \mathbb{H}^B \to \mathbb{H}^A \otimes \mathbb{H}^B$ be a linear operator. We define the partial trace on B, $\mathrm{tr}^B\,(\rho) : \mathbb{H}^A \to \mathbb{H}^A$ as the only linear operator $L^A : \mathbb{H}^A \to \mathbb{H}^A$ that satisfies

$$\forall K : \mathbb{H}^A \to \mathbb{H}^A, \qquad \mathrm{tr}\left(KL^A\right) = \mathrm{tr}\left((K \otimes \mathbb{1}_B)M\right)$$

Similarly we can define $\mathrm{tr}^A\,(M) : \mathbb{H}^B \to \mathbb{H}^B$

Explicitly (by repeating the same calculations we've done previously) we get

$$\begin{aligned}\mathrm{tr}^A\,(M) &= \sum_{i,k=1}^{n_A} |e_i\rangle\langle e_k| \left(\sum_{j=1}^{n_B} \underbrace{\langle e_i \otimes f_j | M e_k \otimes f_j \rangle}_{=M_{(i,k)(k,j)}} \right) \\ \mathrm{tr}^B\,(M) &= \sum_{j,l=1}^{n_B} |f_j\rangle\langle f_l| \left(\sum_{i=1}^{n_A} \langle e_i \otimes f_j | M e_i \otimes f_l \rangle \right)\end{aligned}$$

To show that the definition is well posed, that is L^A is actually unique, suppose that L^A and $\widetilde{L^A}$ are such that $\forall K$, $\mathrm{tr}\left(KL^A\right) = \mathrm{tr}\left(K\widetilde{L^A}\right)$. Then we get $\mathrm{tr}\left(K(L^A - \widetilde{L^A})\right) = 0$. It can be proved that the map $(K, C) \mapsto \mathrm{tr}\,(KC)$ defines a scalar product (see Frobenius inner product), the fact that $\forall K$, $\mathrm{tr}\left(K(L^A - \widetilde{L^A})\right) = 0$ implies that $L^A - \widetilde{L^A} = 0$, hence L^A is actually unique.

Proposition 3.2.12

The following properties hold

1. $\forall M, N$ operators, $\forall \lambda \in \mathbb{C}$, it holds $\mathrm{tr}^A\,(M + \lambda N) = \mathrm{tr}^A\,(M) + \lambda \mathrm{tr}^A\,(N)$.
   More generally $\forall K : \mathbb{H}^B \to \mathbb{H}^B$, $\mathrm{tr}^A\,((\mathbb{1}_A \otimes K)M) = K\mathrm{tr}^A\,(M)$
2. If M is self-adjoint and non-negative, then $\mathrm{tr}^A\,(M)$ is also self-adjoint and non-negative
3. $\mathrm{tr}\left(\mathrm{tr}^A\,(M)\right) = \mathrm{tr}\,(M)$

Similar properties also hold for $\mathrm{tr}^B\,(\cdot)$

In particular, given $\rho \in D(\mathbb{H}^A \otimes \mathbb{H}^B)$ we have $\mathrm{tr}^A\,(\rho) = \rho^B \in D(\mathbb{H}^B)$ and $\mathrm{tr}^B\,(\rho) = \rho^A \in D(\mathbb{H}^A)$ which are the reduced density operators.

Exercise 3.2.13. Prove that if $M : \mathbb{H}^A \otimes \mathbb{H}^B \otimes \mathbb{H}^C \to \mathbb{H}^A \otimes \mathbb{H}^B \otimes \mathbb{H}^C$, then $\mathrm{tr}^{AB}(M) = \mathrm{tr}^A\left(\mathrm{tr}^B(M)\right)$, where $\mathrm{tr}^{AB}(M)$ is defined by intepreting the tensor product $\mathbb{H}^A \otimes \mathbb{H}^B \otimes \mathbb{H}^C$ as $(\mathbb{H}^A \otimes \mathbb{H}^B) \otimes \mathbb{H}^C$

Exercise 3.2.14. Prove that if $U^B : \mathbb{H}^B \to \mathbb{H}^B$ is unitary then

$$\forall M, \qquad \mathrm{tr}^B(M) = \mathrm{tr}^B\left((\mathbb{1}_A \otimes U^B)M(\mathbb{1}_A \otimes U^B)^*\right)$$

Example 3.2.15. Let $\mathbb{H}^A = \mathbb{H}^B\mathbb{C}^2$, let's compute $\mathrm{tr}^B(|\Phi^+\rangle\langle\Phi^+|)$.

$$\begin{aligned}\mathrm{tr}^B\left(|\Phi^+\rangle\langle\Phi^+|\right) = \mathrm{tr}^B\left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) &= \\ &= \frac{1}{2}\left[\mathrm{tr}^B(|00\rangle\langle 00|) + \mathrm{tr}^B(|00\rangle\langle 11|) + \mathrm{tr}^B(|11\rangle\langle 00|) + \mathrm{tr}^B(|11\rangle\langle 11|)\right] = \\ &= \frac{1}{2}[|0\rangle\langle 0| + |1\rangle\langle 1|] = \\ &= \frac{1}{2}\mathbb{1}_A\end{aligned}$$

Note that this is not a pure state!

October 28th, 2022

3.3 Entanglement

As we know, given the systems $\mathbb{H}^A$ and $\mathbb{H}^B$ we can build the composite system $\mathbb{H}^A \otimes \mathbb{H}^B$. If we take $|\varphi\rangle \in \mathbb{H}^A$ and $|\psi\rangle \in \mathbb{H}^B$ we can obtain the state $|\varphi\rangle \otimes |\psi\rangle \in \mathbb{H}^A \otimes \mathbb{H}^B$, but not every state in the tensor product $\mathbb{H}^A \otimes \mathbb{H}^B$ can be represented as $|\varphi\rangle \otimes |\psi\rangle$ for some $|\varphi\rangle \in \mathbb{H}^A$ and $|\psi\rangle \in \mathbb{H}^B$.

Example 3.3.1. Let's consider Bell's states in a two-qubit composite system, that is $\mathbb{H}^A = \mathbb{H}^B = \mathbb{C}^2$.

Suppose there existed two states $|\varphi\rangle \in \mathbb{H}^A$, $|\psi\rangle \in \mathbb{H}^B$ such that $|\varphi\rangle \otimes |\psi\rangle = |\Phi^+\rangle$. If that was the case, we could write with respect to the computational base

$$|\varphi\rangle = \varphi_0 |0\rangle + \varphi_1 |1\rangle, \qquad |\psi\rangle = \psi_0 |0\rangle + \psi_1 |1\rangle$$

which yields

$$|\varphi\rangle \otimes |\psi\rangle = \varphi_0\psi_0 |00\rangle + \varphi_0\psi_1 |01\rangle + \varphi_1\psi_0 |10\rangle + \varphi_1\psi_1 |11\rangle$$

If we now impose $|\varphi\rangle \otimes |\psi\rangle = |\Phi^+\rangle$ we get the system

$$\begin{cases}\varphi_0\psi_0 = \varphi_1\psi_1 = \frac{1}{\sqrt{2}} \\ \varphi_0\psi_1 = \varphi_1\psi_0 = 0\end{cases}$$

but this system has no solution.

This means that Bell's state $|\Phi^+\rangle$ (and similarly for the other three states) can't be represented as $|\varphi\rangle \otimes |\psi\rangle$

Definition 3.3.2 – Separable and entangled states

A state in $\mathbb{H}^A \otimes \mathbb{H}^B$ is called separable if it can be written as $|\varphi\rangle \otimes |\psi\rangle$ for some $|\varphi\rangle \in \mathbb{H}^A$ and $|\psi\rangle \in \mathbb{H}^B$. Otherwise, that state is called entangled

Note that if $|\Psi\rangle = |\varphi\rangle \otimes |\psi\rangle$, then by iterpreting it as a mixed state we get

$$\begin{aligned}\rho_\Psi &= |\Psi\rangle\langle\Psi| = \\ &= (|\varphi\rangle \otimes |\psi\rangle)(\langle\varphi| \otimes \langle\psi|) = \\ &= (|\varphi\rangle\langle\varphi|) \otimes (|\psi\rangle\langle\psi|)\end{aligned}$$

Definition 3.3.3

A mixed state $\rho \in D(\mathbb{H}^A \otimes \mathbb{H}^B)$ is called separable if it can be written

$$\rho = \sum_{j\in I} p_j \rho_j^{(A)} \otimes \rho_j^{(B)}$$

with $\sum_{j\in I} p_j = 1$ and $\rho_j^{(A)} \in D(\mathbb{H}^A)$, $\rho_j^{(B)} \in D(\mathbb{H}^B)$. Otherwise, it's called entangled

Theorem 3.3.4

The definition of separable/entangled for pure states and mixed states are consistent

For example, we've shown that Bell's states are entangled. The state $|\Psi\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$ is separable as we can write

$$|\Psi\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

Theorem 3.3.5

Let $|\Psi\rangle \in \mathbb{H}^A \otimes \mathbb{H}^B$ be a pure state. Then $|\Psi\rangle$ is separable if and only if both $\rho^A_{|\Psi\rangle}$ and $\rho^B_{|\Psi\rangle}$ are pure state

Example 3.3.6. We can apply this criteria to $|\Phi^+\rangle$. We saw that $\rho^A_{\Phi^+} = \text{tr}^B\left(|\Phi^+\rangle\langle\Phi^+|\right) = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$, so $\rho^A_{\Phi^+} = \frac{1}{2}\mathbb{1}$.
We also know that a mixed state ρ is pure if and only if $\rho^2 = \rho$.

Since $(\rho^A_{\Phi^+})^2 = \left(\frac{1}{2}\mathbb{1}\right)^2 = \frac{1}{4}\mathbb{1} \neq \frac{1}{2}\mathbb{1} = \rho^A_{\Phi^+}$ we get that $|\Phi^+\rangle$ is entangled

Consider the composite system $\mathbb{H}^{ABCD} = \mathbb{H}^A \otimes \mathbb{H}^B \otimes \mathbb{H}^C \otimes \mathbb{H}^D$ with $\mathbb{H}^A \cong \mathbb{H}^B \cong \mathbb{H}^C \cong \mathbb{H}^D \cong \mathbb{C}^2$ qubits.

We can prepare the system $\mathbb{H}^{ABCD}$ in the state $|\Phi\rangle$ defined as

$$|\Phi\rangle = \left|\Psi^-\right\rangle^{AB} \otimes \left|\Psi^-\right\rangle^{CD}$$

where

$$\left|\Psi^-\right\rangle^{AB} = \frac{|01\rangle^{AB} - |10\rangle^{AB}}{\sqrt{2}}$$
$$\left|\Psi^-\right\rangle^{CD} = \frac{|01\rangle^{CD} - |10\rangle^{CD}}{\sqrt{2}}$$

Notice (or prove as an exercise) that

$$\begin{aligned}|\Phi\rangle =&\frac{1}{2}(|0101\rangle - |1001\rangle - |0110\rangle + |1010\rangle) = \\ =&\frac{1}{2}(\left|\Psi^+\right\rangle^{AD} \otimes \left|\Psi^+\right\rangle^{BC} - \left|\Psi^-\right\rangle^{AD} \otimes \left|\Psi^-\right\rangle^{BC} - \\ &- \left|\Phi^+\right\rangle^{AD} \otimes \left|\Phi^+\right\rangle^{BC} + \left|\Phi^-\right\rangle^{AD} \otimes \left|\Phi^-\right\rangle^{BC}\end{aligned}$$

We can define the observables

$$\Sigma_z = \mathbb{1} \otimes \sigma_z \otimes \sigma_z \otimes \mathbb{1}, \qquad \Sigma_x = \mathbb{1} \otimes \sigma_x \otimes \sigma_x \otimes \mathbb{1}$$

As we've seen in the 2 qubit system, $(\sigma_x \otimes \sigma_x)$ and $(\sigma_z \otimes \sigma_z)$ commute, so one can reasonably believe (and it can be proven formally) that Σ_z and Σ_x commute.
In the 2 qubit system, measuring $(\sigma_x \otimes \sigma_x)$ and $(\sigma_z \otimes \sigma_z)$ allowed us to prepare the system in one of the Bell's state[1]. Similarly, we can measure Σ_x and Σ_z to detect in which of the states $\left|\Phi^\pm\right\rangle^{BC}$, $\left|\Psi^\pm\right\rangle^{BC}$ the subsystem $\mathbb{H}^{BC}$ is. Given the decomposition of $|\Phi\rangle$, knowing the state of the system $\mathbb{H}^{BC}$ will tell us the state of $\mathbb{H}^{AD}$.

For example, if we measure $\Sigma_x\Sigma_z$ and obtain $(+1,+1)$, we know that $\mathbb{H}^{BC}$ is in state $\left|\Phi^+\right\rangle^{BC}$, which means that the global system $\mathbb{H}^{ABCD}$ must be in state $\left|\Phi^+\right\rangle^{AD} \otimes \left|\Phi^+\right\rangle^{BC}$. This is because, in the prepared state $|\Phi\rangle$, the subsystem $\mathbb{H}^{AD}$ is entangled with the subsystem $\mathbb{H}^{BC}$.

[1] that is, it collapsed the system in one of those states and allowed us to detect which state it collapsed to. By repeating this procedure untill the desired result, we can prepare the system in a desired state

4 | Quantum circuits

4.1 Quantum copier

In a classical system it's taken for granted that once you have a set of bits (or even just a single bit) you can copy and share them as much as you like. We would like to do the same thing in a quantum system, that is we would like to be able to duplicate or clone the state of a qubit to another qubit, without losing the original.

Formally, a quantum copier is an operator $K : \mathbb{H} \otimes \mathbb{H} \to \mathbb{H} \otimes \mathbb{H}$ such that given a fixed state $|w\rangle \in \mathbb{H}$ (on which K depends), and any state $|\varphi\rangle \in \mathbb{H}$, it maps $|\varphi\rangle \otimes |w\rangle \mapsto K\, |\varphi\rangle \otimes |w\rangle = |\varphi\rangle \otimes |\varphi\rangle$.

Note that we must require that the second system is prepared in a predeterminated state $|w\rangle$ and not just any state, because an operator that sends the state $|\varphi\rangle \otimes *$ to $|\varphi\rangle \otimes |\varphi\rangle$ would not be injective and clearly not unitary, which means it couldn't be built as a quantum operator.

Theorem 4.1.1 – No cloning theorem

There is no quantum copier

Proof. We will prove this theorem on a 2-qubit system (that is $\mathbb{H} \cong \mathbb{C}^2$).
Fix any $|w\rangle \in \mathbb{H}$. Ad absurdum, suppose there existed a quantum copier K. Then we would get

$$K(|0\rangle \otimes |w\rangle) = |0\rangle \otimes |0\rangle = |00\rangle$$
$$K(|1\rangle \otimes |w\rangle) = |1\rangle \otimes |1\rangle = |11\rangle$$

If we instead apply K on $\frac{|0\rangle+|1\rangle}{\sqrt{2}} \otimes |w\rangle$ we should get (using the copying property)

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

but by linearity we should get

$$K\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes |w\rangle\right) = \frac{1}{\sqrt{2}}(K(|0\rangle \otimes |w\rangle) + K(|1\rangle \otimes |w\rangle)) = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

This is clearly absurd, which means a quantum copier can't exist □

This means that when building quantum circuits, we won't be free of copying a reusing multiple times a given qubit, which will be quite a restriction.

4.2 EPR states and Bell telephone

Suppose two parties (which will be referred to from now on as Alice and Bob) have each their own qubit, and a 2-qubit system $\mathbb{H}^A \otimes \mathbb{H}^B$ is built from these two qubits.
Suppose we prepare this system in Bell's state

$$\begin{aligned} \left|\Phi^+\right\rangle &= \frac{1}{\sqrt{2}}(|{\uparrow_{\hat{z}}}\rangle \otimes |{\uparrow_{\hat{z}}}\rangle + |{\downarrow_{\hat{z}}}\rangle \otimes |{\downarrow_{\hat{z}}}\rangle) = \\ &= \frac{1}{\sqrt{2}}(|{\uparrow_{\hat{x}}}\rangle \otimes |{\uparrow_{\hat{x}}}\rangle + |{\downarrow_{\hat{x}}}\rangle \otimes |{\downarrow_{\hat{x}}}\rangle) \end{aligned}$$

Suppose Alice measures σ_z (which means we measure $\sigma_z^A \otimes \mathbb{1}_B$ on the global system) on her quibit and gets 1. Then her qubit must be in state $|{\uparrow_{\hat{z}}}\rangle$, and by entanglement Bob's qubit must be on $|{\uparrow_{\hat{z}}}\rangle$ even though he took no measurement.
The same would happen if Alice measured σ_x. This would be a way of sending information instantly and is at the base of the EPR paradox.

Let's see how we can use this mechanism to send information (this concept is called Bell telephone). We can encode classical bits in the following way:

- if Alice wants to send a 0, she can measure σ_z making Bob's qubit either $|0\rangle$ or $|1\rangle$
- if Alice wants to send a 1, she can measure σ_x making Bob's qubit either $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$ or $\frac{|0\rangle-|1\rangle}{\sqrt{2}}$

The thing is that just by a single measurement, Bob cannot distinguish whether the qubit was in state $|0\rangle$, $|1\rangle$, $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$ or $\frac{|0\rangle-|1\rangle}{\sqrt{2}}$. If he could measure σ_z multiple times, he could know if Alice measured σ_z (in which case he would get the same result every time) or if Alice measured σ_x (in which case he would get random results). The issue is that Bob would need multiple copies of the same bit, and as we've shown, a qubit cannot be cloned.

An argument can be made for using multiple qubits to send a single bit of information, but even in this case it would be prone to errors as it might happen that Alice measured σ_x and by chance Bob could measure n times and get the same (but technically random) result, leading him to believe that Alice measured σ_z.

November 2$^{\text{nd}}$, 2022

4.3 Classical gates

In classical computing, transforming a state means modifying a finite sequence of binary digits into another finite sequence of binary digits. At a low level, this is done by a combination of binary gates, that we can imagine as blackboxes that take an input $(x_1, \ldots, x_n)$ and return an output $(o_1, \ldots, o_m)$ and are described by defining the result on every possible input via a truth table. More precisely:

Definition 4.3.1 – Gates

An elementary classical gate is a function $e : \{0,1\}^n \to \{0,1\}$.
A classical gate is a function $g : \{0,1\}^n \to \{0,1\}^m$ composed of the product of m elementary gates.

A classical gate is fully described by a truth table, that is a table where every possible input is paired with the corresponding output, such as

x_1	$\dots$	x_{n-1}	x_n	o_1	$\dots$	o_m
0	$\dots$	0	0	*	$\dots$	*
0	$\dots$	0	1	*	$\dots$	*
0	$\dots$	1	0	*	$\dots$	*
0	$\dots$	1	1	*	$\dots$	*
	$\vdots$				$\vdots$	
1	$\dots$	1	1	*	$\dots$	*

Example 4.3.2. We will use $\oplus$ to indicate the addition modulo N for some case-specific N, and $\boxplus$ to indicate the bitwise addition modulo 2. Here are some notable examples of classical gates

- NOT : $\{0,1\} \to \{0,1\}$, which can be written as NOT$(x) = 1 \oplus x$, with the following truth table

x	o
0	1
1	0

- AND : $\{0,1\}^2 \to \{0,1\}$, which can be written as AND$(x_1, x_2) = x_1 x_2$, with the following truth table

x_1	x_2	o
0	0	0
0	1	0
1	0	0
1	1	1

- XOR : $\{0,1\}^2 \to \{0,1\}$, which can be written as XOR$(x_1, x_2) = x_1 \oplus x_2$, with the following truth table

x_1	x_2	o
0	0	0
0	1	1
1	0	1
1	1	0

- OR : $\{0,1\}^2 \to \{0,1\}$, which can be written as OR$(x_1, x_2) = x_1 \oplus x_2 \oplus (x_1 x_2)$, with the following truth table

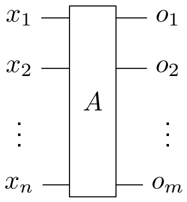
x_1	x_2	o
0	0	0
0	1	1
1	0	1
1	1	1

- TOF$\{0,1\}^3 \to \{0,1\}^3$, the Toffoli gate, which can be written as TOF$(x_1, x_2, x_3) = (x_1, x_2, x_1x_2 \oplus x_3)$, with the following truth table

x_1	x_2	x_3	o_1	o_2	o_3
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

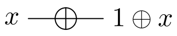
Interpreting $\{0,1\}^3$ as binary numbers, the Toffoli gate is the permutation $(6, 7)$

Graphically, we express a generic gate $g : \{0,1\}^n \to \{0,1\}^m$ as the following diagram

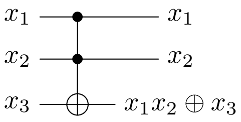


but there some special notations for some gates, such as

- the NOT gate is represented as



- the Toffoli gate is represented as



where the black dot means that we're applying the identity on that bit but we're also using that bit as a control to decide whether to apply the NOT gate on x_3

Definition 4.3.3

We say that a gate is reversible if the associated function is a bijection

Note that the NOT gate and the Toffoli gate are reversible, while the AND, XOR and OR gate are not reversible.

The notion of reversibility is particularly important in quantum computation because we can only apply unitary operators, which are necessarily reversible. This means that in quantum computation we cannot apply AND, XOR and OR gates

The workaround to solve this problem is the use of auxiliary qubits, that are usually called "ancilla" qubits, such that when prepared in a specific state, by using a different gate we can emulate a non bijective gate. For example

- TOF$(x_1,\, x_2,\, 0) = (x_1,\, x_2,\, x_1x_2 \oplus 0) = (x_1,\, x_2,$ AND$(x_1,\, x_2))$
- TOF$(1,\, x_2,\, x_3) = (1,\, x_2, 1, x_2 \oplus x_3) = (1, x_2,$ XOR$(x_2,\, x_3))$
- TOF$(1,\, 1,\, x_3) = (1,\, 1,\, 1 \oplus x_3) = (1,\, 1,$ NOT$(x_3))$

Other useful gates in the classical setting are:

- Id : $x_1 \mapsto x_1$
- FALSE : $x_1 \mapsto 0$
- TRUE : $x_1 \mapsto 1$
- COPY : $x_1 \mapsto (x_1,\, x_1)$

Given $g_1, \ldots, g_k$ gates, we define $F(g_1, \ldots, g_k)$ the set of gates that can be built using $g_1, \ldots, g_k$ as building blocks, according to the following rules:

1. $g_1, \ldots, g_k \in F(g_1, \ldots, g_k)$
2. Padding is allowed, where padding is defined as

$$\begin{aligned} P^{(n)}_{y_1,\ldots,y_l,j_1,\ldots,j_l} :\ & \{0,1\}^n \xrightarrow{\qquad\qquad} \{0,1\}^{n+l} \\ & (x_1, \ldots, x_n) \longmapsto (x_1, \ldots, x_{j_1 - 1}, y_1, x_{j_1}, \ldots) \end{aligned}$$

   That is, the padding operator takes an n-bit state $(x_1, \ldots, x_n)$, l indices $j_1, \ldots, j_l$ and l bits $y_1, \ldots, y_l$ and inserts the given bits y_i in the corresponding position j_i in the given state, obtaining a new $(n+l)$-bit state

3. Restriction and reorderings are allowed, and defined as

$$\begin{aligned} r^{(n)}_{j_1,\ldots,j_l} :\ & \{0,1\}^n \xrightarrow{\qquad} \{0,1\}^{l} \\ & (x_1, \ldots, x_n) \longmapsto (x_{j_1}, \ldots x_{j_l}) \end{aligned}$$

   where $l \leq n$ and $j_1, \ldots, j_l$ are distinct

4. Composition of gates is allowed, that is

$$h_1, h_2 \in F(g_1, \ldots, g_k) \Rightarrow h_1 \circ h_2 \in F(g_1, \ldots, g_k)$$

5. Cartesian product is allowed, that is given $h_1 : \{0,1\}^n \to \{0,1\}^m$ and $h_2 : \{0,1\}^p \to \{0,1\}^q$ such that $h_1, h_2 \in F(g_1, \dots, g_k)$, then $h_1 \times h_2 \in F(g_1, \dots, g_k)$, where

$$\begin{aligned} h_1 \times h_2 : \qquad & \{0,1\}^{n+p} \longrightarrow \{0,1\}^{m+q} \\ & (x_1, \dots, x_n, x_{n+1}, \dots, x_{n+p}) \longmapsto (h_1(x_1, \dots, x_n), h_2(x_{n+1}, \dots, x_{n+p})) \end{aligned}$$

Example 4.3.4. Suppose we start with the set of elementary gates $(\mathsf{AND}, \mathsf{XOR})$ and we would like to build a Toffoli gate. It seems reasonable to be able to do so, given that the Toffoli gate requires only AND and XOR of a specific combination of bits. Indeed we have

$$\begin{aligned} &(\mathsf{Id} \times \mathsf{Id} \times \mathsf{XOR}) \circ (\mathsf{Id} \times \mathsf{Id} \times \mathsf{AND} \times \mathsf{Id}) \circ r^{(5)}_{13245} \circ (\mathsf{COPY} \times \mathsf{COPY} \times \mathsf{Id})(x_1, x_2, x_3) = \\ &= (\mathsf{Id} \times \mathsf{Id} \times \mathsf{XOR}) \circ (\mathsf{Id} \times \mathsf{Id} \times \mathsf{AND} \times \mathsf{Id}) \circ r^{(5)}_{13245}(x_1, x_1, x_2, x_2, x_3) = \\ &= (\mathsf{Id} \times \mathsf{Id} \times \mathsf{XOR}) \circ (\mathsf{Id} \times \mathsf{Id} \times \mathsf{AND} \times \mathsf{Id})(x_1, x_2, x_1, x_2, x_3) = \\ &= (\mathsf{Id} \times \mathsf{Id} \times \mathsf{XOR})(x_1, x_2, x_1 x_2, x_3) = \\ &= (x_1, x_2, x_1 x_2 \oplus x_3) = \\ &= \mathsf{TOF}(x_1, x_2, x_3) \end{aligned}$$

Definition 4.3.5 – Universal set of gates

We say that a set of gates $g_1, \dots, g_k$ is universal if any gate can be built starting with them, that is

$$\forall g \text{ gate, } g \in F(g_1, \dots, g_k)$$

Theorem 4.3.6

The (classical) Toffoli gate is reversible, and $\{\mathsf{TOF}\}$ is universal

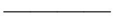
4.4 Quantum gates

Definition 4.4.1 – Quantum gates

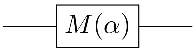
A quantum n-gate is a unitary operator $U : \mathbb{H}^{\otimes n} \to \mathbb{H}^{\otimes n}$, where $\mathbb{H} \cong \mathbb{C}^2$ is the Hilbert space of a single qubit.
$\mathbb{H}^{\otimes n}$, that is the composite system of n qubits on which we apply the quantum gates, is called a quantum register

Here are some examples of quantum gates that we will use often. The first set is a set of unary gates, that is gates that act on a single qubit

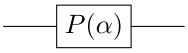
- Id, expressed in matrix form as $\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and drawn as



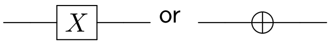
- Phase factor, expressed in matrix form as $e^{i\alpha}\mathbb{1} = e^{i\alpha}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and drawn as



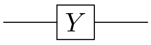
- Phase shift, expressed in matrix form as $|0\rangle\langle 0| + e^{i\alpha}\,|1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{pmatrix}$ and drawn as



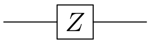
- QNOT, expressed in matrix form as $X = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and drawn as



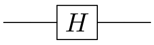
- Pauli Y, expressed in matrix form as $Y = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ and drawn as



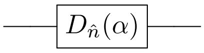
- Pauli Z, expressed in matrix form as $Z = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and drawn as



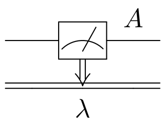
- Hadamard, expressed in matrix form as $H = \frac{\sigma_x+\sigma_z}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ and drawn as



- Spin rotation, expressed in matrix form as $D_{\hat{n}}(\alpha)$ and drawn as



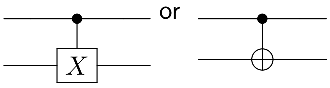
- Measurement of an observable A with result λ, drawn as



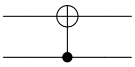
  If no observable A is specified, we mean σ_z, so we're measuring with respect to the computational basis

The following are examples of binary gates, that is quantum gates that act on a register of two qubits

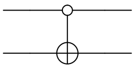
- CNOT (controlled not) controlling on the first qubit, $\Lambda^1(X) = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes X = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ and drawn as



- CNOT controlling on the second qubit, $\Lambda_1(X) = \mathbb{1} \otimes |0\rangle\langle 0| + X \otimes |1\rangle\langle 1| = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$ and drawn as



- CNOT controlling $|0\rangle$ instead of $|1\rangle$, $\Lambda^{|0\rangle}(X) = |0\rangle\langle 0| \otimes X + |1\rangle\langle 1| \otimes \mathbb{1} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ and drawn as

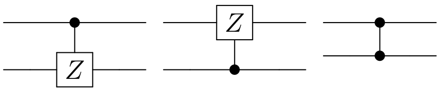


- Controlled Z controlled on some qubit. Note that

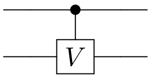
$$\Lambda^1(Z) = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes Z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$\Lambda_1(Z) = \mathbb{1} \otimes |0\rangle\langle 0| + Z \otimes |1\rangle\langle 1| = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

which means $\Lambda^1(Z) = \Lambda_1(Z)$. For this reason one we usually refer to this gate generically as CZ without specifying where the control is applied. As for the circuit, we use any of the following equivalently

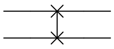


- Given a generic unary gate V, we can express controlled-V as $\Lambda^1(V) = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes V = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & v_{11} & v_{12} \\ 0 & 0 & v_{21} & v_{22} \end{pmatrix}$ and drawn as

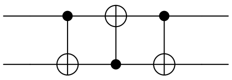


- Swap, expressed in matrix form as $S = |00\rangle\langle 00| + |11\rangle\langle 11| + |01\rangle\langle 10| + |10\rangle\langle 01| = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

  and drawn as

  Note that the Swap gate can be built as a composition of 3 CNOT gates as

Just like in the classical case, we can define rules to combine quantum gates. Given $U_1, \dots, U_k$ unitary operators, the set $F(U_1, \dots, U_k)$ of gates that can be obtained from $U_1, \dots, U_k$ is generated by the following rules:

1. $U_1, \dots, U_k \in F(U_1, \dots, U_k)$
2. $\mathbb{1}^{\otimes n} \in F(U_1, \dots, U_k)$
3. $V_1, V_2 \in F(U_1, \dots, U_k) \Rightarrow V_1 V_2 \in F(U_1, \dots, U_k)$
4. If $V_1, V_2 \in F(U_1, \dots, U_k)$ with $V_1 \in \mathcal{U}(\mathbb{H}^{\otimes n_1})$ and $V_2 \in \mathcal{U}(\mathbb{H}^{\otimes n_2})$, then $V_1 \otimes V_2 \in F(U_1, \dots, U_k)$

As in the classical case, we say that a set $\{U_1, \dots, U_k\}$ is universal if any unitary operator on $\mathbb{H}^{\otimes n}$ belongs to $F(U_1, \dots, U_k)$

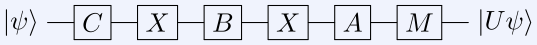
November 4^{th}, 2022

Theorem 4.4.2

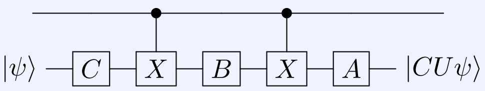
The set $\{M(\alpha), D_{\hat{y}}(\beta), D_{\hat{z}}(\gamma), \Lambda^1(X)\}_{\alpha,\beta,\gamma \in \mathbb{R}}$ is universal

Example 4.4.3. Given $U \in \mathcal{U}(\mathbb{C}^2)$, let's write $\Lambda^1(U)$ in terms of the operators from the previous theorem.

Recall that any such U can be decomposed as $U = e^{i\alpha} A \sigma_x B \sigma_x C$ with $A = D_{\hat{z}}(\beta) D_{\hat{y}}\left(\frac{\gamma}{2}\right)$, $B = D_{\hat{y}}\left(-\frac{\gamma}{2}\right) D_{\hat{z}}\left(-\frac{\delta+\beta}{2}\right)$ and $C = D_{\hat{z}}\left(\frac{\delta-\beta}{2}\right)$.
Note that we can build σ_x using a global phase, which means $\sigma_x \in F$, hence $U \in F$.
For U we have



Since we also have that $ABC = \mathbb{1}$, if we ignore the global phase given by M we can modify the circuit to obtain a circuit equivalent to $\Lambda^1(U)$ like this



To be more precise, in order to get a circuit that acts precisely like $\Lambda^1(U)$ we would need to add a gate of partial phase $P(\alpha)$. More details are given in the reference book.

Example 4.4.4. Let's now write a doubly-controlled version of U.

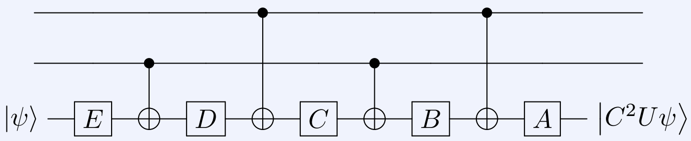
One way to achieve such a circuit is by using the more general decomposition of unitary operator that is

$$U = EXDXCXBXA$$

where the operators A, B, C, D and E are such that

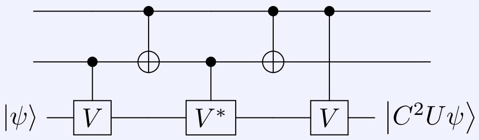
$$\begin{aligned} EDXCBXA &= \mathsf{Id} \\ EXDCXBA &= \mathsf{Id} \\ EDCBA &= \mathsf{Id} \end{aligned}$$

Then by building the following circuit



we would have found the solution. The issue is that to be formally correct, we would need to prove that such a decompositio is always possible, which is fairly hard.

Another approach is to prove that for any unitary U there exists another unitary V such that $V^2 = U$. This decomposition is easier to prove, and gives the solution



4.5 Quantum circuits

There are three types of circuits:

- a plain quantum circuit is a circuit where the used register is composed only of the input/output register, that is $\mathbb{H}^{\otimes n} = \mathbb{H}^{\mathsf{I/O}}$
- a circuit with ancilla is a circuit where the register has some I/O qubits and some auxillary qubits, that is $\mathbb{H}^{\otimes n} = \mathbb{H}^{\mathsf{I/O}} \otimes \mathbb{H}^{\otimes m}$
- a composite circuit is the composition of quantum circuits and classical operations

We will now give a more formal definition of these circuits

Definition 4.5.1 – Plain circuit

A plain circuit is a composition of L "elementary" gates $U_1, \dots, U_L \in \mathcal{U}(\mathbb{H}^{\mathsf{I/O}})$ which acts as $U = U_L \dots U_1$. It's said to be of depth L and if the system is initially in state $\rho \in D(\mathbb{H}^{\mathsf{I/O}})$, after applying U the state will be $U\rho U^*$

Before we give the definition of circuit with ancilla, we will prove an useful theorem

Theorem 4.5.2

Let $\mathbb{H}^{\mathsf{I/O}}, \mathbb{H}^W$ be Hilbert spaces. Let $|w_i\rangle, |w_f\rangle$ be states in $\mathbb{H}^W$ and let $\hat{U} \in \mathcal{U}(\mathbb{H}^{\mathsf{I/O}} \otimes \mathbb{H}^W)$ such that

$$\forall |\psi\rangle \in \mathbb{H}^{\mathsf{I/O}}, \qquad \hat{U}(|\psi\rangle \otimes |w_i\rangle) = (U|\psi\rangle) \otimes |w_f\rangle$$

for some fixed U. Then the map $|\psi\rangle \mapsto U|\psi\rangle$ is unitary on $\mathbb{H}^{\mathsf{I/O}}$ and if $\rho \in D(\mathbb{H}^{\mathsf{I/O}})$ we get

$$U\rho U^* = \mathsf{tr}^W \left(\hat{U}(\rho \otimes (|w_i\rangle\langle w_i|))\hat{U}^*\right)$$

Proof. First, we prove that U is a linear operator. It holds that

$$\begin{aligned}
[U(|\psi_1\rangle + |\psi_2\rangle)] \otimes |w_f\rangle &= \hat{U}(|\psi_1\rangle + |\psi_2\rangle) \otimes |w_i\rangle = \\
&= \hat{U}|\psi_1\rangle \otimes |w_i\rangle + \hat{U}|\psi_2\rangle \otimes |w_i\rangle = \\
&= (U|\psi_1\rangle) \otimes |w_i\rangle + (U|\psi_2\rangle) \otimes |w_i\rangle = \\
&= (U|\psi_1\rangle + U|\psi_2\rangle) \otimes |w_i\rangle
\end{aligned}$$

which implies linearity. Let's prove that U is unitary

$$\begin{gathered}
\|(U|\psi\rangle) \otimes |w_f\rangle\|_{\mathbb{H}^{\mathsf{I/O}} \otimes \mathbb{H}^W} = \|(U|\psi\rangle)\|_{\mathbb{H}^{\mathsf{I/O}}} \underbrace{\||w_f\rangle\|_{\mathbb{H}^W}}_{=1} = \|U|\psi\rangle\|_{\mathbb{H}^{\mathsf{I/O}}} \\
\| \\
\|\hat{U}|\psi\rangle \otimes |w_i\rangle\|_{\mathbb{H}^{\mathsf{I/O}} \otimes \mathbb{H}^W} = \||\psi\rangle \otimes |w_i\rangle\|_{\mathbb{H}^{\mathsf{I/O}} \otimes \mathbb{H}^W} = \||\psi\rangle\|_{\mathbb{H}^{\mathsf{I/O}}} \underbrace{\||w_i\rangle\|_{\mathbb{H}^W}}_{=1} = \||\psi\rangle\|_{\mathbb{H}^{\mathsf{I/O}}}
\end{gathered}$$

so U is also unitary. Finally, given $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$ with $p_j \geq 0$, $\sum_j p_j = 1$ and $\{\psi_j\}$ ONB, by linearity we have

$$\mathsf{tr}^W \left(\hat{U}(\rho \otimes |w_i\rangle\langle w_i|)\hat{U}^*\right) = \sum_j p_j \mathsf{tr}^W \left(\hat{U}(|\psi_j\rangle\langle\psi_j| \otimes |w_i\rangle\langle w_i|)\hat{U}^*\right)$$

which means that we only have to prove the identity on pure states. Indeed we have

$$\begin{aligned}
&\mathsf{tr}^W \left(\hat{U}(\underbrace{|\psi\rangle\langle\psi| \otimes |w_i\rangle\langle w_i|}_{=(|\psi\rangle\otimes|w_i\rangle)(\langle\psi|\otimes\langle w_i|)})\hat{U}^*\right) = \qquad\qquad \text{because } \begin{matrix}\hat{U}|\psi\rangle\otimes|w_i\rangle= \\ =(U|\psi\rangle)\otimes|w_f\rangle\end{matrix} \\
&= \mathsf{tr}^W ([(U|\psi\rangle) \otimes |w_f\rangle][(\langle\psi| U^*) \otimes \langle w_f|]) = \\
&= \mathsf{tr}^W ((U|\psi\rangle\langle\psi| U^*) \otimes |w_f\rangle\langle w_f|) = \\
&= U|\psi\rangle\langle\psi| U^*
\end{aligned}$$

where the last step holds because of the more general proposition

$$\mathsf{tr}^B (\rho_A \otimes \rho_B) = \rho_A$$

which one can prove as an exercise □

Definition 4.5.3 – Circuit with ancilla

A unitary $U \in \mathcal{U}(\mathbb{H}^{\mathsf{I/O}})$ is implemented by a quantum circuit with ancilla system $\mathbb{H}^W$ and states $|w_i\rangle, |w_f\rangle \in \mathbb{H}^W$ if there exists a plain circuit $\hat{U}$ on $\mathbb{H}^{\mathsf{I/O}} \otimes \mathbb{H}^W$ such that $\forall |\psi\rangle \in \mathbb{H}^{\mathsf{I/O}}$ it holds

$$\hat{U} |\psi\rangle \otimes |w_i\rangle = (U |\psi\rangle) \otimes |w_f\rangle$$

or equivalently

$$U\rho U^* = \mathrm{tr}^W \left(\hat{U}(\rho \otimes |w_i\rangle\langle w_i|)\hat{U}^* \right)$$

Definition 4.5.4

Given a function $f : \mathbb{N} \to \mathbb{N}$ we can write an implementation of f on $n \geq 1$ qubits, which is the unitary operator U_f such that

$$\begin{aligned} U_f : \ \mathbb{H}^n \otimes \mathbb{H}^n &\longrightarrow \mathbb{H}^n \otimes \mathbb{H}^n \\ |x\rangle^n \otimes |y\rangle^n &\longmapsto |x\rangle^n \otimes |y \boxplus f(x)\rangle^n \end{aligned}$$

where we interpret x and y both as integers and finite strings of bits.

Note that it's not necessarily the case that the image of a binary string of length n (in this case x) is also a binary string of length n. If it's smaller, this is no issue and to perform $y \boxplus f(x)$ we just pad $f(x)$ with leading zeros. If $f(x)$ is longer than n bits, then we cannot represent $y \boxplus f(x)$ in the register $\mathbb{H}^n$. In this case we usually resort to use $f(x) \mod 2^n$ instead of simply $f(x)$

Example 4.5.5. If we pick $f = \mathrm{Id}$, that is $f(x) = (x)$, then we use a slightly different notation $U_f = U_\boxplus$ which is the unitary $|x\rangle \otimes |y\rangle$ to $|x\rangle \otimes |y \boxplus x\rangle$

5 | Quantum algorithms

Quantum algorithms usually follow this general structure:

1. Prepare the input $|\psi\rangle \in \mathbb{H}^{\mathsf{I/O}}$
2. Implement U_f for some (problem dependant) $f : \mathbb{N} \to \mathbb{N}$
3. Do some "clever transformations"
4. Measure/observe the output

We'll show some examples of possible interpretations of these four steps in order to get familiar with quantum algorithms

1. Many quantum algorithms use the initial state $|0\rangle^n \otimes |0\rangle^m$, that is the input register is prepared in state $|0\rangle^n$ and the ancilla register is prepared in state $|0\rangle^m$. The reason behind this is that it's generally assumed that preparing a state of the computational basis (specifically the $|0\rangle$) is easy or at least possible (we will see that the issue of not being able to prepare a specific initial state will be a problem we will have to work around for some algorithms).

   Another commonly used initial state for the input register is

   $$|\varphi_0\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle^n$$

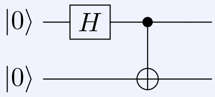
   which is useful because it represents a superposition where all the numbers from 0 to $2^n - 1$ have the same aplitude.

   Note that we can generate this state by applying $H^{\otimes n}$ to the initial state $|0\rangle^n$, where $H^{\otimes n}$ is the Hadamard gate on all the n qubits. Indeed we have

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

   Note that this state is in superposition, but it's not entangled (in fact we've just shown how to separate it)

Example 5.0.1. If we instead want to produce an entangled state we can consider the following circuit U with initial state $|00\rangle$



If we compute $U|00\rangle$ we get

$$\begin{aligned} U|00\rangle &= \Lambda^1(X)H\otimes\mathbb{1}|00\rangle = \\ &= \Lambda^1(X)\left(\frac{|0\rangle+|1\rangle}{\sqrt{2}}\otimes|0\rangle\right) = \\ &= \frac{|0\rangle\otimes|0\rangle+|1\rangle\otimes|1\rangle}{\sqrt{2}} = \\ &= |\Phi^+\rangle \end{aligned}$$

As an exercise, the reader can compute $U|01\rangle$, $U|10\rangle$ and $U|11\rangle$

The usefulness of this state comes from the fact that if we now apply some unitary U_f for $f:\mathbb{N}\to\mathbb{N}$ we get (assuming that the ancilla register has initial state $|0\rangle^m$)

$$U_f(|\varphi_0\rangle\otimes|0\rangle^m) = \frac{1}{\sqrt{2^n}}\sum_{x=0}^{2^n-1}|x\rangle\otimes|f(x)\rangle$$

which means that somehow we encoded or calculated all the possible outputs of f with only one evaluation of U_f. The problem is that it's not yet clear how to retrieve this all information from the superposition

4. An example of observable we can measure is $\Sigma_z^j = \mathbb{1}\otimes\cdots\otimes\mathbb{1}\otimes\sigma_z\otimes\mathbb{1}\otimes\cdots\otimes\mathbb{1}$, that is the observable that measures σ_z on the j-th qubit.

   Notice that if $j\neq j'$, then $[\Sigma_z^j,\Sigma_z^{j'}]=0$, that is they commute, which means that given a state $\rho\in D(\mathbb{H}^n)$ we can measure all the Σ_z^j for $j=1,\dots,n$ in any order, and obtain a "binary" string in $\{1,-1\}^n$ which becomes an actual binary string with the mapping $1\mapsto 0$, $-1\mapsto 1$. By doing so we get a binary string s which represents a number $x\in\{0,\dots,2^n-1\}$ which is also (maybe not surprisingly) the state in which the system currently is (more precisely, $|x\rangle^n$ is the current state)

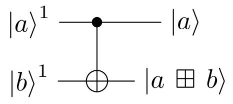
Remark 5.0.1 (On future measurements)**.** As we've just pointed out, when we measure all the Σ_z^j, there is a bijection between the string in $\{1,-1\}^n$ that we get from the measurement and the state $|s\rangle$ with $s\in\{0,1\}^n$ which is the state after the measurement. Of course, if we know one we know the other, and practically we will use most often the value of the state instead of the value of the measurement, so most of the times we will consider, as output of the measurment, the state itself, and not the string of eigenvalues.

Note that this can create confusion. Consider a system of one qubit and suppose we measure on the computational basis (that is we measure σ_z) and get "1". Because we don't specify what this "1" means, it both could be that we measured and

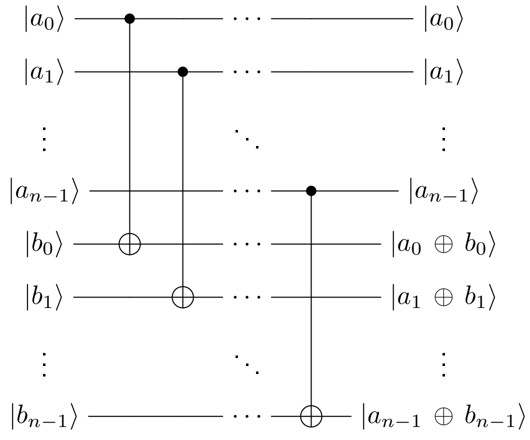
> obtained the eigenvalue "1", corresponding to the final state $|0\rangle$, or that we measured some eigenvalue (which a posteriori is -1) and we got the final state $|1\rangle$. To avoid this confusion, we will try to use the convention where just a number "n" means that n is the eigenvalue we obtained by the measurement, and "$|m\rangle$" means that m is the state we "obtained" by the measurement (ignoring the actual eigenvalue or string of eigenvalues).
>
> Furthermore, if we say that we "measure" without specifying the observable, it's meant as measuring with respect to the computational basis, so we're measuring all the Σ_z^j for that given register

November 9th, 2022

2. Consider the case $f = \mathsf{Id}$, so we want to implement $U_\boxplus$, that acts like $U_\boxplus |a\rangle \otimes |b\rangle = |a\rangle \otimes |a \boxplus b\rangle$. For $n = 1$ the implementation is quite simply the following circuit



For $n > 1$ it gets only slightlt more complicated, being



Suppose instead that we want to implement an operator U_+ such that

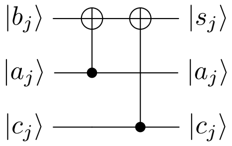
$$\begin{aligned} U_+ :\ \mathbb{H}^n \otimes \mathbb{H}^{n+1} &\longrightarrow \mathbb{H}^n \otimes \mathbb{H}^{n+1} \\ |a\rangle^n \otimes |b\rangle^{n+1} &\longmapsto |a\rangle^n \otimes |a+b\rangle^{n+1} \end{aligned}$$

for all $a, b \in \{0, \dots, 2^n - 1\}$. Note that even though (given the dimension of the second register) b could assume values bigger than $2^n - 1$, we can only require that U_+ acts as shown only for values of $b \leq 2^n - 1$. Otherwise, the addition might cause an overflow that wouldn't be representable in the second register for dimension-related issues. Actually, this is precisely the reason why we use a second register of dimension $n + 1$ instead of n, because the addition of two n-bit numbers can be an $(n + 1)$-bit number.

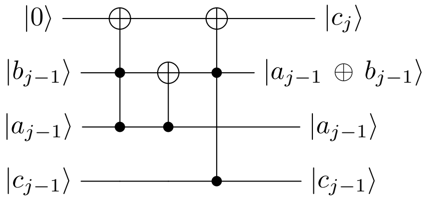
Let $a = \sum_{j=0}^{n-1} 2^j a_j$ and $b = \sum_{j=0}^{n-1} 2^j b_j$ with $a_j, b_j \in \{0, 1\}$.
Then we can write $a + b = \sum_{j=0}^{n-1} 2^j s_j + 2^n c_n$, c_n being the carry digit. We can find explicit values of s_j and c_n with the following recursive formulas

$$s_j = a_j \oplus b_j \oplus c_j$$
$$c_j = \begin{cases} 0 & \text{if } j = 0 \\ (a_{j-1} b_{j-1}) \oplus (a_{j-1} c_{j-1}) \oplus (b_{j-1} c_{j-1}) & \text{otherwise} \end{cases}$$

Let's implement U_+. First, we will implement two subcircuits (sometimes called routines) U_s and U_c that respectively calculate s_j and c_j. For U_s we have the following circuit

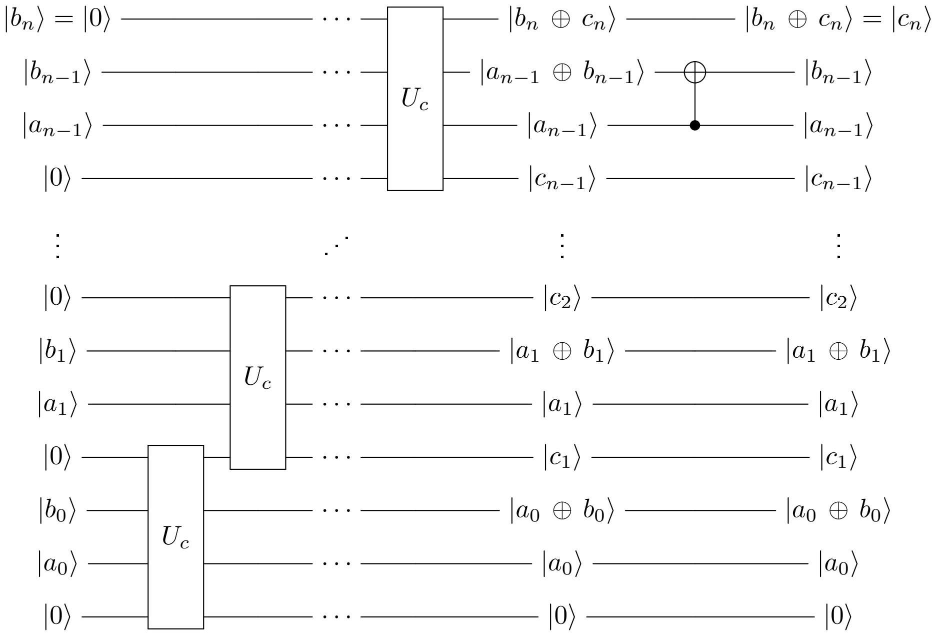


For computing c_j, we use a gate that also produces a side effect (namely changing $|b_{j-1}\rangle$ into $|a_{j-1} \oplus b_{j-1}\rangle$) as the resulting gate will be cheaper (that is, it will use fewer elementary gates). The gate U_c is

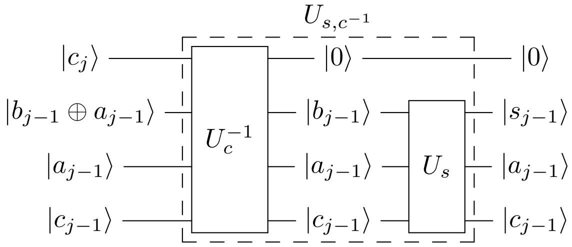


We're now ready to build the whole circuit. The initial state will be $|b\rangle^{n+1} \otimes |a\rangle^n \otimes |0\rangle^n$, but shuffled (that is, we're implying some Swap gates that only add visual complexity). The first step of the circuit will be the following

Step 1

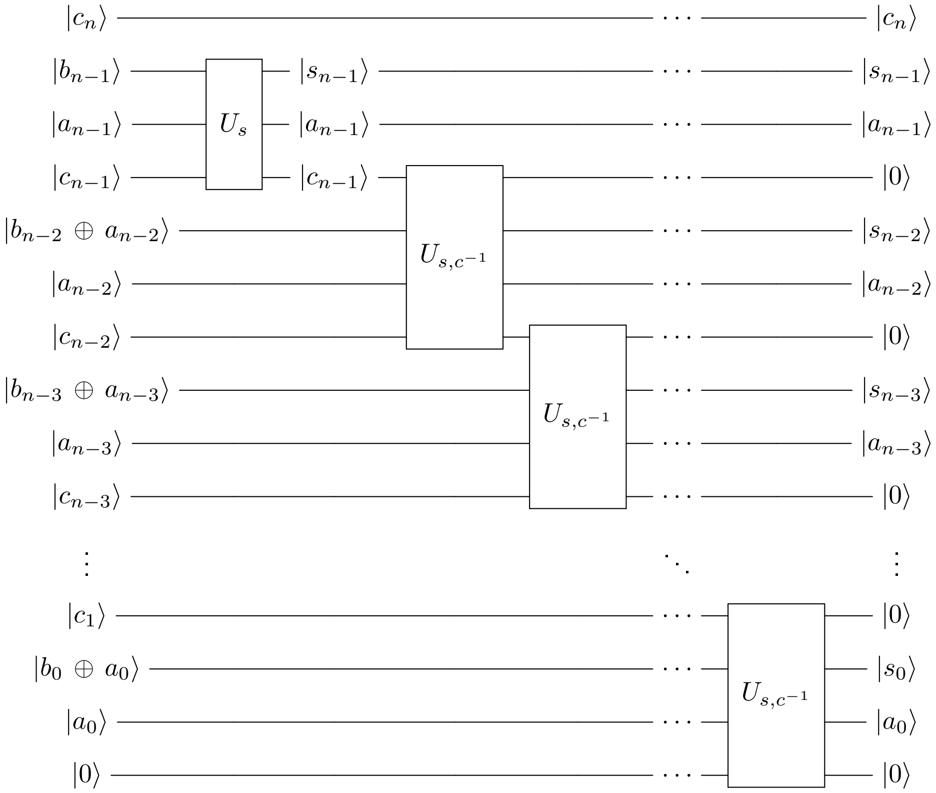


For the second step we first consider the action of the gate $U_{s,c^{-1}} = (U_s \otimes \mathbb{1})(U_c^{-1})$



The second step of the circuit is now defined as the following circuit

Step 2



The construction of this circuit requires $\mathcal{O}(n)$ elementary gates

To define some possible extensions of this building procedure, we first define $\mathbb{H}^{<N} \subset \mathbb{H}^{\otimes n}$ as $\mathbb{H}^{<N} = \text{Span}\{|a\rangle^n \mid a \in \{0, \dots, N-1\}\}$ Then, other operators that we could consider are

- $U_{+\%N} = U_{+ \mod N} : \mathbb{H}^{<N} \otimes \mathbb{H}^{<N} \to \mathbb{H}^{<N} \otimes \mathbb{H}^{<N}$ that sends $|a\rangle^n \otimes |b\rangle^n$ to $|a\rangle^n \otimes |a+b \mod N\rangle^n$
- $U_{- \mod N} = U^*_{+ \mod N}$
- $U_{\cdot c \mod N}$ that sends $|a\rangle \otimes |b\rangle$ to $|a\rangle \otimes |b + ac \mod N\rangle$
- $U_{b^\cdot \mod N}$ that sends $|a\rangle$ to $|b^a \mod N\rangle$

5.1 Quantum Fourier Transform

Remark 5.1.1 (Discrete Fourier Transform). Recall that the Discrete Fourier Transform (*DFT*) is the map $F : \mathbb{C}^N \to \mathbb{C}^N$ defined as

$$F(a_0, \dots, a_{N-1}) = (b_0, \dots, b_{N-1})$$

where

$$b_h = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} a_n \exp\left(2\pi i \frac{hn}{N}\right)$$

Note that $F^*F = \text{Id}_{\mathbb{C}^N}$

The fastest classical algorithm to compute the DFT is the Fast Fourier Transform (*FFT*) with complexity of $\mathcal{O}(N\log(N))$. From now on we will fix $N = 2^n$. This means that the DFT is an operator $F : \mathbb{C}^{2^n} \cong \mathbb{H}^{\otimes n} \to \mathbb{H}^{\otimes n}$. We would like to build it as a quantum circuit.

To define the corresponding quantum version of the DFT, the Quantum Fourier Transform (*QFT*), we need to understand the behaviour of the DFT on elements of the computational basis of $\mathbb{H}^{\otimes n}$, seen as elements of $\mathbb{C}^{2^n}$.
Let $k \in \{0, \dots, 2^n - 1\}$ and consider the element $|k\rangle$ of the computational basis. To apply the DFT to $|k\rangle$ we have to interpret $|k\rangle$ as a vector [1]. The corresponding vector in $\mathbb{C}^{2^n}$ is e_k, which means that when computing b_h all but the k-th terms of the sum are zero, and we get

$$b_h = \frac{1}{\sqrt{2^n}} \exp\left(2\pi i \frac{hk}{2^n}\right)$$

Because b_h is the coefficient of e_h, which corresponds to the state $|h\rangle$ in of the computational basis in the correspondence $\mathbb{C}^{2^n} \cong \mathbb{H}^{\otimes n}$, we can define the Quantum Fourier Transform as follows

Definition 5.1.1 – Quantum Fourier Transform

The Quantum Fourier Transform (*QFT*) is the operator $QFT : \mathbb{H}^{\otimes n} \to \mathbb{H}^{\otimes n}$ defined on the computational basis as

$$QFT\,|k\rangle = \frac{1}{\sqrt{2^n}} \sum_{h=0}^{2^n-1} \exp\left(2\pi i \frac{hk}{2^n}\right) |h\rangle$$

and extended by linearity

The matrix representation of the QFT is $QFT = (QFT_{k,h})_{k,h=0}^{2^n-1}$ where $QFT_{k,h} = \frac{1}{\sqrt{2^n}} e^{2\pi i \frac{kh}{2^n}}$.

Our goal is to express the QFT as a composition of elementary gates. To do so, we first introduce some notation and useful results

Definition 5.1.2

Given an integer $x \in \{0, \dots, 2^n - 1\}$ and its binary decomposition $x = \sum_{j=0}^{n-1} 2^j x_j$ we use the notation $[0.x_k \dots x_0]$ to express the fractional number that has as its binary digits (after the period) the last $k+1$ binary digits of x.
"$0.x_k \dots x_0$" would be the actual binary representation of the number $[0.x_k \dots x_0]$

One can easily notice that $[0.x_k \dots x_0] = 2^{-(k+1)} \sum_{j=0}^{k} 2^j x_j$ and also that $[0.x_k \dots x_0]$ is the fractional part (that is, the reminder modulo 1) of the value $2^{-(k+1)} \sum_{j=0}^{n-1} 2^j x_j = 2^{-(k+1)} x$. Similarly, the fractional part of $2^{k-n}x$ for $k < n$ is $[0.x_{n-k-1} \dots x_0]$

[1]: Be sure not to get confused. So far we've seen two ways of getting a vector out of a state $|k\rangle$ of the computational basis: either as a vector of the canonical base of $\mathbb{C}^{2^n}$ (which would be e_k of length 2^n), or as a vector of the binary digits of the number k, (which would be some $(k_{n-1}, \dots, k_0)$ of length n), and some confusion might arise given that (being it an element of the computational basis) the end result is always a binary vector, but with very different lengths and meanings. Right now we mean the vector $e_k \in \mathbb{C}^{2^n}$

Lemma 5.1.3

Given $|x\rangle^n$ element of the computational basis of $\mathbb{H}^{\otimes n}$, that is $x \in \{0, \dots, 2^n - 1\}$, and given its binary decomposition $x = \sum_{j=0}^{n-1} 2^j x_j$, we have

$$\begin{aligned} QFT\,|x\rangle^n &= \frac{1}{\sqrt{2^n}} \bigotimes_{k=0}^{n-1} \Big(|0\rangle + e^{2\pi i x 2^{-(k+1)}}\, |1\rangle \Big) = \\ &= \frac{1}{\sqrt{2^n}} \bigotimes_{k=0}^{n-1} \Big(|0\rangle + e^{2\pi i [0.x_k \dots x_0]}\, |1\rangle \Big) \end{aligned}$$

Proof. It holds

$$\begin{aligned} QFT\,|x\rangle^n &= \frac{1}{\sqrt{2^n}} \sum_{\xi=0}^{2^n-1} e^{2\pi i \frac{x\xi}{2^n}}\, |\xi\rangle = && \text{decomposing } \xi = \sum_{j=0}^{n-1} 2^j \xi_j \\ &= \frac{1}{\sqrt{2^n}} \sum_{\xi=0}^{2^n-1} \left(\prod_{j=0}^{n-1} e^{2\pi i x 2^j \xi_j 2^{-n}} \right) \left(\bigotimes_{j=n-1}^{0} |\xi_j\rangle \right) = && \text{by distributing the factors of the productory onto the tensor product} \\ &= \frac{1}{\sqrt{2^n}} \sum_{\xi=0}^{2^n-1} \bigotimes_{j=n-1}^{0} e^{2\pi i x \xi_j 2^{j-n}}\, |\xi_j\rangle = \\ &= \frac{1}{\sqrt{2^n}} \bigotimes_{j=n-1}^{0} \Big(|0\rangle + e^{2\pi i x 2^{j-n}}\, |1\rangle \Big) = \\ &= \frac{1}{\sqrt{2^n}} \bigotimes_{k=0}^{n-1} \Big(|0\rangle + e^{2\pi i x 2^{-(k+1)}}\, |1\rangle \Big) \end{aligned}$$

To conclude the proof we just notice that at the exponent in the last equality we only care about the fractional part of $x2^{-(k+1)}$, because any integer part, being it multiplied by $2\pi i$ as an exponent with base e, would end up being an additional factor of 1. □

To express the QFT as a composition of elementary gates, we first notice that the result of the lemma can be expressed as

$$\begin{aligned} QFT\,|x\rangle^n &= \frac{1}{\sqrt{2^n}} \bigotimes_{j=0}^{n-1} \left(|0\rangle + \underbrace{e^{2\pi i [0.x_j \dots x_0]}}_{=\prod_{k=0}^{j} \exp(2\pi i x_k 2^{j-k-1})} |1\rangle \right) = \\ &= \frac{1}{\sqrt{2^n}} \bigotimes_{j=0}^{n-1} \left(|0\rangle + \prod_{k=0}^{j} \exp(\pi i x_k 2^{j-k})\, |1\rangle \right) \end{aligned}$$

This means that to implement the QFT we just need to find a way to cleverly modify the phase of the qubits.

Notice that because the Hadamard gates acts as

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

we can express it more concisely (although in a more complicated way, which will be useful for our case) as

$$H|x_j\rangle = \frac{|0\rangle + e^{\pi i x_j}|1\rangle}{\sqrt{2}}$$

simply by the fact that $e^{\pi i \cdot 1} = -1$ and $e^{\pi i \cdot 0} = 1$.

To add the remaining phase to the qubit, remember the phase shift operator $P(\theta) = |0\rangle\langle 0| + e^{i\theta}|1\rangle\langle 1|$. If we imlpement a controlled version of $P(\theta)$, using the same $e^{\pi i x} = (-1)^x$ for $x \in \{0, 1\}$ that we used for the Hadamard gate, we would get an operator that acts as

$$|x_k\rangle \otimes \frac{|0\rangle + e^{i\alpha}|1\rangle}{\sqrt{2}} \mapsto |x_k\rangle \otimes \frac{|0\rangle + e^{i\alpha}e^{i\theta x_k}|1\rangle}{\sqrt{2}}$$

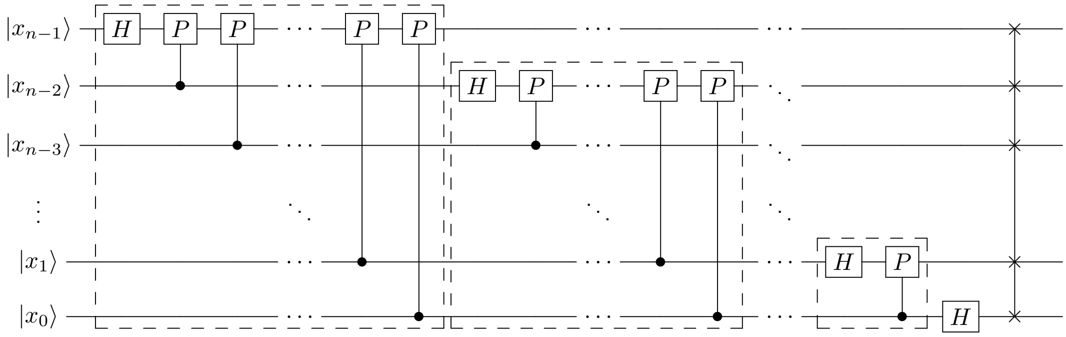
which means we have a way to add the remaining phases to the qubit, containing the value of the bits $x_{j-1}, \dots, x_0$.

Let $H_j = \mathbb{1} \otimes \dots \mathbb{1} \otimes H \otimes \mathbb{1} \otimes \dots \otimes \mathbb{1}$, that is the Hadamard gate applied to the j-th qubit, and let P_{jk} be the controlled version of the shift operator $P(\theta_{jk})$ acting on the j-th qubit with control on the k-th qubit, where $\theta_{jk} = \pi 2^{j-k}$. Then we can express the QFT as

$$QFT = S^{(n)} \prod_{j=0}^{n-1} \left(\prod_{k=0}^{j-1} P_{jk} \right) H_j$$

where $S^{(n)}$ means reversing the order of the qubits.

We can also express the QFT as the following circuit (with the omission of the indices on the gates P_{jk} which are obvious from the position of the gate and the position of the control)



November 16^{th}, 2022

5.2 Deutch's Problem

We'll now showcase an artificial problem to show how quantum algorithms can be extremely more efficient than classical algorithm. The problem is artificial in the sense that it's not actually a problem where the solution is useful for real life applications, more than it's useful to show the potential of quantum algorithms.

Suppose there existed a function $f : \{0,1\}^n \to \{0,1\}$ such that

- we have a way to compute f, but we don't have any analitic representation of f
- we know that f is either constant (which means either $f \equiv 0$ or $f \equiv 1$) or f is balanced (that is f assumes value 0 on exactly half of the inputs and 1 on the remaining half)

We want to determine with certainty whether f is constant or balanced, with the least number of evaluations possible.

The classical solution for this problem is able to give certainty only with $2^{n-1} + 1$ evaluations in the worst case.
The quantum solution for this problem, that is the Deutsch-Jozsa algorithm that we're now going to define, solves this problem with only 1 evaluation, which is an impressive improvement.

We've already seen that given a function $f : \{0,1\}^n \to \{0,1\}$, if we implement a gate U_f and apply it to the state $|\varphi_0\rangle^n \otimes |0\rangle$ we get the result

$$U_f |\varphi_0\rangle^n \otimes |0\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \otimes |f(x)\rangle$$

where the state $|\varphi_0\rangle^n = H^{\otimes n} |0\rangle^n$ is the (equal) superposition of all the elements of the computational basis.

Proposition 5.2.1

Given $x, y \in \{0, \dots, 2^n - 1\}$ and interpreting them as binary vectors of length n, that is $x, y \in \{0,1\}^n$, it holds

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

where $x^\intercal y = \sum_{i=0}^{n} x_i y_i$

Proof. By induction on n
$[n = 1]$ The thesis is $H|x\rangle = \frac{1}{\sqrt{2}}(|0\rangle + (-1)^x |1\rangle)$ for $x \in \{0,1\}$, which is trivially checked manually for $x = 0$ and $x = 1$

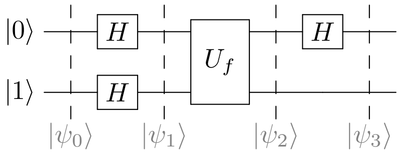
$[n \to n+1]$ Let $\tilde{x} \in \{0,1\}^{n+1}$, then by rewriting $\tilde{x} = 2^n x_n + x$ where $x_n \in \{0,1\}$,

$x \in \{0,1\}^n$, $H^{\otimes n+1} |\tilde{x}\rangle^{n+1}$ can be written as

$$\begin{aligned}
H^{\otimes n+1} |\tilde{x}\rangle^{n+1} &= \left(H \otimes H^{\otimes n}\right) |x_n\rangle \otimes |x\rangle^n = \\
&= \left(\frac{|0\rangle + (-1)^{x_n} |1\rangle}{\sqrt{2}}\right) \otimes \left(\frac{1}{\sqrt{2^n}} \sum_{y \in \{0,1\}^n} (-1)^{x^\intercal y} |y\rangle\right) = \\
&= \frac{1}{\sqrt{2^{n+1}}} \sum_{y \in \{0,1\}^n} \left((-1)^{x^\intercal y} |0\rangle \otimes |y\rangle + (-1)^{x^\intercal y + x_n} |1\rangle \otimes |y\rangle\right) = \\
&= \frac{1}{\sqrt{2^{n+1}}} \sum_{y \in \{0,1\}^n} \left((-1)^{x^\intercal y + x_n \cdot 0} |0\rangle \otimes |y\rangle + (-1)^{x^\intercal y + x_n \cdot 1} |1\rangle \otimes |y\rangle\right) = \\
&= \frac{1}{\sqrt{2^{n+1}}} \sum_{\tilde{y} \in \{0,1\}^{n+1}} (-1)^{\tilde{x}^\intercal \tilde{y}} |\tilde{y}\rangle
\end{aligned}$$

□

The Deutsch-Jozsa algorithm for $n = 1$ is defined as the following circuit



Let's analyse the intermediate states

$$\begin{aligned}
|\psi_0\rangle &= |0\rangle \otimes |1\rangle \\
|\psi_1\rangle &= \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}
\end{aligned}$$

Note that

$$U_f \left(|x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) = \frac{|x\rangle \otimes |f(x)\rangle - |x\rangle \otimes |1 \oplus f(x)\rangle}{\sqrt{2}}$$

so if $f(x) = 0$ we get $\frac{|x,0\rangle - |x,1\rangle}{\sqrt{2}}$, and if $f(x) = 1$ we get $\frac{|x,1\rangle - |x,0\rangle}{\sqrt{2}}$.
This means that we can write

$$U_f \left(|x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) = (-1)^{f(x)} \frac{|x,0\rangle - |x,1\rangle}{\sqrt{2}}$$

and we get

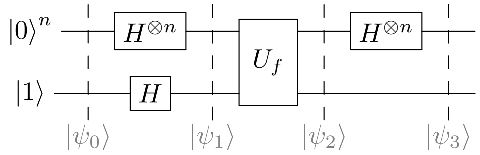
$$|\psi_2\rangle = \begin{cases} \pm \left[\frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}\right] & \text{if } f(0) = f(1) \\ \pm \left[\frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}\right] & \text{if } f(0) \neq f(1) \end{cases}$$

where the ± depends on the actual value of f but only adds up to a global phase, so it's irrelevant.
For $|\psi_3\rangle$ we get

$$|\psi_3\rangle = \begin{cases} |0\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} & \text{if } f(0) = f(1) \\ |1\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} & \text{if } f(0) \neq f(1) \end{cases}$$

This means that if we measure the first qubit and we get $|0\rangle$, the function is constant, and if we get $|1\rangle$ the function is balanced

The Deutsch-Jozsa algorithm for $n > 1$ is only slightly more complicated, and is given by the following circuit



Similarly we have the following partial states

$$
\begin{aligned}
|\psi_0\rangle &= |0\rangle^n \otimes |1\rangle \\
|\psi_1\rangle &= \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} |x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \\
|\psi_2\rangle &= \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} (-1)^{f(x)} |x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \\
|\psi_3\rangle &= \frac{1}{2^n} \sum_{x\in\{0,1\}^n} \sum_{y\in\{0,1\}^n} (-1)^{x^\intercal y + f(x)} |y\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}
\end{aligned}
$$

Note that in the last state, the amplitude of $|0\rangle^n$ in the first register is $\frac{1}{2^n}\sum_{x\in\{0,1\}^n}(-1)^{f(x)}$, which means it's ± 1 if the function is constant and $|0\rangle$ if the function is balanced. If we now measure the observable $|0\rangle^{n\,n}\langle 0|$ we can detect whether f is constant or balanced

5.3 Superdense coding

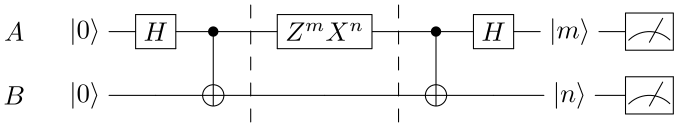
A more realistic problem where quantum algorithms allow us to be more efficient than the classical solution is in superdense coding, that is encoding and transmitting n classical bits in less than n qubits. Note that it is simply not possible to encode n classical bits in less than n bits (compression algorithms do exist but they either work on stronger assumptions or cannot guarantee that for any possible input the output will be less than n bits, just from a dimensional point of view).

Suppose we have a two qubit system in state $|\Phi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$. We then give the first qubit to Alice and the second qubit to Bob. Alice will perform "some operation" on her qubit to encode the information that she wants to send(in this case, two classical bits of information) on the global state, then give her qubit back to Bob. If the two bits that Alice wants to send are (m, n), the operation Alice will perform will be $Z^m X^n$, changing the global state according to the following table

Classical bits to encode	Operator to apply	Global state after operator
$(0,0)$	$\mathbb{1}$	$\frac{\|00\rangle+\|11\rangle}{\sqrt{2}} = \|\Phi^+\rangle$
$(0,1)$	X	$\frac{\|10\rangle+\|01\rangle}{\sqrt{2}} = \|\Psi^+\rangle$
$(1,0)$	Z	$\frac{\|00\rangle-\|11\rangle}{\sqrt{2}} = \|\Phi^-\rangle$
$(1,1)$	ZX	$\frac{-\|10\rangle+\|01\rangle}{\sqrt{2}} = \|\Psi^-\rangle$

Notice that the four possible outcomes form a basis (in particular, Bell's basis) which means by measuring with respect to this basis Bob can tell in which state the global system is, understand

which operation was applied and retrieve the bits. To simplify this, we can transform Bell's basis to the computational basis by applying a CNOT and an Hadamard gate. After this, the global system will be in the computational basis' state corresponding to the two bits of information that Alice wanted to encode, and this information can be obtained by Bob just by measuring with respect to the computational basis.
An example circuit, corresponding to the case of Alice encoding 10, is given by the following



where the barriers represent the moment where Alice recieves and sends her qubit. Note that the initial state for this circuit is actually $|0\rangle \otimes |0\rangle$, and the state $|\Phi^+\rangle$ is produced.

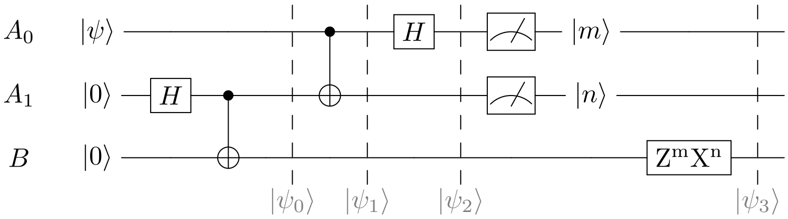
Exercise 5.3.1. One can prove that the reduced density matrix of Alice's qubit after her operation is the same independently of which operator she decides to apply, which means that an eavesdropper cannot obtain any information without modyfing the system

5.4 Teleportation

In the previous problem we were able to use an entangled system to send two bits of information by transmitting only one qubit. We will now consider a problem which can be seen as the inverse of the previous problem. We want to be able to send some state $|\psi\rangle$, and we will achieve this by using an entangled state and sending two bits of information.

Note that by the cloning theorem, we must destroy the original state in order to recreate it on a different register.

Suppose that Alice and Bob have each one qubit of an entangled pair in state $|\Phi^+\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$. Suppose that Alice also has a qubit in a state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ which she is not able to recreate, but would like to send to Bob (in this case "sending a qubit to Bob" means making so that Bob's qubit transforms in that state). Consider the following circuit



Again, the entangled state $|\Phi^+\rangle$ is produced from the state $|00\rangle$ rather than being in the initial state.

The partial states are

$$\begin{aligned}
|\psi_0\rangle =& |\psi\rangle \otimes \frac{|00\rangle + |11\rangle}{\sqrt{2}} \\
|\psi_1\rangle =& \frac{1}{\sqrt{2}} (\alpha |0\rangle \otimes (|00\rangle + |11\rangle) + \beta |1\rangle (|10\rangle + |01\rangle)) \\
|\psi_2\rangle =& \frac{1}{2}(|00\rangle (\alpha |0\rangle + \beta |1\rangle) + \\
& + |01\rangle (\alpha |1\rangle + \beta |0\rangle) + \\
& + |10\rangle (\alpha |0\rangle - \beta |1\rangle) + \\
& + |11\rangle (\alpha |1\rangle - \beta |0\rangle)) \\
|\psi_3\rangle =& |m\rangle \otimes |n\rangle \otimes |\psi\rangle
\end{aligned}$$

The reason why the final state has $|\psi\rangle$ on Bob's register is best manually checked: suppose for example that after the measurement, Alice measured $|mn\rangle = |10\rangle$. By checking the superpositon before the measurement $|\psi_2\rangle$, then we know that after Alice's register collapsed to $|10\rangle$, Bob's register must have collapsed to $\alpha |0\rangle - \beta |1\rangle$, which means that after applying Z, Bob's register is in state $\alpha |0\rangle + \beta |1\rangle = |\psi\rangle$.
One can check as an exercise that the other possible values of (m, n) yield the same final state with the opportune measurement

November 18^{th}, 2022

5.5 Shor's algorithm

Given a composite number N the goal of the factorization problem is to find a non-trivial factorization of that number, ie: find $p, q \in \mathbb{N}$ such that $pq = N$ and $p, q \neq 1$. The case of N even is trivial (assuming the number N is stored in any reasonable representation like binary, decimal, hexadecimal, ...) and there exist classical algorithms that efficiently solve the case of N being the power of a prime. For this reason we usually focus on the case of N being an odd composite prime that is the product of at least two distinct primes. Factorizing composite numbers is also at the core of many cryptographic primitives for protocols currently in use that assume that the factorization problem is difficult to solve. Indeed, the fastest classical algorithm currently known is the Numbered Field Sieve algorithm (*NFS*) which has a complexity of $\mathcal{O}(e^{c(\log N)^{\frac{1}{3}} \log\log N})$ (or, if you let n be the number of binary digits of N, you can express the complexity as $\mathcal{O}(e^{cn^{\frac{1}{3}} \log n})$; the number of binary digits, which is also the number of bits needed to store N, is the usual unit for calculating the complexity of such algorithms).

Shor's algorithm, proposed in 1994, requires only $\mathcal{O}((\log N)^3 \log\log N)$, using a register of $\mathcal{O}(\log N)$ qubits.
As a meter of comparison, consider that checking if a given number $b < N$ divides N requires $\mathcal{O}(\log N)$ steps, and computing the GCD of b and N (for example via Euclid's Algorithm) requires $\mathcal{O}((\log N)^2)$.

An important note is that Shor's algorithm is probabilistic, as in it has only a constant probability $p \in (0, 1)$ of returning a correct non-trivial solution. It is always possible, though, to run the algorithm multiple times (which would only add up to a moltiplicative factor in the complexity calculation) until the probability of retrieving a non-trivial factorization is at least $1-\varepsilon$ for some small value of $\varepsilon > 0$.

The key idea behind Shor's algorithm is to reduce the factorization problem to the problem of finding the period[2] r of a periodic function $f : \mathbb{N} \to \mathbb{N}$, which has an easy-to-implement solving algorithm in the quantum setting. The actual function f that we will find the period of is a function of the form $f_{b,N}(n) = b^n \mod N$, for some $b < N$. If $GCD(b, N) = 1$ then such a function is indeed periodic with period $r = \text{ord}_N(b)$, that is the order of b in the multiplicative group $(\mathbb{Z}/N\mathbb{Z})^*$.

The classical method to compute *a* period for such a function requires $\mathcal{O}(N)$ steps in the worst case (it requires calculating powers of $b \mod N$ until you find a repetition). On the other side, Shor's algorithm for the calculation of the period of a function computes the period of the function in only $\mathcal{O}((\log N)^3 \log\log N)$ steps (note that it has the same complexity as the whole of Shor's algorithm for factorization, as it is its most expensive part), which is an exponential improvement.

5.5.1 Definition and remarks

We can describe Shor's algorithm as composed of 3 steps, of which we will discuss implementation and correctness afterwards

- Input: an odd, composite integer N that is the product of at least two different primes
- Output: a non-trivial factorization of N
- Steps:
  1. **Choosing b**
     We start with a random value of $b < N$ and compute $GCD(b, N)$ (which requires $\mathcal{O}((\log N)^2)$ steps) with Euclid algorithms. If $GCD(b, N) \neq 1$ then we're done, as clearly $GCD(b, N)$ divides N and would give us a non-trivial factorization. If $GCD(b, N) = 1$, we procede with step 2
  2. **Quantum routine**
     Use a quantum algorithm (that we will describe later) to compute (with high probability) the period $r = \text{ord}_N(b)$ of $f_{b,N}$. If r is odd, we go back to step 1. Otherwise, we procede to step 3
  3. **Finding the factorization**
     Compute $GCD(b^{\frac{r}{2}} + 1, N)$. If the result is N, we go back to step 1. Otherwise the pair $GCD(b^{\frac{r}{2}} + 1, N), GCD(b^{\frac{r}{2}} - 1, N)$ gives a non-trivial factorization of N

As one can see from the many branches that require to restart the algorithm and the presence of a probabilistic subroutine, this algorithm is itself probabilistic.

The first issue we're going to address is the correctness of step 3. Once we get to step 3 we have values b and r such that $b^r \equiv 1 \mod N$ and r even. We can write $(b^{\frac{r}{2}})^2 - 1 \equiv 0 \mod N$ which can be factored as

$$(b^{\frac{r}{2}} + 1)(b^{\frac{r}{2}} - 1) \equiv 0 \mod N$$

which implies

$$N \mid (b^{\frac{r}{2}} + 1)(b^{\frac{r}{2}} - 1)$$

[2]: when we talk about "a" period of a periodic function $f : \mathbb{N} \to \mathbb{N}$ we mean any value k such that $\forall n \in \mathbb{N}$, it holds $f(n+k) = f(n)$. When we talk about "the" period of such a function f we mean the smallest possible value for such a k

Notice that since r is the (smallest) period, it must be $b^{\frac{r}{2}} - 1 \not\equiv 0 \mod N$. This implies that $N \not| (b^{\frac{r}{2}} - 1)$, which combined with $N \,|\, (b^{\frac{r}{2}} + 1)(b^{\frac{r}{2}} - 1)$ implies that N and $(b^{\frac{r}{2}} + 1)$ must have a common non-trivial factor, that is $GCD(b^{\frac{r}{2}} + 1, N) \neq 1$. If we also have that $GCD(b^{\frac{r}{2}} + 1, N) \neq N$ then $GCD(b^{\frac{r}{2}} + 1, N)$ is a non-trivial factor of N, and the remaining factor is exactly $GCD(b^{\frac{r}{2}} - 1, N)$.

About the likelihood of a random b resulting in a number coprime with N with an even period $r = \text{ord}_N(b)$, the following theorem (which we won't prove) reassures us

Theorem 5.5.1

Let $N = \prod_{j=1}^{k} p_j^{v_j}$ with $\{p_j\}$ distinct odd primes, with $k \geq 2$ and $v_j \geq 1$.
Let $\Omega = \{c \in \{0, \dots, N-1\} \,|\, GCD(c, N) = 1\}$. Then

- $|\Omega| = \phi(N) = \prod_{j=1}^{k} p_j^{v_j - 1}(p_j - 1)$
- $|\{b \in \Omega \,|\, r = \text{ord}_N(b) \text{ is even } \wedge N \not| (b^{\frac{r}{2}} + 1)\}| \geq \phi(N) \left(1 - \frac{1}{2^{k-1}}\right)$

Proposition 5.5.2

There exists a constant $c > 0$ such that $\frac{\phi(n)}{n} \geq \frac{c}{\log \log n}$ definitely in n

5.5.2 Quantum routine

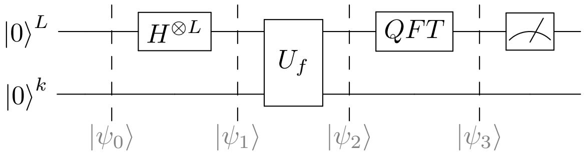
We'll now focus on step 2 of Shor's algorithm, which is the quantum routine for determining the period of a function of the form $f_{b,N}$ with $GCD(b, N) = 1$.

The problem we will solve is actually a bit more general, in that we will find the period r of an unknown function $f : \mathbb{N} \to \mathbb{N}$ with the following assumptions:

- $\exists L \geq 2$ such that $r < 2^{\frac{L}{2}}$, which means we have an upper bound for r. In our case we have $L = \lfloor 2 \log_2 N \rfloor + 1$
- f restricted to one period is injective and $\exists k \geq 1$ such that $\forall n$, $f(n) < 2^k$, which means we can store the values of $f(n)$ in k classical bits
- $U_f : \mathbb{H}^L \otimes \mathbb{H}^k \to \mathbb{H}^L \otimes \mathbb{H}^k$ is implementable with $\mathcal{O}(L^c)$ elementary gates. In our case $f_{b,N}$ requires $\mathcal{O}((\log N)^3)$

The algorithm we will see will find the period r with probability of at least $\frac{c'}{\log L}$ with at most $\mathcal{O}(L^{\max\{c,3\}})$ elementary gates, where c' is a constant that doesn't depend on anything, and is just cumbersome to write. To find the period r with probability of at least, for example, $\frac{1}{2}$, we need to repeat the algorithm s times where s is such that $(1 - \frac{c'}{\log L})^s \leq \frac{1}{2}$. As L grows, s needs to be proportional to $\log L$. This repetition is what gives the factor $\log \log N$ in the complexity of the algorithm.

The circuit for this routine is the following



For the partial states $|\psi_0\rangle$, $|\psi_1\rangle$ and $|\psi_2\rangle$ we get

$$\begin{aligned}
|\psi_0\rangle &= |0\rangle^L \otimes |0\rangle^k \\
|\psi_1\rangle &= \frac{1}{\sqrt{2^L}} \sum_{x=0}^{2^L-1} |x\rangle^L \otimes |0\rangle^k \\
|\psi_2\rangle &= U_f\,|\psi_1\rangle = \frac{1}{\sqrt{2^L}} \sum_{x=0}^{2^L-1} |x\rangle^L \otimes |f(x)\rangle
\end{aligned}$$

Since f is periodic we can rewrite $|\psi_2\rangle$ as

$$|\psi_2\rangle = \frac{1}{\sqrt{2^L}} \sum_{y=0}^{r-1} \sum_{j=0}^{J_y} |y+jr\rangle \otimes |f(y)\rangle$$

where, given $J = \lfloor \frac{2^L-1}{r} \rfloor$ being the total number of complete periods in $\{0,\dots,2^L-1\}$ and $R \equiv 2^L - 1 \mod r$, then J_y is defined as

$$J_y = \begin{cases} J+1 & \text{if } y \le R \\ J & \text{if } y > R \end{cases}$$

Practically speaking, J_y is the number of times the "class" of y appears in $\{0,\dots 2^L-1\}$ (and it might appear one more time than J if a representant of the class of y appears in the last incomplete period, which appens exactly if and only if $y \le R$)
For $|\psi_3\rangle$, remember that $QFT\,|x\rangle^L = \frac{1}{\sqrt{2^L}} \sum_{l=0}^{2^L-1} e^{2\pi i \frac{lx}{2^L}}\,|l\rangle$, which yields

$$|\psi_3\rangle = QFT \otimes \mathbb{1}_k\,|\psi_2\rangle = \frac{1}{2^L} \sum_{y=0}^{r-1} \sum_{j=0}^{J_y} \sum_{l=0}^{2^L-1} e^{2\pi i \frac{l(y+jr)}{2^L}}\,|l\rangle \otimes |f(y)\rangle$$

When we then measure on the register $\mathbb{H}^L$ the system will collapse to a state $|z\rangle$ of the canonical basis, where the probability of obtaining $|z\rangle$ is

$$\mathbb{P}(z) = \langle A_z \rangle_{|\psi_3\rangle} = \|A_z\,|\psi_3\rangle\,\|^2$$

Note that the observable is $A_z = |z\rangle\langle z| \otimes \mathbb{1}_k = A_z^2$ because it's a projection, so

$$A_z |\psi_3\rangle = \frac{1}{2^L} \sum_{y=0}^{r-1} \left(\sum_{j=0}^{J_y} e^{2\pi i \frac{z(y+jr)}{2^L}} \right) |z\rangle \otimes |f(y)\rangle$$

$$\begin{aligned}
\| A_z |\psi_3\rangle \|^2 &= \frac{1}{2^{2L}} \sum_{y=0}^{r-1} \left| \sum_{j=0}^{J_y} e^{2\pi i \frac{z(y+jr)}{2^L}} \right|^2 = \\
&= \frac{1}{2^{2L}} \sum_{y=0}^{r-1} \left| e^{2\pi i \frac{zy}{2^L}} \sum_{j=0}^{J_y} e^{2\pi i \frac{zjr}{2^L}} \right|^2 = \\
&= \frac{1}{2^{2L}} \sum_{y=0}^{r-1} \left| \sum_{j=0}^{J_y} e^{2\pi i \frac{zjr}{2^L}} \right|^2 = \qquad\qquad a = e^{2\pi i \frac{zr}{2^L}} \\
&= \frac{1}{2^{2L}} \sum_{y=0}^{r-1} \left| \sum_{j=0}^{J_y} a^j \right|^2
\end{aligned}$$

where in the first equality for $\| A_z |\psi_3\rangle \|^2$ we were able to take out the sum because the $f(y)$ are all different, which means the $|z\rangle \otimes |f(y)\rangle$ are all orthogonal. We find that

$$\mathbb{P}(z) \begin{cases} \frac{1}{2^{2L}} \sum_{y=0}^{r-1} (J_y + 1)^2 & \text{if } \frac{zr}{2^L} \text{ is an integer, that is } a = 1 \\ \frac{1}{2^{2L}} \sum_{y=0}^{r-1} \left| \frac{e^{2\pi i \frac{zr}{2^L}} - 1}{e^{2\pi i \frac{zr}{2^L}} - 1} \right|^2 & \text{otherwise} \end{cases}$$

It's still not clear how to extract the period r from this. To extract the value of r, we will use the follwoing theorem from the continuous fractions theory that guarantees that if we're able to obtain an integer z such that $\frac{zr}{2^L}$ is "sufficiently close" to be an integer, then it is possible to obtain the value of r.

Theorem 5.5.3

Given an integer z, suppose the exists an integer l such that

$$\text{A)} \qquad \left| \frac{zr}{2^L} - l \right| \leq \frac{r}{2^{L+1}}$$

then the (possibly reduced) fraction $\frac{l}{r}$ will appear in the fraction approximations of $\frac{z}{2^L}$ given by its continuous fraction representation (which we can manually calculate efficiently).
If it also hold that

$$\text{B)} \qquad GCD(l, r) = 1$$

then the fraction $\frac{l}{r}$ is unreducible, which means it will appear as-is in the fraction approximations and we will know the values of l (which is not interesting for us) and of r

To show that it's likely to get from the measurement a number z that satisfies that property, consider the following results (that we will not prove and take for granted):

- if z is such that $\frac{zr}{2^L}$ is an integer, that is $a = 1$, then $\mathbb{P}(z) \approx \frac{1}{2^{2L}} r \left(\frac{2^L}{r}\right)^2 = \frac{1}{r}$, which is reasonabily higher than the likelyhood given by a uniform distribution
- if z is such that there exists an integer l such that A) holds, then a similar bound $\mathbb{P}(z) \geq \frac{c}{r}$ holds for some fixed constant c

More precisely, the following lemma holds

Lemma 5.5.4

$$\sum_{\substack{z \text{ such that A)} \\ \text{and B) hold}}} \mathbb{P}(z) \geq \frac{c}{\log L}$$

November 23th, 2022

5.6 Quantum phase estimation

Consider the following problem: given U (that we don't know explictly) with eigenstate $|\varphi\rangle^n \in \mathbb{H}^{\otimes n}$ with eigenvalue $e^{2\pi i\theta}$, that is

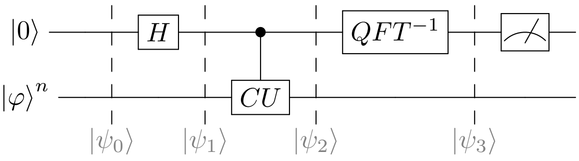
$$U|\varphi\rangle = e^{2\pi i\theta}|\varphi\rangle$$

we would like to retrieve at least an approximation of θ.

Remember that $|\varphi\rangle$ and $e^{2\pi i\theta}|\varphi\rangle$ belong to the same ray, so they give the same density operator $\rho = |\varphi\rangle\langle\varphi|$ and are physically indistinguishable.

The solution will assume that we we're not only able to evaluate U, but also a controlled version CU (which is not necessarily possible given U as a blackbox). The solution will use an ancilla register of t qubits to obtain (with high probability) the value $\lfloor 2^t\theta \rfloor$ from which we can get an approximation of θ to t binary digits.

For $t = 1$, the algorithm is described by the following circuit



The partial states are

$$\begin{aligned}
|\psi_0\rangle &= |0\rangle \otimes |\varphi\rangle^n \\
|\psi_1\rangle &= \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes |\varphi\rangle \\
|\psi_2\rangle &= \frac{|0\rangle \otimes |\varphi\rangle}{\sqrt{2}} + e^{2\pi i\theta}\frac{|1\rangle \otimes |\varphi\rangle}{\sqrt{2}} = \frac{|0\rangle + e^{2\pi i\theta}|1\rangle}{\sqrt{2}} \otimes |\varphi\rangle
\end{aligned}$$

Notice that for $t = 1$ we get $QFT^{-1} = H$ so we get

$$\begin{aligned}|\psi_3\rangle &= \frac{1}{\sqrt{2}}\frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes |\varphi\rangle + \frac{e^{2\pi i\theta}}{\sqrt{2}}\frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes |\varphi\rangle = \\ &= \left[\frac{1}{2}(1 + e^{2\pi i\theta})|0\rangle + \frac{1}{2}(1 - e^{2\pi i\theta})|1\rangle\right] \otimes |\varphi\rangle\end{aligned}$$

therefore when we measure the first qubit we obtain 0 with probability

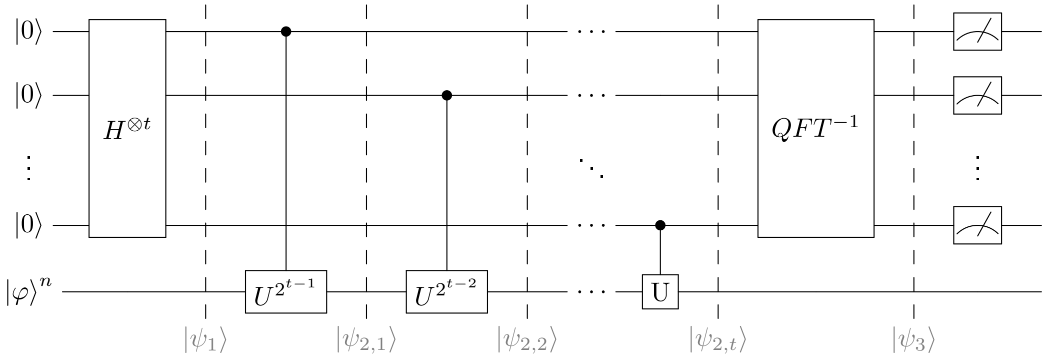
$$\mathbb{P}(0) = \left|\frac{1 + e^{2\pi i\theta}}{2}\right|^2 = \cos^2(\pi\theta)$$

and 1 with probability

$$\mathbb{P}(1) = \left|\frac{1 - e^{2\pi i\theta}}{2}\right|^2 = \sin^2(\pi\theta)$$

Note that $\mathbb{P}(0) > \mathbb{P}(1)$ if and only if $\theta \in \left[0, \frac{1}{4}\right] \cup \left[\frac{3}{4}, 1\right)$. So if θ has only one binary digit, that is if $\theta \in \left\{0, \frac{1}{2}\right\}$ then we obtain exactly θ.

The algorithm for the general case $t > 1$ is given by the following circuit



This circuit gives the following partial states

$$\begin{aligned}|\psi_1\rangle &= \frac{1}{\sqrt{2^t}}(|0\rangle + |1\rangle) \otimes \cdots \otimes (|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle) \otimes |\varphi\rangle \\ |\psi_{2,1}\rangle &= \frac{1}{\sqrt{2^t}}(|0\rangle + e^{2\pi i\theta 2^{t-1}}|1\rangle) \otimes (|0\rangle + |1\rangle) \otimes \cdots \otimes (|0\rangle + |1\rangle) \otimes |\varphi\rangle \\ |\psi_{2,2}\rangle &= \frac{1}{\sqrt{2^t}}(|0\rangle + e^{2\pi i\theta 2^{t-1}}|1\rangle) \otimes (|0\rangle + e^{2\pi i\theta 2^{t-2}}|1\rangle) \otimes \cdots \otimes (|0\rangle + |1\rangle) \otimes |\varphi\rangle \\ &\vdots \\ |\psi_{2,t}\rangle &= \frac{1}{\sqrt{2^t}}(|0\rangle + e^{2\pi i\theta 2^{t-1}}|1\rangle) \otimes \cdots \otimes (|0\rangle + e^{2\pi i\theta 2^{t-2}}|1\rangle) \otimes (|0\rangle + e^{2\pi i\theta}|1\rangle) \otimes |\varphi\rangle \\ &= \frac{1}{\sqrt{2}}\bigotimes_{j=0}^{t-1}\Big(|0\rangle + e^{2\pi i\theta 2^{t-(j+1)}}|1\rangle\Big) \otimes |\varphi\rangle \\ &= \frac{1}{\sqrt{2}}\bigotimes_{j=0}^{t-1}\Big(|0\rangle + e^{2\pi i(2^t\theta)2^{-(j+1)}}|1\rangle\Big) \otimes |\varphi\rangle\end{aligned}$$

Note that by a lemma on the QFT, if $2^t\theta$ is an integer then we get

$$|\psi_3\rangle = (QFT^{-1} \otimes \mathbb{1}_n)\,|\psi_{2,t}\rangle = \left|2^t\theta\right\rangle \otimes |\varphi\rangle$$

in which case we can just measure and obtain $2^t\theta$ and retrieve θ. That said, $2^t\theta$ might not be an integer, in which case we won't get that exact result.

To analize the general case, first consider that

$$\frac{1}{\sqrt{2^t}}\bigotimes_{j=0}^{t-1}(|0\rangle + e^{2\pi i(2^t\theta)2^{-(j+1)}}\,|1\rangle) = \frac{1}{\sqrt{2^t}}\sum_{\xi=0}^{2^t-1} e^{2\pi i\frac{(2^t\theta\xi)}{2^t}}\,|\xi\rangle$$

the proof of which is the reverse of the same lemma on the QFT used for the $2^t\theta$ integer case (in which part of the proof we never use that x is an integer, right until the last step that we don't need now).

Because the inverse of the QFT acts exactly as the QFT, except with a negative sign at the exponent, in the general case we get

$$|\psi_3\rangle = \frac{1}{2^t}\sum_{x=0}^{2^t-1}\sum_{\xi=0}^{2^t-1} e^{2\pi i\frac{(2^t\theta)\xi}{2^t}}\,e^{-2\pi i\frac{x\xi}{2^t}}\,|x\rangle^t \otimes |\varphi\rangle^n$$

which means that fixed any value $x \in \{0, \dots, 2^t - 1\}$ the probability of measuring x is

$$\mathbb{P}(x) = \frac{1}{2^{2t}}\left|\sum_{\xi=0}^{2^t-1} e^{2\pi i\frac{\xi}{2^t}(2^t\theta - x)}\right|^2$$

Using the same result from continuous fractions theory that we mentioned in Shor's algorithm, we know that there exists a unique $\bar{x}$ such that $\left|\frac{\bar{x}}{2^t} - \theta\right| \leq \frac{1}{2^{t+1}}$ that, if measured, would give us the approximation of θ to t bits. It can be proven that the probability of measuring this specific $\bar{x}$ is greater than an universal constant that is "sufficiently high" (somewhere close 40%).

Remark 5.6.1. Note that at the beginning of the algorithm we implicitly made the strong assumption that we were able to prepare the state $|\varphi\rangle$ being an eigenstate of the operator we're considering. It is not always the case that we can prepare an eigenstate of the operator, and we would like to ease the assumptions.

Suppose instead that we're able to prepare a state $|\varphi\rangle = \alpha\,|\varphi_a\rangle + \beta\,|\varphi_b\rangle$ where $|\varphi_a\rangle$ and $|\varphi_b\rangle$ are eigenstates for U, where

$$U\,|\varphi_a\rangle = e^{2\pi i\theta_a}\,|\varphi_a\rangle\,, \qquad U\,|\varphi_b\rangle = e^{2\pi i\theta_b}\,|\varphi_b\rangle$$

Following the same steps as before, and using linearity when needed, we get

$$
\begin{aligned}
|\psi_2\rangle &= \frac{1}{\sqrt{2^t}} \sum_{\xi=0}^{2^t-1} \alpha e^{2\pi i\theta_a} |\xi\rangle \otimes |\varphi_a\rangle + \beta e^{2\pi i\theta_b} |\xi\rangle \otimes |\varphi_b\rangle \\
|\psi_3\rangle &= \frac{1}{2^t} \sum_{x=0}^{2^t-1} \sum_{\xi=0}^{2^t-1} e^{2\pi i\left(\theta_a - \frac{x}{2^t}\right)\xi} \alpha |x\rangle \otimes |\varphi_a\rangle + e^{2\pi i\left(\theta_b - \frac{x}{2^t}\right)\xi} \beta |x\rangle \otimes |\varphi_b\rangle
\end{aligned}
$$

This means that the probability of measuring some integer l is

$$
\begin{aligned}
\mathbb{P}(l) &= \left\| \frac{1}{2^t} \sum_{\xi=0}^{2^t-1} e^{2\pi i\left(\theta_a - \frac{l}{2^t}\right)\xi} \alpha \otimes |\varphi_a\rangle + e^{2\pi i\left(\theta_b - \frac{l}{2^t}\right)\xi} \beta \otimes |\varphi_b\rangle \right\|^2 = |\varphi_a\rangle \perp |\varphi_b\rangle \\
&= \frac{1}{2^t} \left| \sum_{\xi=0}^{2^t-1} e^{2\pi i\left(\theta_a - \frac{l}{2^t}\xi\right)} \right|^2 |\alpha|^2 + \frac{1}{2^t} \left| \sum_{\xi=0}^{2^t-1} e^{2\pi i\left(\theta_b - \frac{l}{2^t}\xi\right)} \right|^2 |\beta|^2 \\
&= |\alpha|^2 \mathbb{P}_{\theta_a}(l) + |\beta|^2 \mathbb{P}_{\theta_b}(l)
\end{aligned}
$$

In the extreme case where both $\theta_a 2^t = l_a$ and $\theta_b 2^t = l_b$ are integers, then we would get l_a with probability $|\alpha|^2$ and l_b with probability $|\beta|^2$

5.6.1 Back to Shor's quantum routine

Now that we know how to perform QPE we can give a different interpretation to Shor's quantum routine for the period of a function. The following algorithm is more of a sketch than a rigorous algorithm, and has the purpose of showing a possible application of the QPE.
The hypothesis under which we were working are

- $\exists L \geq 2$ such that $r < 2^{L/2}$
- f restricted to one period is injective
- $\exists k \geq 1$ such that $\forall n, f(n) < 2^k$

Before we used an operator U_f that acted as $|x\rangle^L \otimes |y\rangle^k \mapsto |x\rangle^L \otimes |y \oplus f(x)\rangle^k$. Now we're going to use a differnet operator that acts as $|f(x)\rangle^k \mapsto |f(x+1)\rangle^k$ and we will also require to be able to build a controlled version CU_f. This is not a strong hypothesis because in the case of $f_{b,N}$, such operator would just be the operator that acts as $|x\rangle \mapsto |xb\rangle$ which is actually very easy to implement.

Consider the following state

$$
|\varphi\rangle = \frac{1}{\sqrt{r}} \sum_{y=0}^{r-1} |f(y)\rangle \in \mathbb{H}^k
$$

Of course, not knowing r, we can't easily prepare this state so the interest in it right now is purely theoretical, but the reason behind it will be clear soon.

Note that

$$\begin{aligned}
U|\varphi\rangle &= \frac{1}{\sqrt{r}}\sum_{y=0}^{r-1} U|f(y)\rangle = \\
&= \frac{1}{\sqrt{r}}\sum_{y=0}^{r-1} |f(y+1)\rangle = \\
&= \frac{1}{\sqrt{r}}\sum_{y=1}^{r} |f(y)\rangle = && f(r)=f(0) \\
&= \frac{1}{\sqrt{r}}\sum_{y=0}^{r-1} |f(y)\rangle
\end{aligned}$$

which means that $|\varphi\rangle$ is an eigenstate of U with eigenvalue 1.
Consider the following (still purely theoretical) class of states

$$|\varphi_s\rangle = \frac{1}{\sqrt{r}}\sum_{y=0}^{r-1} e^{-2\pi i\frac{sy}{r}}|f(y)\rangle, \qquad s\in\{0,\dots r-1\}$$

By repeating the same calculations as before we get that $|\varphi_s\rangle$ is an eigenstate of U with eigenvalue $e^{2\pi i\frac{s}{r}}$. If we were able to prepare a state $|\varphi_s\rangle$ for some s coprime with r, then by using QPE we could obtain $\frac{s}{r}$ and with the same trick from continuous fractions we could obtain r.

The solution to not bein able to prepare any of the states $|\varphi_s\rangle$ comes by observing that

$$\begin{aligned}
\frac{1}{\sqrt{r}}\sum_{s=0}^{r-1}|\varphi_s\rangle &= \frac{1}{r}\sum_{s=0}^{r-1}\sum_{y=0}^{r-1} e^{-2\pi i\frac{sy}{r}}|f(y)\rangle \\
&= \frac{1}{r}\sum_{y=0}^{r-1}\underbrace{\sum_{s=0}^{r-1} e^{-2\pi i\frac{sy}{r}}}_{=r\delta_{y0}}|f(y)\rangle = \\
&= |f(0)\rangle
\end{aligned}$$

This means that $|f(0)\rangle$ (which is Shor's case is $|1\rangle^k$ which is even easier to prepare) is a superposition of eigenstates of U, and that if we apply QPE built with this particular U to the initial state $|f(0)\rangle$ we will get as a result an approximation of the fraction $\frac{s}{r}$, for some integer $s\in\{0,\dots,r-1\}$ that is random because of the superposition. With a relatively high probability ($\phi(r)/r$, which by a previous lemma we know is a decent percentage) the random s corresponding to the approximation we obtained will be coprime with r, and will alow us to retrieve the value of r.

November 9^{th}, 2022

5.7 Grover's algorithm

The problem we want to solve now is the unstructured search problem, that is being able to find elements of a subset S (that we suppose satisfy a certain property) in a set X with no additional structure other than being enumerated. Note that with additional structure, for example if the set is ordered in some way related to the property that element in S satisfy, the problem becomes easier to solve. This is the most general version of the search problem.

In the unstructured case, the classical approach consist in bruteforcing the search resulting in a complexity of $\mathcal{O}(N)$ in the worst case (where $N = |X|$). Grover's algorithm for the unstructured search problem, although being a probabilistic algorithm, has a complexity of $\mathcal{O}(\sqrt{N})$ which is provably optimal.

We will suppose that $N = 2^n$ (if this is not the case we can just pretend that X has more elements than it actually does) and let $m = |S|$. With a bit of an abuse of notation, we will identify X with the set $\{0,1\}^n$ and will treat elements of X as binary string. This does not count as additional structure on X and has the only purpose of simplifying the notation. A solution without this identification would just insert the bijection $\{0,1\}^n \to X$ and its inverse many times in the definition.

We will also suppose that we have an oracle for S, that is we can implement and use a gate U_g where $g : \{0,1\}^n \to \{0,1\}$ is a boolean function that returns 1 if the input belongs to S, and 0 otherwise. Let's also call $S^\perp = X \setminus S$.

> **Remark 5.7.1.** The solution to this problem given these assumption might seem pointless, as one could find it unlikely that we can't easily find elements in S but we can easily have access to such a function g.
>
> Actually, such cases are very common. Consider this: we would like to use this algorithm to solve some different problem, making X the set of all the possible solutions and S the set of only the valid solutions (eg: for the 3 colouring problem, X would be the set of all the possible colourations regardless of them bein valid, and S would be the set of actual solutions). In many cases, solving a problem can be a much more difficult task than verifying if a given possible solution is an actual solution (famously, NP-Hard problems have a supposedly exponential solving algorithm but a polynomial verifying algorithm). This means that we can use the efficiency of the verifying algorithm to build an efficient solving algorithm.
>
> Note that this has no implication on P vs NP as the resulting algorithm will indeed have a quadratic improvement from bruteforce but will still be exponential in n

Given the gate

$$\begin{aligned} \widehat{U_g} :\ & \mathbb{H}^n \otimes \mathbb{H} \longrightarrow \mathbb{H}^n \otimes \mathbb{H} \\ & |x\rangle^n \otimes |y\rangle \longmapsto |x\rangle^n \otimes |y \oplus g(x)\rangle \end{aligned}$$

note that, if we set the second register to the state $\frac{|0\rangle - |1\rangle}{\sqrt{2}} = H|1\rangle$ then on a basis state $|x\rangle$ we get

$$\begin{aligned} \widehat{U_g}\,|x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} &= |x\rangle \otimes \frac{|g(x) - |1 \oplus g(x)\rangle\rangle}{\sqrt{2}} = \\ &= (-1)^{g(x)}\,|x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \end{aligned}$$

More generally, on a non-basis state $|\varphi\rangle = \sum_{x=0}^{2^n-1} \alpha_x |x\rangle$ we get

$$\widehat{U_g} |\varphi\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} = \sum_{x=0}^{2^n-1} (-1)^{g(x)} \alpha_x |x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

This means that $\widehat{U_g}$ reverses the phases of the states belonging to the solution set. For this reason, it makes sense to initialize the I/O register to $|\varphi_0\rangle = H^{\otimes n} |0\rangle^n$, but given that also the ancilla register was initialized to $H|1\rangle$, in practice we will actually use $|0\rangle^n \otimes |1\rangle$ as initial state and just apply an $H^{\otimes n+1}$ at the beginning of the circuit.

Notice also that $\widehat{U_g}$ acts as the identity on the ancilla qubit when the ancilla qubit is $\frac{|0\rangle - |1\rangle}{\sqrt{2}}$, reason for which it makes sense to consider the operator U_g as the operator that acts as $|x\rangle \mapsto (-1)^{g(x)} |x\rangle$ on the computational basis and extended by linearity. Because $\widehat{U_g}$ doesn't change the ancilla qubit from its initial state (because of the specific initial state, and not true in general), we will commit a bit of a notation abuse and for the sake of simplicity consider only U_g and the I/O register, implicitly saying that the ancilla register does exist and is and always remains in state $\frac{|0\rangle - |1\rangle}{\sqrt{2}}$. When more formality will be needed, we will resort to $\widehat{U_g}$ and will use in general the notation of putting a "hat" on operators acting on both I/O and ancilla registers, and not putting a hat on operators that act only on the I/O register.

The algorithm itself is not necessarily easy to understand, but it has a nice geometrical interpretation that gives much of the intuition needed to understand. We will first consider the case of $m = 1$, where there exist a unique solution.

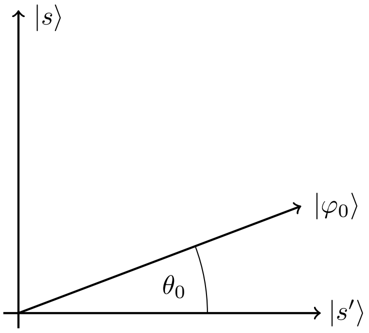
Let $s \in S$ be the unique solution. The initial state $|\varphi_0\rangle$ of the I/O register can be decomposed as a component on $|s\rangle$ and a component orthogonal to $|s\rangle$. More formally, we can define

$$|s'\rangle = \frac{1}{\sqrt{N-1}} \sum_{x \in S^\perp} |x\rangle$$

and get

$$\begin{aligned} |\varphi_0\rangle &= \frac{1}{\sqrt{N}} |s\rangle + \frac{\sqrt{N-1}}{\sqrt{N}} |s'\rangle \\ &= \sin\theta_0 |s\rangle + \cos\theta_0 |s'\rangle \end{aligned}$$

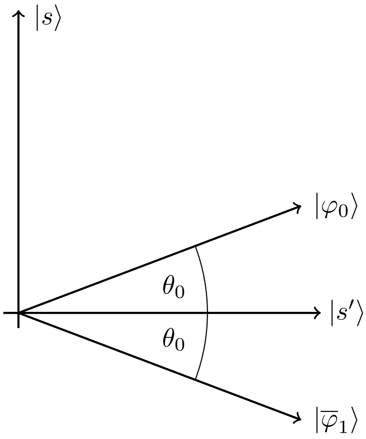
where $\theta_0 = \text{arcsin}\frac{1}{\sqrt{N}}$. Geometrically, we can consider the plane spanned by $|s\rangle$ and $|s'\rangle$ and consider $|\varphi_0\rangle$ on this plane



Note that so far, every basis state is equally likely. Our goal is to amplify the amplitude of the state $|s\rangle$ to make it more likely to measure. If we apply the oracle U_g we get

$$|\overline{\varphi}_1\rangle = U\,|\varphi_0\rangle = -\sin\theta_0\,|s\rangle + \cos\theta_0\,|s'\rangle$$

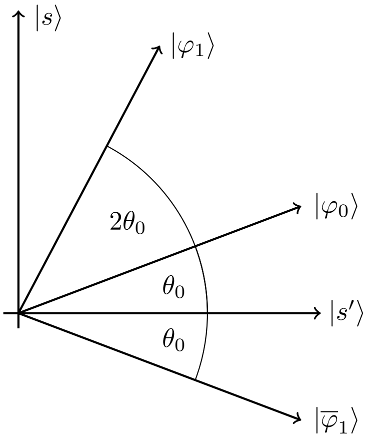
This means that applying U_g reflects a state with respect to $|s'\rangle$. Geometrically we get



To conclude this iteration of the algorithm, we can now get a new state obtained by the previous, reflected with respect to $|\varphi_0\rangle$. The reflection with respect to $|\varphi_0\rangle$ is the operator $R_{\varphi_0} = 2\,|\varphi_0\rangle\langle\varphi_0| - \mathbb{1}^{\otimes n}$, and we get

$$|\varphi_1\rangle = R_{\varphi_0}\,|\overline{\varphi}_1\rangle = \sin(2\theta_0 + \theta_0)\,|s\rangle + \cos(2\theta_0 + \theta_0)\,|s'\rangle$$

as one can clearly see from the geometric interpretation



It's clear to see that (supposing that $\theta_0 \leq \frac{\pi}{6}$ which is true for $N \geq 2m$) then the amplitude of $|s\rangle$ in $|\varphi_1\rangle$ is bigger than the amplitude in $|\varphi_0\rangle$, which means we made it more likely to measure $|s\rangle$.

The general case for a generic m is not much different, and requires only a few changes, namely:

- The "solution" state $|s\rangle$ is now defined as

$$|s\rangle = \frac{1}{\sqrt{m}} \sum_{x \in S} |x\rangle$$

- Similarly, $|s'\rangle$ orthogonal to $|s\rangle$ is defined as

$$|s'\rangle = \frac{1}{\sqrt{N-m}} \sum_{x \in S^{\perp}} |x\rangle$$

- The initial superposition state is now decomposed as

$$\begin{aligned} |\varphi_0\rangle &= \frac{1}{\sqrt{N}} \sum_{x \in X} |x\rangle = \\ &= \sqrt{\frac{m}{N}} |s\rangle + \sqrt{\frac{N-m}{N}} |s'\rangle = \\ &= \sin\theta_0 |s\rangle + \cos\theta_0 |s'\rangle \end{aligned}$$

$$\theta_0 = \text{arcsin}\sqrt{\frac{m}{N}}$$

The rest of the argument stays exactly the same, and by applying first U_g (which acts like a reflection with respect to $|s'\rangle$) and then R_{φ_0} we get

$$|\varphi_1\rangle = \sin(2\theta_0 + \theta_0) |s\rangle + \cos(2\theta_0 + \theta_0) |s'\rangle$$

Definition 5.7.1 – Grover operator

The Grover operator, at the core of Grover's algorithm, is defined as

$$\widehat{G_g} = (R_{\varphi_0} \otimes \mathbb{1})\widehat{U_g}$$

If we iterate this step, we get the following result, which can be proved by induction as an exercise

Proposition 5.7.2

By iterating Grover's operator we get

$$|\varphi_j\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} = \widehat{G_g}^{\,j} |\varphi_0\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

where

$$|\varphi_j\rangle = \sin\theta_j |s\rangle + \cos\theta_j |s'\rangle$$

and

$$\theta_j = (2j+1)\theta_0$$

Our hope is to iterate this step just the right number of times to maximise the amplitude of $|s\rangle$ in $|\varphi_j\rangle$. It's easy to see that the probability of measuring an element of S is

$$\mathbb{P}(S) = \sin^2\theta_j$$

and to maximise this probability we need $\theta_j \approx \frac{\pi}{2}$.

Lemma 5.7.3

Let

$$j_{\frac{N}{m}} = \left\lfloor \frac{\pi}{4\text{arcsin}\sqrt{\frac{N}{m}}} \right\rfloor$$

Then if we perform $j_{\frac{N}{m}}$ iterations of Grover, the probability of getting an element of S upon measurement is greater than $1 - \frac{m}{N}$

Proof. By definition it holds

$$j_{\frac{N}{m}} \leq \frac{\pi}{4\theta_0} < j_{\frac{N}{m}} + 1$$

and by multiplying by $2\theta_0$ we get

$$2\theta_0 j_{\frac{N}{m}} \leq \frac{\pi}{2} < 2\theta_0 j_{\frac{N}{m}} + 2\theta_0 = \theta_{j_{\frac{N}{m}}} + \theta_0$$

By adding θ_0 to the first inequality we get

$$\theta_{j_{\frac{N}{m}}} \leq \frac{\pi}{2} + \theta_0$$

which combined with the other inequality we get

$$\frac{\pi}{2} - \theta_0 < \theta_{j_{\frac{N}{m}}} \leq \frac{\pi}{2} + \theta_0$$

Since $\mathbb{P}(S) = \sin^2 \theta_{j_{\frac{N}{m}}}$ and $\theta_{j_{\frac{N}{m}}} > \frac{\pi}{2} - \theta_0$ we get

$$\begin{aligned} \mathbb{P}(S) = \sin^2 \theta_{j_{\frac{N}{m}}} &\geq \\ &\geq \sin^2\left(\frac{\pi}{2} - \theta_0\right) = \\ &= \cos^2(\theta_0) = \\ &= 1 - \sin^2(\theta_0) = \\ &= 1 - \frac{m}{N} \end{aligned}$$

□

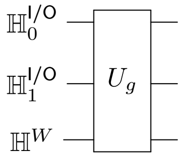
Note that (initially) any additional iteration after the $j_{\frac{N}{m}}$-th iteration actuall worsens. If we expand with a simple Taylor expansion we get $j_{\frac{N}{m}} = \mathcal{O}\left(\sqrt{\frac{N}{m}}\right)$, so the entirety of Grover's algorithm requires $\mathcal{O}\left(\sqrt{\frac{N}{m}}\right)$ applications of Grover operator. The complexity in terms of elementary gates remains unclear, as we haven't discussed how to implement Grover's operator using only elementary gates.

For this purpose, notice that $R_{\varphi_0} = H^{\otimes n} R_{|0\rangle} H^{\otimes n}$, which means it can be implemented with $2n+1$ elementary gates where $n = \log N$, so the actual cost of Grover's algorithm in terms of elementary gates is $\mathcal{O}\left(\sqrt{\frac{N}{m}} \log N\right)$ assuming that the oracle is a blackbox of constant cost.

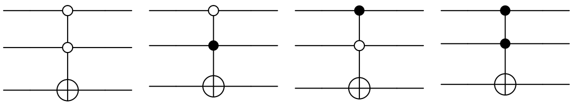
November 30^{th}, 2022

Let's implement Grover's algorithm for $N = 4$ and $m = 1$. We will actually analize two different methods of implementing Grover for this specific case.

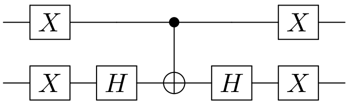
The first method we're going to consider is one where, consistently with how we've described Grover's algorithm, the oracle requires one ancilla qubit. Graphically we get



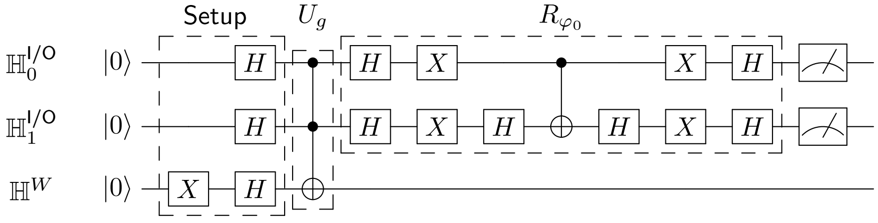
The search set X is formed by the basis states $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. For an actual implementation on something like qiskit we also have to implement the oracle that we would normally take for granted. To implement it, we need to build a gate that can detect one of those basis state. For example, for target states $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$ we can respectively use the following gates



For R_{φ_0} we will use the previously seen equivalent gate $H^{\otimes 2}(2\,|00\rangle\langle 00| - \mathbb{1}_2)H^{\otimes 2}$, and we can implement $2\,|00\rangle\langle 00| - \mathbb{1}_2$ with the following circuit



In this case we get $\theta_0 = \text{arcsin}\sqrt{\frac{1}{4}} = \frac{\pi}{6}$ which means that after $j_4 = 1$ iteration we get an angle of $\theta_1 = 3\theta_0 = \frac{\pi}{2}$, which means we get the target state with certainty instead of just with high probability. This happens everytime $\frac{N}{m} = 4$, for example if $N = 8$ and $m = 2$. The complete circuit (with target state $|11\rangle$) will be



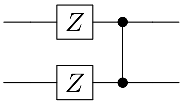
The second method we will consider for implementing Grover's for $N = 4$ and $m = 1$ will use a slightly different oracle that doesn't need an ancilla.

Suppose that the target state that we want to detect is $|11\rangle$. For this, we consider the oracle given by the matrix

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

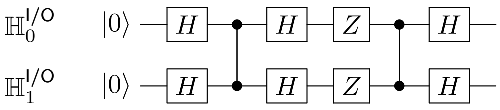
which is the matrix of the CZ gate.

Another way to optimize the circuit is by implementing a smarter version of the reflection $2\,|00\rangle\langle 00| - \mathbb{1}_2$ with the following equivalent circuit



which also makes use of the CZ gate. Note that this is equivalent to the given reflection only up to a global phase (it adds a -1 factor to the global phase), but a global phase won't change the correctness of the circuit.

The complete circuit becomes



5.8 Quantum counting

A really important assumption that we used in Grover's algorithm (particularly, to determine how many iterations to do) was knowing $m = |S|$. We would like now to solve this problem, that is being able to (approximally) count the number of elements in S.

We saw that Grover's operator, restricted to the plane spanned by $|s\rangle\,, |\varphi_0\rangle$, acted as a rotation of an angle $2\theta_0$, which is described by the matrix

$$\begin{pmatrix} \cos 2\theta_0 & -\sin 2\theta_0 \\ \sin 2\theta_0 & \cos 2\theta_0 \end{pmatrix}$$

which is a unitary operator with eigenvectors $\begin{pmatrix} i \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -i \\ 1 \end{pmatrix}$ and respective eigenvalues $e^{i\theta_0}$ and $e^{-i\theta_0}$. We've also studied the QPE algorithm, which allowed us to get an extimation of the phase of an eigenvalue of such operators. Supposing that we have access to the oracle that allows us to build Grover's operator, the idea behind quantum counting is to retrieve an approximation of θ_0 using QPE and from this value, obtain m as $m = N \sin^2 \theta_0$

December 2nd, 2022

5.9 Harrow-Hassidim-Lloyd algorithm

We will now study the Harrow-Hassidim-Lloyd (*HHL*) algorithm, originally proposed in 2008, which is used to find solutions of systems of linear equations.

More precisely, given $A \in \mathbb{C}^{N \times N}$ and $b \in \mathbb{C}^N$, we want to find $x \in \mathbb{C}^N$ such that $Ax = b$. Without any additional structure, classical algorithms for this general problem take $\mathcal{O}(N^3)$ steps. We will actualy focus on a more specific case that we will now describe

Remark 5.9.1 (Conditioning number)**.** The conditioning number of an invertible matrix $A \in \mathbb{C}^{N\times N}$ is given by

$$k(A) = \|A\|\|A^{-1}\|$$

which clearly depends on the choice of the norm.

If A is hermitian and positive, and we choose $\|\cdot\|_2$ as the norm, we get

$$k(A) = \frac{\lambda_{\max}}{\lambda_{\min}}$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are respectively the maximum and the minimun eigenvalues of A

Definition 5.9.1 – s-sparse matrix

Given a matrix $A \in \mathbb{C}^{N\times N}$, we say that A is s-sparse if every row of A contains at most s non-zero elements

The most efficient classical algorithm for the $Ax = b$ problem with A being hermitian, positive and s-sparse matrix is the conjugate gradient algorithm which takes $\mathcal{O}\left(Nsk(A)\log\frac{1}{\varepsilon}\right)$ to get a result closer than ε (in some norm) to an actual solution.

The HHL algorithm for $Ax = b$ with A being hermitian, positive and s-sparse takes only $\mathcal{O}(\log Ns^2(k(A))^2\frac{1}{\varepsilon})$. This looks like an exponential improvement, but to be completely fair the result given by the HHL algorithm is not exactly the same type of result given by other classical algorithms, in that it doesn't give a numerical solution but it is able to produce a state $\frac{|x\rangle}{\|x\|}$ with x being a solution, which allows us to evaluate observables on x without having a numerical expression for it. Still, in many applications we don't actually want the numerical solution, just some observable properties of the solution, in which case the result given by the HHL algorithm is good enough.

A recent result (Wassing, 2018) showed that if A is not sparse you can solve the problem on a quantum device in $\mathcal{O}(\sqrt{N}\log N(k(A))^2)$ steps.

The HHL algorithm can be summed up in two steps

1. **Hamiltonian Simulation**
   We use QPE on $U = e^{2\pi i A}$ (so we need an efficient implementation of U). We will take for granted how to obtain the operator $e^{2\pi i A}$ from the matrix A.

2. **Branch selection step**
   We measure the ancilla and according to the outcome we perform some quantum operations on the I/O

The I/O register will be $\mathbb{H}^{\otimes n_b}$ where n_b is chosen big enough to encode b. The ancilla register will be $\mathbb{H}^{\otimes t} \otimes \mathbb{H}$ where t will be the precision that we will use for the QPE and will influence the precision of the final result.
We will also make use of an inversion gate, defined on the computational basis as follows. Given $|l\rangle^t$

in the computational basis, the inversion gate acts as:

$$|l\rangle \otimes |0\rangle \mapsto |l\rangle \otimes \left(f(l)\,|1\rangle + \sqrt{(}1 - f^2(l))\,|0\rangle\right)$$
$$|l\rangle \otimes |1\rangle \mapsto |l\rangle \otimes \left(-\sqrt{1 - f^2(l)}\,|1\rangle + f(l)\,|0\rangle\right)$$

where the function f is defined as

$$f(l) = \begin{cases} \frac{2^t \lambda_{\min}}{l} = \frac{1}{l}c & \text{if } l > c \\ 0 & \text{otherwise} \end{cases}$$

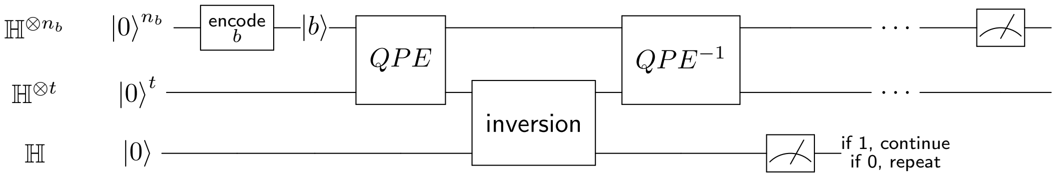
where $\lambda_{\min}$ is the minimum eigenvalue of A.

We will assume that $N = 2^{n_b}$ which means that $b = (\widetilde{b_j})_{j=0}^{N-1}$ and we can encode b to a state $|b\rangle$ defined as

$$|b\rangle = \sum_{j=0}^{N-1} \widetilde{b_j}\,|j\rangle$$

assuming that $\|b\| = 1$ (if this is not the case we can just rescale it and solve an equivalent problem)

We can now define the HHL algorithm as



After encoding the vector b, the global state is $|b\rangle \otimes |0\rangle^t \otimes |0\rangle$. Then we perform QPE with unitary $U = e^{2\pi i A}$ on $|b\rangle$, and to understand the result, note that by Bertrand's theorem we can decompose A as

$$A = \sum_{k=0}^{N-1} \lambda_k\,|\varphi_k\rangle\langle\varphi_k|$$

with $|\varphi_k\rangle$ eigenvectors and λ_k such that $U\,|\varphi_k\rangle = e^{2\pi i \lambda_k}\,|\varphi_k\rangle$. Because the eigenvectors form a base, we can also decompose the state $|b\rangle$ as

$$|b\rangle = \sum_{k=0}^{N-1} b_k\,|\varphi_k\rangle$$

After the QPE we get the following state

$$\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} \beta(l, \lambda_k) b_k\,|\varphi_k\rangle \otimes |l\rangle^t \otimes |0\rangle$$

where

$$\beta(l, \lambda_k) = \frac{1}{2^t}\sum_{j=0}^{2^t-1} e^{2\pi i\left(\lambda_k - \frac{l}{2^t}\right)j}$$

If we suppose that all the eigenvalues λ_k can be written as $2^t\lambda_k = l_k$ for some $l_k \in \{0,\dots,2^t-1\}$, which to be fair is a strong supposition, we can simplify the state as

$$\sum_{k=0}^{N-1} b_k \left|\varphi_k\right\rangle \otimes \left|2^t\lambda_k\right\rangle^t \otimes \left|0\right\rangle$$

When we apply the inversion we get

$$\sum_{k=0}^{N-1} b_k \left|\varphi_k\right\rangle \otimes \left|2^t\lambda_k\right\rangle^t \otimes \left(\frac{c}{2^t\lambda_k}\left|1\right\rangle + \sqrt{1-\left(\frac{c}{2^t\lambda_k}\right)^2}\left|0\right\rangle\right)$$

Then we apply QPE^{-1} we get

$$\sum_{k=0}^{N-1} \frac{c}{2^t\lambda_k} b_k \left|\varphi_k\right\rangle \otimes \left|0\right\rangle^t \otimes \left|1\right\rangle + \sqrt{1-\left(\frac{c}{2^t\lambda_k}\right)^2} b_k \left|\varphi_k\right\rangle \otimes \left|0\right\rangle^t \otimes \left|1\right\rangle$$

If we measure the "isolated" ancilla qubit, we get 1 with probability

$$\begin{aligned}\mathbb{P}(1) &= \left\|\sum_{k=0}^{N-1} b_k \left|\varphi_k\right\rangle \otimes \left|0\right\rangle^t \frac{c}{2^t\lambda_k}\right\|^2 = \\ &= \left\|\sum_{k=0}^{N-1} b_k \left|\varphi_k\right\rangle \otimes \left|0\right\rangle^t \frac{\lambda_{\min}}{\lambda_k}\right\|^2 = \\ &= \sum_{k=0}^{N-1} |b_k|^2 \left(\frac{\lambda_{\min}}{\lambda_k}\right)^2 \geq \\ &\geq \left(\frac{\lambda_{\min}}{\lambda_{\max}}\right)^2 = \\ &= k(A)^{-2}\end{aligned}$$

If we get 1, then the rest of the global state must be collapsed to

$$\begin{aligned}\frac{c}{2^t}\sum_{k=0}^{N-1} \frac{b_k}{\lambda_k}\left|\varphi_k\right\rangle \otimes \left|0\right\rangle^t &= \frac{c}{2^t}\left|A^{-1}b\right\rangle \otimes \left|0\right\rangle^t = \\ &= \frac{\left|A^{-1}b\right\rangle}{\|A^{-1}b\|} \otimes \left|0\right\rangle^t\end{aligned}$$

so the final global state will be $\frac{|x\rangle}{\|x\|} \otimes \left|0\right\rangle^t \otimes \left|1\right\rangle$

December 9^{th}, 2022

5.10 Quantum walks

The last class of problems we will consider are about quantum walks. The goal is to find and analyze a correspondent in the quantum setting of random walks. We'll first introduce classical random walks with all the related concepts and then we will jump to quantum walks.

Definition 5.10.1 – Undirected graph

A graph is a pair of sets (V, E) were V is a set of vertices or nodes, and E is a set of edges connecting two of the nodes in the set V.
We say that the graph:

- is undirected if the edges are undirected, that is "connects via an edge to" is a symmetric relation.
- is connected if every node is reachable from any other node only jumping through edges.
- has no multiple edges if given a pair of nodes, at most one edge connects them.
- has no loops if no edge connects a node to itself
- is d-regular if every node has exactly d neighbors

We will use d-regular, undirected, connected graphs with no multiple edges or loops. Examples of such graphs are cycles (2-regular) or n-dimensional hypercubes (n-regular)

Definition 5.10.2 – Adjacency matrix

Given a graph $G = (V, E)$ we can construct its adjacency matrix as the matrix $A = (A_{ij})_{ij}$ where

$$A_{ij} = \begin{cases} 1 & \text{if } v_i \text{ and } v_j \text{ are connected by an edge} \\ 0 & \text{otherwise} \end{cases}$$

In a classical setting, a random walk on G is a stocastic process without memory (Markov chain) where for each node i we have a discrete probability function that associates to each neighbor j of i a probability, which represents the probability of jumping to j when being in state i.

The natural way to represent these probability distributions is by the transition matrix

Definition 5.10.3 – Transition matrix

The (row) transition matrix is a matrix $P = (p_{ij})_{ij}$ where p_{ij} is the probability of j given by the discrete probability function associated to i. This matrix is clearly row-stocastic.

We can also define a column transition matrix which is the same as the previous, but transposed, which is also clearly column-stocastic.

Note that because the graph is undirected, the adjacency matrix A is symmetric, but the transition matrix P might not be. If we also suppose that the graph is d-regular and that all the probability functions are uniform, then P is symmetric and more precisely it holds $P = \frac{1}{d}A$.

In the contex of random walks, a position vector v is a vector of probabilities of size $|V|$ that represents the current state. If the position vector is a vector of the canonical basis (ie it has only one non-zero entry, and that entry is equal to 1) then we interpret it as knowing with certainty that the current position is the node correspondig to the non-zero entry. Otherwise, if it's a linear combination of elements of the canonical basis, we interpret is as giving us the probability of the current position being on the respective node.

By applying P we get the position vector after 1 step of the random walk. By iterating steps, the process typically converges to a stationary distribution (that is a uniform distribution on all the nodes), although this is not always the case (for example, consider the case of a cycle with an even number of nodes). We will assume that this is always the case for the sake of simplicity.

We now want to transpose the concept of random walks to the quantum domain. There are mainly two ways to formalize this:

- Coined quantum walks (Aharonov, 1993). This type of quantum walks is more well suited for d-regular graphs
- Szegdy quantum walks (Szegdy, 2004), which is a more general approach. In this approach instead of using a position state to represent nodes, we use a position state to represent edges

Note that we've only discussed (and will only discuss) the case of discrete random walks, that is the case where the position can only change discretely. There are also continuous random walks.

5.10.1 Coined quantum walks

Let $G = (V, E)$ be a d-regular, undirected, connected graph with no multiple edges or loops. Let $|V| = N = 2^n$. We will need a register $\mathbb{H}_T = \mathbb{H}_c \otimes \mathbb{H}_p$ where the first subregister is called the "coin space" and the second subregister is called the "position space". A basis state in $\mathbb{H}$ is $|k,p\rangle = |k\rangle \otimes |p\rangle$ which is interpreted as representing a position on node $p \in \{0, \ldots, N-1\}$ with a coin toss resulting in k. On a cycle graph of length $N = 2^n$ we have $\mathbb{H}_c = \mathbb{H}$, that is only one qubit, and $\mathbb{H}_p = \mathbb{H}^{\otimes n}$.

In the classical setting we used the transition matrix as operator to get to the next step of the walk. In this context, we will make use of two operators, the coin operator and the shift operator.

For the coin operator one usually chooses the Hadamard gate, as it gives (starting from a basis state) equal probability to every (basis state) outcome. Formally

$$\begin{aligned} C : \ \mathbb{H}_T &\longrightarrow \mathbb{H}_T \\ |k,p\rangle &\longmapsto |Hk,p\rangle \end{aligned}$$

that is $C = H \otimes \mathbb{1}$ (without specifying the dimensions of H or $\mathbb{1}$). The shift operator models the concept of "jumping to the next node". It's defined as

$$\begin{aligned} S : \ \mathbb{H}_T &\longrightarrow \mathbb{H}_T \\ |k,p\rangle &\longmapsto \left|k, p + (-1)^k \mod N\right\rangle \end{aligned}$$

Note that some authors also impose that the shift operator must also flip the coin, that acting as a NOT on the first register.

The walk operator is given by the composition $W = SC$ and each application of W models a single quantum walk step. This is sometiems referred as the "flip-flop quantum walk".

Remark 5.10.1. The reason for why we also need a coin space beside the position space is that in quantum algorithms we can only operate with unitary operators, which are necessarily invertible hence not random. This means that we can't just randomly go from a position vector to another position vector, but we somehow need to keep track of the randomness (the coin toss) so that the operation is at least invertible.

We can define this process as an algorithm, like the following:

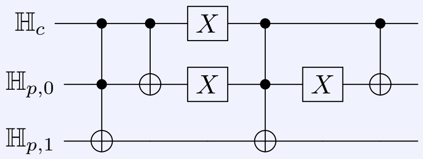
1. Prepare the initial state
2. Repeat k times
   - 2.1 Apply C
   - 2.2 Apply S
3. Measure

Note that there is no "randomness" in the body of the algorithm, as everything happening in step 2 is deterministic. The only source of randomness is the measure, which makes the system collapse to some possibly random state.

Another important observation is that since every step is unitary, we can't possibly approach some limit stationary distribution as unitary operators preserver distances. For this reason we usually use the notion of "limiting distribution" defined as the limit of partial time averages

Example 5.10.4. Let's implement a circuit for quantum walks for a 4-cycle graph. In this case we have $\mathbb{H}_T = \mathbb{H}_C \otimes \mathbb{H}_p$ where $\mathbb{H}_c = \mathbb{H}$ and $\mathbb{H}_p = \mathbb{H}^{\otimes 2}$.

The coin operator is $C = H \otimes \mathbb{1}$. For the shift operator, one way of implementing it is the following circuit



For example, with an initial state of $|0\rangle \otimes |00\rangle$ we get

$$\begin{aligned} |0\rangle \otimes |00\rangle &\mapsto |0\rangle \otimes |00\rangle \mapsto \\ &\mapsto |0\rangle \otimes |00\rangle \mapsto \\ &\mapsto |1\rangle \otimes |10\rangle \mapsto \\ &\mapsto |1\rangle \otimes |11\rangle \mapsto \\ &\mapsto |1\rangle \otimes |01\rangle \mapsto \\ &\mapsto |1\rangle \otimes |11\rangle \end{aligned}$$

5.10.2 Quantum walks on an n-dimensional hypercube

Definition 5.10.5 – n-dimensional hypercube graph

An n-dimensional hypercube graph is a graph where the set of nodes is the set of binary strings of length n, and two nodes are adjacent if they have Hamming-distance 1 (ie they differ only by one character)

With an n-dimensional hypercube we have $N = 2^n$ and $\mathbb{H}_T = \mathbb{H}_c \otimes \mathbb{H}_p$ with basis $\{|a, v\rangle \mid 0 \leq a \leq n-1,\ v \in \{0,1\}^n\}$. We interpret the result $a \in \{0, \ldots, n-1\}$ of the coin toss as which digit (0-indexed) of the current node we have to change to get to the next node.

This means that the shift operator can be defined on a basis as $S\,|a, v\rangle = |a, v \boxplus e_a\rangle$ where e_a is the a-th element of the canonical basis of $|0,1\rangle^n$. As for the coin operator, this time we will use a Grover reflection, that is

$$G = \frac{2}{n} u u^\intercal - \mathbb{1}$$

where $u = (1, \ldots, 1)^\intercal$. As before we have $W = SG$. This example, that we're referring to as an example of a coined quantum walk, is actually a specific case of a Szegedy quantum walk. In general, Szegedy quantum walks on d-regular graphs become equivalent to a coined Grover quantum walks.

The reason why we're discussing quantum walks on the hypercube is because we can use them to implement a search algorithm on the nodes of the hypercube, as follows.

Let M be a set of marked nodes and $m = |M|$. We start the quantum walk on a "suitable state" (which could be a node or a superposition of nodes) and we perform some steps until we "reach" a marked node. The interesting fact is that we can estimate a priori how many steps we have to take to have a probability of landing on a marked node.

As we've mentioned, in this formalism the position state represent an edge instead of a node. More precisely, we will need two subregisters: one for the current node and one for the previous node. Together they identify an edge. We can define the state

$$|p_x\rangle = \sum_y \sqrt{p_{xy}}\,|y\rangle$$

which is a superposition of all the neighbors of x, each with an amplitude corresponding to the (square root of the) associated discrete probability. We can then define

$$|G\rangle = \frac{1}{\sqrt{m}} \sum_{x \in M} |x\rangle \otimes |p_x\rangle$$

which is a superposition of all the marked ("good") states. Similarly we define

$$|B\rangle = \frac{1}{\sqrt{N-m}} \sum_{x \notin M} |x\rangle \otimes |p_x\rangle$$

Now we set $\varepsilon = \frac{m}{N}$ and $\theta = \text{arcsin}\sqrt{\varepsilon}$ so that we can write a superposition of all the edges as

$$|U\rangle = \frac{1}{\sqrt{N}} \sum_x |x\rangle \otimes |p_x\rangle = \sin\theta\,|G\rangle + \cos\theta\,|B\rangle$$

Just like we did for Grover, we can define a quantum walk with initial state $|U\rangle$ where we repeat $\mathcal{O}(\frac{1}{\sqrt{\varepsilon}})$ times the following step routine:

1. apply reflection with respect to $|B\rangle$
2. apply reflection with respect to $|U\rangle$

Finally, we measure. Note that to apply a reflection with respect to $|B\rangle$ we need a phase oracle (much like Grover)