
Generant Imatges amb un Ordinador Quàntic

TREBALL DE RECERCA DE BATXILLERAT
IES MIQUEL TARRADELL

Autor:

Tomàs Ockier Poblet
2nd Batxillerat
ockier1@gmail.com

Tutor:

Tomàs Ockier Poblet



Institut Miquel Tarradell

3 d'octubre de 2021
Barcelona, Barcelona

Índex

I	Marc Teòric	6
1	Àlgebra Lineal	7
1.1	Vectors i Espais Vectorials	7
1.2	Operadors Lineals	10
1.1	Tipus d'Operadors Lineals	11
1.3	Producte Interior i Producte Exterior	12
1.1	Producte Interior	12
1.1.1	Properties of the Inner Product	14
1.2	Orthonormal and orthogonal vectors	14
1.3	Outer product	16
1.4	Tensor product	16
1.1	Properties of the Tensor Product	19
1.5	Trace	19

<i>ÍNDEX</i>	2
2 Quantum Computation	21
2.1 Quantum State and Superpositions	21
2.2 Qubits i operacions quàntiques	24
2.1 Representació geomètrica d'un qubit	26
2.2 Operacions per a només un qubit	27
2.3 Operacions per a múltiples qubits	28
2.3 Mesurament quàntic	28
II Part Experimental	30
III Conclusions	31
Appendices	33
A More Linear Algebra	34
A.1 Gram–Schmidt Procedure	34
A.2 Dirac Notation Crash Course	36
A.3 More on the Partial Trace	37
B Quantum Computation vs Quantum Mechanics	38
B.1 Normalizing	38
C Polarization of a photon	41

Introducció

Desde hace más de un año, me he dedicado a estudiar computación cuántica durante mi tiempo libre. Buscaba investigar un campo relacionado con la mecánica cuántica, pero sin que sea muy complicado, que se pueda entender a un nivel teórico y que me entusiasme.

La Computación Cuántica encaja perfectamente con esos criterios. Es más sencilla que la mecánica cuántica debido a que no está basada en cálculo o ecuaciones diferenciales, se basa en la álgebra lineal, utilizando valores discretos, vectores y matrices. Además si se trabaja a un nivel teórico sencillo, no se tienen en consideración las interpretaciones físicas, lo cual simplifica mucho las cosas. Cuanto más me adentraba, más ganas tenía de seguir.

Mi parte favorita de este campo es el Quantum Machine Learning que consiste en diseñar y aplicar conceptos de Machine Learning a los ordenadores cuánticos, como por ejemplo implementar cuánticamente las famosas Redes Neuronales, que están detrás de la mayoría de inteligencias artificiales que vemos hoy en día [1].

QML es un campo de investigación joven y en crecimiento debido a que sus algoritmos son ideales para implementarlos con los ordenadores cuánticos actuales, los cuales no son muy potentes. Ejemplos de estas implementaciones serían [insertar aplicaciones aquí], etc.

De entre todos los tipos de algoritmos me he centrado en las Redes Neuronales Cuánticas, análogas cuánticas de las Redes Neuronales tan utilizadas hoy en día para hacer gran variedad de tareas. Me he interesado particularmente en ellas debido a que tenía experiencia en el pasado con las RNs clásicas y había visto que existen frameworks de software para trabajar con ellas como TensorFlow Quantum [2] que me podían ayudar.

Para adentrarme en el campo de QML, he tenido que adquirir conocimientos en álgebra lineal, cálculo y física. Dentro de QML en concreto me he dedicado a leer papers que me interesan y en un par de ocasiones intentar implementar los algoritmos detallados en esos papers. Puede parecer algo imposible en principio debido a que no tengo acceso directo a un ordenador cuántico, no obstante estos no son necesarios debido a que las operaciones cuánticas pueden ser simuladas en un ordenador corriente de escritorio (con ciertas limitaciones). Pero puedo tener acceso a ordenadores cuánticos ya que IBM permite acceder a los

suyos mediante IBM Quantum Experience [3], aunque nunca he dado uso de ello debido a que no lo veía necesario.

En este trabajo de investigación me he propuesto implementar mediante código uno de los algoritmos que he visto en un paper, una Red Adversaria Generativa Cuántica (GAN, en inglés) [4] que genera imágenes a partir de un circuito cuántico [5]. Como objetivo tengo verificar una sugerencia que hacen los autores del paper: implementar una función no-lineal en una parte del algoritmo que podría mejorar el rendimiento de este. Mi hipótesis al igual que los autores (aunque ellos lo comentan muy brevemente) es que el algoritmo va a reducir ligeramente el número de interacciones que son necesarias para llegar a su punto óptimo. Es decir, el modelo con la función no-lineal va a necesitar menos operaciones que lo entren conseguir los mismos resultados que el modelo sin la función.

Part I

Marc Teòric

Capítol 1

Àlgebra Lineal

Quan vaig començar a buscar informació sobre computació quàntica, en vaig ràpidament donar compte que necessitava molt més coneixement matemàtic, degut a que no entenia gairebé res dels llibres sobre computació quàntica. Arran aquell temps, una serie de vídeos sobre àlgebra lineal en va captar l'atenció, que es justament la branca de les matemàtiques sobre la qual es basa la computació quàntica. Els vídeos son les lliçons que dona el Professor Gilbert Strang al Institut Tecnològic de Massachusetts (MIT en anglès) [6, 7]. Una vegada havia vist gairebé tots els vídeos, ja tenia bastants conceptes apresos.

Aquelles lliçons es van ajudar a entendre les matemàtiques de *Quantum Computation and Quantum Information* [8] i *Quantum Computing: A Gentle Introduction*. A poc a poc, vaig anar aprenent els fonaments matemàtics de la computació quàntica i mecànica quàntica.

En aquesta secció aniré explicant els conceptes bàsics de l'àlgebra lineal, per formar els coneixements en matemàtiques necessaris per poder comprendre aquest treball.

1.1 Vectors i Espais Vectorials

Els objectes fonamentals de l'àlgebra lineal són els espais vectorials. Un espai vectorial es el conjunt de tots els vectors que tenen les mateixes dimensions. Per exemple \mathbb{R}^3 seria el espai vectorial de tots els vectors de 3 dimensions, aquests

vectors normalment s'utilitzen per representar punts en un espai tridimensional. En computació quàntica un tipus d'espais vectorials en concret són utilitzats: Els espais de Hilbert, en altres paraules, un espai vectorial amb un producte interior [9]. Els espais de Hilbert segueixen un conjunt de productes i compleixen unes certes normes, en aquest capítol presentaré una part d'aquestes normes i productes, la quantitat que és necessària. S'ha de tenir en compte que els espais de Hilbert són molt més complicats que el que es representa en aquest treball, també que d'aquí en endavant, quan mencionï espai vectorial hem referiré a un espai de Hilbert, d'ha no ser que s'especifiqui el contrari.

Els espais vectorial estan definits per les seves bases, un set de vectors $B = \{|v_1\rangle, \dots, |v_n\rangle\}$ es una base vàlida per l'espai V , si cada vector $|v\rangle$ en l'espai es pot escriure com $|v\rangle = \sum_i a_i |v_i\rangle$ per $|v_i\rangle \in B$. Els vectors en B són linealment independent entre ells.

La notació estàndard pels conceptes de àlgebra lienal en mecànica quàntica es la notació de Dirac, en la qual es representa un vector com $|\psi\rangle$. On ψ es la etiqueta del vector. Un vector $|\psi\rangle$ amb n dimensions també pot ser representat com una matriu columna que te la forma:

$$|\psi\rangle = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{n-1} \\ z_n \end{bmatrix}$$

On els nombres complexos $(z_1, z_2, \dots, z_{n-1}, z_n)$ són els seus elements. Un vector escrit com a $|\psi\rangle$ també s'anomena *ket*.

La adició d'un par de vectors en un espai de Hilbert es definida per ¹:

$$|\psi\rangle + |\varphi\rangle = \begin{bmatrix} \psi_1 \\ \vdots \\ \psi_n \end{bmatrix} + \begin{bmatrix} \varphi_1 \\ \vdots \\ \varphi_n \end{bmatrix}$$

¹Els vectors d'aquesta definició tenen els seus elements representats per la seva etiqueta i un subscrit e.g. el vector $|\psi\rangle$ te un element qualsevol ψ_1 i el seu primer element es ψ_1 . Aquesta notació es seguirà utilitzant al llarg del treball.

A més a més, hi ha una multiplicació per un escalar² definida per:

$$z|\psi\rangle = z \begin{bmatrix} \psi_1 \\ \vdots \\ \psi_n \end{bmatrix} = \begin{bmatrix} z\psi_1 \\ \vdots \\ z\psi_n \end{bmatrix}$$

On z es un escalar i $|\psi\rangle$ un vector. Cal que notar que cada element del vector es multiplica per el escalar.

Degut a que els espais de Hilbert son complexos tenen un conjugat complex definit per escalar com a: Per un escalar complex $z = a + bi$, el seu conjugat z^* es igual a $a - bi$.

Aquesta noció pot ampliar per a vectors i matrius, agafant el conjugat de totes les seves entrades/elements:

$$|\psi\rangle^* = \begin{bmatrix} \psi_1 \\ \vdots \\ \psi_n \end{bmatrix}^* = \begin{bmatrix} \psi_1^* \\ \vdots \\ \psi_n^* \end{bmatrix}$$

$$A^* = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix}^* = \begin{bmatrix} A_{11}^* & \cdots & A_{1n}^* \\ \vdots & \ddots & \vdots \\ A_{m1}^* & \cdots & A_{mn}^* \end{bmatrix}$$

Amb $|\psi\rangle$ sent un vector de dimensions n , i A sent una matriu de dimensions $m \times n$.

Un altre concepte important es la transposada, representada per el supercrit T que "rota" un vector o una matriu. Un vector columna amb una dimensió $n, 1$ es transforma amb un vector fila amb una dimensió $1, n$ ³:

$$|\psi\rangle^T = \begin{bmatrix} \psi_1 \\ \vdots \\ \psi_n \end{bmatrix}^T = [\psi_1 \quad \cdots \quad \psi_n]$$

El mateix és veritat per les matrius, una matriu $m \times n$ transposada es conver-

²Un numero qualsevol en \mathbb{R} .

³En realitat els vectors columna son matrius amb dimensió $n, 1$ però he estat ometent el 1. Quan hem refereix a les dimensions de un vector qualsevol, només diré un numero, no obstant, especificaré si és un vector columna o un vector fila.

teix en una matriu $n \times m$. Per exemple:

$$A^T = \begin{bmatrix} 2 & 3 \\ 6 & 4 \\ 2 & 5 \end{bmatrix}^T = \begin{bmatrix} 2 & 6 & 2 \\ 3 & 4 & 5 \end{bmatrix}$$

La combinació de un conjugat complex i la transposada s'anomena el conjugat Hermitià, la seva notació es una \dagger supercrita. Per un vector $|\psi\rangle$ el seu conjugat Hermitià $|\psi\rangle^\dagger$ és:

$$|\psi\rangle^\dagger = (|\psi\rangle^*)^T = [\psi_1^* \quad \dots \quad \psi_n^*] = \langle\psi|$$

El conjugat Hermitià compleix que $|\psi\rangle^\dagger = \langle\psi|$ i $\langle\psi|^\dagger = |\psi\rangle$.

El conjugat Hermitià de un vector columna $|\psi\rangle$ s'anomena *bra* o vector dual. En la notació de Dirac un vector dual s'escriu com $\langle\psi|$.

1.2 Operadors Lineals

Per poder operar amb vectors i fer operacions amb ells, s'utilitzen les matrius, que també son denominades mapes lineal o operadors lineals, que són noms que descriuen millor com funcionen aquests objectes. La definició formal de un operador lineal pot ser bastant complicada, per aquesta raó, utilitzaré termes més informals al en aquesta secció.

Bàsicament, un operador lineal transforma un vector en un altre vector, aquest vector poden o no ser de espais diferents [10]. Més formalment, per un vector $|v\rangle$ en un espai V i un vector $|w\rangle$ en un espai W , un operador lineal A entre els vectors, fa l'acció:

$$A|v\rangle = |w\rangle$$

En altres paraules, l'operador mapa un element del espai vectorial V cap a un espai vectorial W . Els operadors lineals han de complir les següents operacions:

1. Adició de Vectors:

Per els vectors $|\psi\rangle$ i $|\varphi\rangle$ en un mateix espai vectorial, i un operador lineal A :

$$A(|\psi\rangle + |\varphi\rangle) = A|\psi\rangle + A|\varphi\rangle$$

2. Multiplicació Escalar:

Per el vector $|\psi\rangle$, el escalar z i el operador lineal A :

$$A(z|\psi\rangle) = zA|\psi\rangle$$

Aquestes afirmacions tenen que ser veritat per tots els vectors i tots els escalars en els espais on els operadors actuen. Cal notar que un operador lineal no te perquè ser una matriu necessàriament, fer exemple, les derivades i les integrals son operadors lienals, això es pot provar fàcilment al veure que compleixen els criteris especificats posteriorment. No obstant, les derivades i les integrals usualment no s'apliquen a vectors, sinó a les funcions, però es possible aplicar-les a vectors ⁴.

Les matrius només son la representació matricial del operadors lienals [11].

1.1 Tipus d'Operadors Lineals

En la secció actual, exposaré els tipus bàsics d'operadors lienals que són indispensables en la teoria presentada en aquest capítol i la rest del treball.

1. Operador Zero

Qualsevol espai vectorial té un vector zero expressat en notació de Dirac com a 0 , degut a que $|0\rangle$ es un altre concepte totalment diferent en CQ i IQ⁵. El vector zero es aquell vector que per qualsevol vector $|\psi\rangle$ i qualsevol escalar z , es compleix que: $|\psi\rangle + 0 = |\psi\rangle$ i $z0 = 0$.

El operador zero també s'escriu com a 0 i es defineix com l'operador que mapa qualsevol vector al vector zero: $0|\psi\rangle = 0$.

2. Matriu inversa

Un matriu quadrada⁶ A és invertible si existeix una matriu A^{-1} de manera que $AA^{-1} = A^{-1}A$. A^{-1} es la matriu inversa de A . La manera més ràpida de saber si una matriu es invertible es veient si el seu determinant no és zero.

⁴No et preocupis, que es clar que les aplicaré a vectors :D.

⁵Computació Quàntica i Informació Quàntica.

⁶Una matriu quadrada és una matriu amb dimensions $n \times n$, on $n \in \mathbb{N}$.

3. Operador Identitat

Per a qualsevol espai vectorial V existeix un operador identitat I que es definit com $I|\psi\rangle = |\psi\rangle$, aquest operador no fa cap canvi als vectors als quals opera. Cal notar també que per qualsevol matriu A i la seva inversa és veritat que $AA^{-1} = I$

4. Operador Unitari

Un operador unitari es qualsevol operador que no altera la norma dels vectors al quals es aplicat, per tant, una matriu es unitària si $AA^\dagger = I$ Per convertir qualsevol operador en unitari, es divideix les seves entrades entre la norma del operador.

5. Operadors Hermitians

Un operador Hermitià o *self-adjoint operator* en anglès, es qualsevol operador que el seu conjugat Hermitià es ell mateix: $A = A^\dagger$

Una altre cosa a tenir en compte es que existeix un operador únic A en un espai de Hilbert, de manera que per qualsevol vectors $|\psi\rangle$ i $|\varphi\rangle$, es compleix que:

$$\langle\psi|(A|\varphi\rangle) = (A^\dagger\langle\psi|)|\varphi\rangle$$

Aquest operador es conegut com el *adjoint* o conjugat Hermitià de A .

1.3 Producte Interior i Producte Exterior**1.1 Producte Interior**

Un vector dual $\langle\psi|$ i un vector $|\varphi\rangle$ combinats formen el producte interior $\langle\psi|\varphi\rangle$, el qual efectua una operació que agafa els dos vectors com a input i produeix un nombre complex com a output:

$$\langle a|b\rangle = a_1b_1 + a_2b_2 + \dots + a_{n-1}b_{n-1} + a_nb_n = z$$

Amb $z, a_i, b_i \in \mathbb{C}$. Quan hem refereixo a un producte interior, normalment diré "el producte interior de dos vectors", quan en realitat es una operació entre un vector dual i un vector.

El equivalent d'aquest producte en un espai real de dos dimensions \mathbb{R}^2 es el producte escalar, que es expressat com a :

$$\langle a|b \rangle = ||a\rangle||_2 \cdot ||b\rangle||_2 \cos \theta \quad (1.1)$$

Amb $||\cdot||_2$ sent la norma ℓ^2 definida com a $||\psi\rangle||_2 = \sqrt{\psi_1^2 + \dots + \psi_n^2}$ amb θ sent l'angle entre els vectors $|a\rangle$ i $|b\rangle$. Com he dit l'equació (1.1) és equivalent al producte interior, no obstant, segons el que he vist, no es usada àmpliament ja que interpretar θ com un angle entre vectors de dimensions altes no té molt de sentit. En contrast, he vist aquest producte presentat en la seva interpretació geomètrica⁷ com el producte entre un vector fila i un vector columna:

$$\langle a|b \rangle = \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Ja he definit la norma I have already defined the ℓ_2 norm as the square root of the sum of the squared entries of a vector:

$$||a\rangle||_2 = \sqrt{\sum_i |a_i|^2}$$

Nonetheless, the more common definition is based upon the inner product. As you can see the inner product of a vector by itself is the sum of the squared entries:

$$\langle a|a \rangle = a_1 a_1 + \dots + a_n a_n = a_1^2 + \dots + a_n^2 = \sum_i a_i^2$$

Thus, the norm can be defined as the square root of the inner product of a vector:

$$||a\rangle||_2 = \sqrt{\langle a|a \rangle} \quad (1.2)$$

When the norm is applied to a 2 dimensional vector you can see that is the same as the length of that vector, that is because norm and length are the same concepts, however, the norm is the generalized length that can be applied to a vector of any dimension.

⁷Els detalls exactes de l'interpretació geomètrica estan fora del domini d'aquest treball, malgrat que m'agradaria molt parlar sobre el tema.

From what I understand some properties of the length of a two dimensional vector do not hold with the norm of a vector that has more than 2 dimensions. In other words, the norm behaves in similar ways like the distance from the origin (which is the length), thus they are not the exact same thing. Moreover, there are different types of norm⁸ that are used in different types of scenarios. That is why I am referring to a ℓ^2 norm, a specific type of norm that is also named Euclidean norm which is used to define the ℓ^2 distance or Euclidean distance, widely used to measure the distance of two points in a 2D space or a 3D space in high school [12].

1.1.1 Properties of the Inner Product

The basic properties of the inner product are as follows:

1. Is linear in the second argument $(z_1 \langle a| + z_2 \langle c|) |b\rangle = z_1 \langle a|b\rangle + z_2 \langle c|b\rangle$
2. Conjugate symmetry $\langle a|b\rangle = (\langle b|a\rangle)^*$
3. $\langle a|a\rangle$ is non-negative and real, except in the case of $\langle a|a\rangle = 0 \Leftrightarrow |a\rangle = 0$

1.2 Orthonormal and orthogonal vectors

From the concept of norm comes the concepts a pair of orthogonal vectors and a pair of orthonormal vectors⁹.

Looking at the equation (1.2) we can see that if the inner product of a vector is one, the norm of this vector is also one. A vector that has norm one is named a unit vector. Therefore, if the inner product of a vector is one, that vector is a unit vector.

A pair of non-zero vectors are orthogonal if their inner product is zero. For two non-zero 2 dimensional vectors, if their inner product is equal to zero, you can

⁸But not different types of length, that I know of at least.

⁹Funny note, when I encountered these two terms for the first time in QC and QI [citation], I taught they were the same thing and a week passed until I realized. It was such a difficult mess to understand everything else with these two terms confused.

see that they are perpendicular to each other by looking at equation (1.1):

For $|a\rangle$ and $|b\rangle \neq 0$:

$$\text{If } \langle a|b\rangle = 0 \text{ then: } \| |a\rangle \|_2 \cdot \| |b\rangle \|_2 \cos \theta = 0$$

Because $|a\rangle$ and $|b\rangle$ are non-zero vectors, their norms can't be zero.

Thus the remainder term $\cos \theta$ is equal to zero.

Therefore, the angle θ has to be $\frac{\pi}{2}$.

However, thinking that perpendicularity and orthogonality are the same concepts is a mistake, since, it only holds when looking at 2 dimensional vectors. As with norm and length, orthogonality is the generalized concept of perpendicularity that works for high dimensional vectors.

When we mix the concepts of unit vector and orthogonal vectors we arrive at the term orthonormality [9]. A pair of non-zero vectors are orthonormal when both are unit vectors and there are orthogonal to each other:

$$|a\rangle \text{ and } |b\rangle \text{ are orthonormal if } \begin{cases} \langle a|b\rangle = 0 \\ \langle a|a\rangle = 1 \\ \langle b|b\rangle = 1 \end{cases}$$

Orthonormal vectors are important, they are broadly used in quantum computation as well as quantum mechanics because they form the basis for the vector spaces on which the quantum states are located.

One thing to point out is that I have been talking about a pair of vectors when referring to orthonormal vectors, however, orthonormality can be extended to a set of vectors. If a set has all unit vectors and the vectors are orthogonal to each other, the set is orthonormal. The set of vectors $B = \{|\beta_1\rangle, |\beta_2\rangle, \dots, |\beta_{n-1}\rangle, |\beta_n\rangle\}$ is orthonormal if $\langle \beta_i | \beta_j \rangle = \delta_{ij} \forall i, j$ [9] where δ_{ij} is the Kronecker delta defined as :

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

maybe introduce them before if you talk about basis later on

1.3 Outer product

The outer product is a function that takes two vectors -expressed as $|a\rangle \langle b|$, with $|a\rangle$ and $|b\rangle$ being vectors- and produces a linear operator as output. Unlike the inner product, there is no analog for the outer product on the mathematics taught in high school¹⁰, and it is a bit difficult to understand it as it can take two vectors from different spaces as input. It is defined as follows:

For a vector $|v\rangle$ and $|v'\rangle$ of dimensions m and a vector $|w\rangle$ of dimension n . The output is a linear operator A of dimensions $m \times n$ in the space $M_{m \times n}$:

$$|v\rangle \langle w| = A \text{ with } A \in \text{Mat}_{m \times n}.$$

Whose action is defined by:

$$(|v\rangle \langle w|) |v'\rangle \equiv |w\rangle \langle v|v'\rangle = \langle v|v'\rangle |w\rangle \quad (1.3)$$

From equation (1.3) the usefulness and meaning of the outer product are hard to comprehend, so I will look at the way to compute it next to clarify how it works. For two vectors $|a\rangle$ and $|b\rangle$ of dimensions m and n respectively, their outer product is computed multiplying each element of $|a\rangle$ by each element of $|b\rangle$ forming a matrix of size $m \times n$:

$$|a\rangle \langle b| = \begin{bmatrix} a_1 b_1 & a_1 b_2 & \cdots & a_1 b_n \\ a_2 b_1 & a_2 b_2 & \cdots & a_2 b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_m b_1 & a_m b_2 & \cdots & a_m b_n \end{bmatrix}$$

The usefulness of the outer product will be shown in future sections.

1.4 Tensor product

The last product to mention is the tensor product, represented with the symbol \otimes . This product is used to create larger vector spaces by combining smaller vector spaces. The formal explanation of this concept is quite difficult, so I will focus

¹⁰The analog of the inner product would be the dot product.

on explaining the way to compute it by using the matrix representation of this product, named the Kronecker product.

For a $m \times n$ matrix A and a $p \times q$ matrix B the output of their Kronecker product [13] is a $pm \times qn$ matrix:

$$\begin{aligned}
 A \otimes B &= \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} \\
 &= \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & \cdots & a_{11}b_{1q} & \cdots & \cdots & a_{1n}b_{11} & a_{1n}b_{12} & \cdots & a_{1n}b_{1q} \\ a_{11}b_{21} & a_{11}b_{22} & \cdots & a_{11}b_{2q} & \cdots & \cdots & a_{1n}b_{21} & a_{1n}b_{22} & \cdots & a_{1n}b_{2q} \\ \vdots & \vdots & \ddots & \vdots & & & \vdots & \vdots & \ddots & \vdots \\ a_{11}b_{p1} & a_{11}b_{p2} & \cdots & a_{11}b_{pq} & \cdots & \cdots & a_{1n}b_{p1} & a_{1n}b_{p2} & \cdots & a_{1n}b_{pq} \\ \vdots & \vdots & & \vdots & \ddots & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \ddots & \vdots & \vdots & & \vdots \\ a_{m1}b_{11} & a_{m1}b_{12} & \cdots & a_{m1}b_{1q} & \cdots & \cdots & a_{mn}b_{11} & a_{mn}b_{12} & \cdots & a_{mn}b_{1q} \\ a_{m1}b_{21} & a_{m1}b_{22} & \cdots & a_{m1}b_{2q} & \cdots & \cdots & a_{mn}b_{21} & a_{mn}b_{22} & \cdots & a_{mn}b_{2q} \\ \vdots & \vdots & \ddots & \vdots & & & \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{p1} & a_{m1}b_{p2} & \cdots & a_{m1}b_{pq} & \cdots & \cdots & a_{mn}b_{p1} & a_{mn}b_{p2} & \cdots & a_{mn}b_{pq} \end{bmatrix}
 \end{aligned}$$

Note that $a_{ij}B$ is a scalar multiplication by a matrix, with a_{ij} being the scalar and B being the matrix.

Here is a clearer example with two 2×2 matrices, note that each entry of the first matrix is multiplied by the second matrix:

$$\begin{aligned}
 \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \otimes \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} &= \begin{bmatrix} 1 \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} & 2 \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} \\ 3 \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} & 4 \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} \end{bmatrix} \\
 &= \begin{bmatrix} 1 \times 0 & 1 \times 5 & 2 \times 0 & 2 \times 5 \\ 1 \times 6 & 1 \times 7 & 2 \times 6 & 2 \times 7 \\ 3 \times 0 & 3 \times 5 & 4 \times 0 & 4 \times 5 \\ 3 \times 6 & 3 \times 7 & 4 \times 6 & 4 \times 7 \end{bmatrix} = \begin{bmatrix} 0 & 5 & 0 & 10 \\ 6 & 7 & 12 & 14 \\ 0 & 15 & 0 & 20 \\ 18 & 21 & 24 & 28 \end{bmatrix}
 \end{aligned}$$

One important piece of notation to take into account is \otimes , used to represent the equivalent of the sum (noted with \sum), but instead of addition the Kronecker product is used. In other words \otimes denotes the Kronecker product of a finite number of terms. To clarify here's an example with the identity matrix:

With \mathbb{I} as the matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and n as a power of 2:

$$\mathbb{I}_n = \overset{\log_2 n}{\otimes} \mathbb{I}$$

Here is the case for $n = 8$:

$$\mathbb{I}_8 = \overset{\log_2 8}{\otimes} \mathbb{I} = \overset{3}{\otimes} \mathbb{I} = \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The Kronecker product also works for vectors in the same way, with a scalar-vector multiplication:

For the vectors $|\psi\rangle$ and $|\varphi\rangle$ of dimensions n and m respectively:

$$|\psi\rangle \otimes |\varphi\rangle = \begin{bmatrix} \psi_1 |\varphi\rangle \\ \psi_2 |\varphi\rangle \\ \vdots \\ \psi_m |\varphi\rangle \end{bmatrix} = \begin{bmatrix} \psi_1 \varphi_1 \\ \psi_1 \varphi_2 \\ \vdots \\ \psi_1 \varphi_m \\ \vdots \\ \vdots \\ \psi_n \varphi_1 \\ \psi_n \varphi_2 \\ \vdots \\ \psi_n \varphi_m \end{bmatrix}$$

Note that the Kronecker product can also be taken between a vector and a matrix or vice-versa, however this form isn't as common as the other two.

1.1 Properties of the Tensor Product

The basic properties of the tensor product are as follows [14, 15]:

1. Associativity:

$$A \otimes (B + C) = A \otimes B + A \otimes C$$

$$(zA) \otimes B = A \otimes (zB) = z(A \otimes B)$$

$$(A \otimes B) \otimes C = A \otimes (B \otimes C)$$

$$A \otimes 0 = 0 \otimes A = 0$$

2. Non-commutative ¹¹:

$$A \otimes B \neq B \otimes A$$

1.5 Trace

The trace of a matrix is just the sum of the elements on the main diagonal, the one that goes from top to bottom and left to right.

Here's a matrix A with its main diagonal marked:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

And its trace, denoted by $\text{Tr}[A]$ is:

$$\text{Tr}[A] = 1 + 1 + 1 = 3$$

More formally, the trace of a n -dimensional squared matrix A is:

$$\text{Tr}[A] = \sum_{i=1}^n a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}$$

¹¹One cool thing is that $A \otimes B$ and $B \otimes A$ are permutation equivalent: $\exists P, Q \Rightarrow A \otimes B = P(B \otimes A)Q$ where P and Q are permutation matrices.

The trace of a matrix as the following properties:

1. Linear operator:

Because the trace is a linear mapping, it follows that:

$\text{Tr}[A + B] = \text{Tr}[A] + \text{Tr}[B]$ and $\text{Tr}[zA] = z \text{Tr}[A]$, for all squared matrices A and B and all scalars z .

2. Trace of a Kronecker product:

$$\text{Tr}[A \otimes B] = \text{Tr}[A] \text{Tr}[B]$$

3. Transpose has the same trace:

$$\text{Tr}[A] = \text{Tr}[A^T]$$

4. Trace of a product is cyclic:

For a $m \times n$ matrix A and a $n \times m$ matrix B :

$$\text{Tr}[AB] = \text{Tr}[BA]$$

One very useful way to compute the trace of an operator is through the Gram-Schmidt procedure¹² and an outer product. Using Gram-Schmidt to represent the unit vector $|\psi\rangle$ with an orthonormal basis $|i\rangle$ which includes $|\psi\rangle$ as the first element, is true that:

$$\text{Tr}[A |\psi\rangle \langle\psi|] = \sum_i \langle i| A |\psi\rangle \langle\psi|i\rangle = \langle\psi| A |\psi\rangle$$

¹²See [A.1](#) for the definition of the Gram-Schmidt procedure.

Capítol 2

Quantum Computation

After some amount of math theory, it is time to start talking about quantum mechanics, in this chapter I am going to introduce the basic concepts of quantum computation and quantum information.

Quantum mechanics is a mathematical framework or rather a set of theories used to describe and explain the physical properties of atoms, molecules, and subatomic particles. It is the framework of all quantum physics including quantum information science. The right way of presenting quantum computation is through the formal quantum mechanics postulates because, with them, the statements made in quantum computation do not seem to come from anywhere [16]. However, to not complicate this section more than it is, I will do my best to explain the concepts and math of quantum computing just on their own, without presenting more generalized concepts from quantum mechanics, unless it is totally necessary to do so.

2.1 Quantum State and Superpositions

Per descriure com evolucionen els sistemes físics a través del temps, es necessita representar els sistemes d'alguna manera. En computació quàntica és representen per estats quàntics, els quals són algun tipus de distribucions de probabilitat per els possibles resultats d'una mesura on un sistema quàntic [17].

Imaginat que tens un boli, però que no saps de quin color és, no obstant, saps que pot ser vermell o blau. Per esbrinar de quin color és, pots provar a escriure

amb el boli per veure el color de la pinta, o en altres paraules, fer una mesura. Saps que hi ha un 50% de probabilitat de que sigui vermell i un 50% de que sigui blau. En aquesta situació hipotètica tindries el teu sistema quàntic (el boli), una manera de mesurar-lo (escriure alguna cosa) i una llista amb els possibles resultats (50% vermell, 50% blau), només et falta una manera de representar-lo tot matemàticament, el estat quàntic. Per tant, per què no intentem guardar la informació que sabem del boli en un vector?

Si posem cada probabilitat de treure un resultat en un vector tenim que:

$$\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

On la primera entrada és la possibilitat de que el boli sigui vermell i la segona entrada de que sigui blau, per fer-ho més senzill aquí està en colors:

$$\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

Cal remarcar que aquest vector està normalitzat amb la norma ℓ_1 , definida com la suma de les entrades d'un vector¹, en altres paraules la norma ℓ_1 d'aquest vector és 1.

Llavors, hem d'escollir una operació matemàtica per poder extreure la informació del vector, com que el output ha de ser un número, podem provar a utilitzar un producte interior. Però primer s'ha de representar el vector com una combinació lineal de les seves bases:

$$0.5 |0\rangle + 0.5 |1\rangle = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

On $|0\rangle$ és el vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, i, $|1\rangle$ és el vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Per tant, per trobar la probabilitat, s'agafa el producte interior del vector amb la base corresponent a la probabilitat, com a continuació:

$$\langle 0|v\rangle = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} = 0.5$$

$$\langle 1|v\rangle = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} = 0.5$$

¹Amb $|a\rangle$ sent un vector, la norma ℓ_1 , denotada per $\|\cdot\|_1$ és $\| |a\rangle \|_1 = \sum_i a_i$.

Aquest procediment és bastant senzill, com a un altre exemple, per representar un boli amb 6 colors possibles amb una possibilitat aleatòria d'escriure amb un color del 6 possibles, l'estat d'aquest boli és²:

$$|w\rangle = \begin{bmatrix} 0.25 \\ 0.3 \\ 0.1 \\ 0.1 \\ 0.1 \\ 0.05 \end{bmatrix}$$

Ara veure la probabilitat de per exemple treure el color verd, utilitzant la tercera base, la que correspon al color verd:

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.25 \\ 0.3 \\ 0.1 \\ 0.1 \\ 0.1 \\ 0.05 \end{bmatrix} = 0.1$$

Deixant els bolis a una banda, ara els podem substituir per un sistema físic quàntic, com per exemple un fotó. Els fotons tenen certes propietats que poden ser mesurades, com la seva polarització³. Al mirar als fotons com ones en que oscil·len en el camp electromagnètic, la polarització és l'orientació geomètrica de l'ona. La polarització pot ser interpretada com un angle respecte a la direcció de propagació.

posar una
figura aquí

Al definir les bases del estat de polarització com vertical i horitzontal, denotades per els vectors $|\rightarrow\rangle$ i $|\uparrow\rangle$, respectivament. Podem definir un estat de superposició entre les bases, representat com $|\nearrow\rangle$ [18]:

$$|\nearrow\rangle = \alpha |\rightarrow\rangle + \beta |\uparrow\rangle$$

On α i β són números complexos. Un estat en superposició es simplement un estat on l'angle de polarització no és 0 ni $\frac{\pi}{2}$. No ens tenim que preocupar de la descripció matemàtica exacte de la polarització dels fotons al parlar de compu-

²Le posat colors a els elements per claredat.

³Concretament, són una propietat que les ones transversals tenen, el tipus d'ona de les ones electromagnètiques, que són els fotons en realitat.

tació quàntica perquè el que importa és l'informació que porten aquests estat, no la física dels sistemes que representen. Aquesta informació s'ha d'agafar mitjançant mesures, com amb els bolis. Aquestes mesures en el cas de la polarització dels fotons seria passar-los per diversos filtres de polarització, els quals deixen passar el fotó o l'absorbeixen, tot això d'una manera probabilística, es a dir, dependent de quin sigui l'estat tenen certa possibilitat de ser absorbits o no.

Per clarificar, el filtre lo que fa es col·lapsar el fotó en els dos possibles estats de polarització, l'estat en el quan el filtre es orientat o l'estat perpendicular a aquest. Si col·lapsa en l'estat del filtre el fotó passa pel el filtre, en cas de col·lapsar en l'altre estat, es absorbit. La manera en la que els fotons col·lapsen és probabilística, si agafen un filtre que està orientat horitzontalment respecte al fotons que li arriben, un fotó orientat en horitzontalment té un 100% de poder passar, mentre que un fotó polaritzat verticalment té un 0% de probabilitat de poder passar. I si un fotó està polaritzat en un angle just entre vertical i horitzontal, es a dir a 45°, tindrà un 50% de possibilitats de passar i un 50% de no poder-hi.

Aquesta és la manera amb la qual els sistemes quàntics es comporten, a través de la probabilitat, on els estats que els representen tenen la informació sobre quines són aquestes possibilitats. No obstant, la, polarització dels fotons són un sistema físic concret, quan es parla de computació quàntica és millor treballar amb conceptes més generals per poder expressar tants algoritmes com sigui possible i poder implementar aquests algoritmes en tants ordinadors quàntics com sigui possible. Per saqueta raó, en la branca de la informació quàntica i la computació quàntica es treballa amb qubits, en comptes de diversos sistemes físics.

2.2 Qubits i operacions quàntiques

Ordinadors modern representen informació a través de cadenes de zeros i uns, anomenats bits. Tot, des de imatges a lletres o instruccions electròniques. Per exemple, la lletra t és representada per la cadena de bits 01110100, codificat a través de codi binari. Tot el que fas en un ordinador es codifica i representa en codi binari.

Degut a que estem molt acostumats a codi binari, en el camp de la computació quàntica també s'utilitza, no obstant, en comptes de bits s'utilitzen qubits.

Un qubit es l'anàleg d'un bit, en altres paraules, és la unitat mínima d'informació utilitzada pels ordinadors quàntics. En els qubits podem aplicar propietats quàntiques com la superposició o l'entrellaçament⁴. Si un bit pot estar en l'estat 0 o en l'estat 1, un qubit pot estar en una combinació d'aquests estats, en un estat enmig del 0 i el 1. És una combinació lineal dels vectors que representen aquests estats, $|0\rangle$ i $|1\rangle$:

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$$

On α i β són nombres complexos i $|\psi\rangle$ és un vector en un bidimensional espai de Hilbert⁵. Els vectors $|0\rangle$ i $|1\rangle$ són anomenats els vectors de la base computacionals, representats com:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Per tant el vector $|\psi\rangle$ és:

$$|\psi\rangle = \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

Aquest vector és un estat quàntic vàlid per representar un qubit, no obstant, hi ha un important factor a tenir en compte. El vector ha d'estar normalitzat amb la norma ℓ_2 , per tant, els nombres α i β no poden ser qualsevol nombre, necessiten ser els coeficients que formen un vector amb una norma de 1.

$$\| |\psi\rangle \| = 1$$

Llavors:

$$\begin{aligned} \| |\psi\rangle \| &= \sqrt{|\alpha|^2 + |\beta|^2} = 1 \\ \Rightarrow |\alpha|^2 + |\beta|^2 &= 1 \end{aligned}$$

Definir un qubit com una combinació lineal dels estats fundamentals no és de molta ajuda, per això, elaboraré més sobre aquesta definició: Una cadena de n -bit pot només representar una única combinació d'uns i zeros, mentres que una cadena de n -qubits representa una combinació de totes les possibles combinacions. En el cas d'un qubit, aquest és una combinació dels possibles estats $|0\rangle$ i $|1\rangle$. Considera un qubit com una barreja dels estats possibles, amb cada coeficient

⁴Ja he parlat de la primera amb la polarització dels fotons, del entrellaçament parlaré més endavant.

⁵Un espai vectorial amb un producte interior.

de la combinació lineal sent el nombre que indica quant d'un estat forma part de la barreja.

Una cosa molt interessant passa quan s'augmenta el nombre qubits, la 'quantitat d'informació' creix exponencialment. Per una cadena de n qubits la quantitat d'informació que té, en altres paraules la quantitat de números que representa és 2^n , on aquests números son els coeficients de la combinació lineal. Això es perquè quan afegeixes un qubit el nombre de combinacions possibles creix exponencialment, per tant es necessiten més coeficients per representar aquestes combinacions noves en la combinació lineal. Els qubits necessiten molta més informació per poder representar-los⁶, no com els bits que al ser només una combinació és necessita només saber quina combinació és. No passa res si això no s'entén perfectament, lo important és saber que es necessiten 2^n números complexos⁷ per representar n -qubits i que es necessiten n números binaris per representar n -bits.

Per il·lustrar tot això, 2 qubits es representen amb el vector:

$$\begin{aligned} |\psi\rangle &= \alpha_1 |00\rangle + \alpha_2 |01\rangle + \alpha_3 |10\rangle + \alpha_4 |11\rangle \\ &= \alpha_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \alpha_4 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} \end{aligned}$$

Cal remarcar que els vectors de la base computacional serien les columnes d'una matriu identitat amb dimensions $2^n \times 2^n$, on n és el nombre de qubits.

Per poder representar informació amb qubits simplement es codifica aquesta informació en els qubits, això es pot fer per exemple mitjançant codi binari: Els números 0, 1, 2 i 3 són els bits 00, 01, 10 i 11 respectivament, per tant es poden representar amb dos qubits en els estats $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$, respectivament.

2.1 Representació geomètrica d'un qubit

Una aspecte que sempre m'agrada del qubits és la seva representació geomètrica, la esfera de Bloch 2.1. Si agafem una esfera unitària⁸ que té els seus pol nord

⁶Informació que es manifesta amb els coeficients de la combinació lineal

⁷Són complexos degut a que els coeficients de la combinació lineal són números complexos.

⁸Una esfera que té radi 1 i que per tant qualsevol punt en la seva superfície correspon a un vector en \mathbb{R}^3 unitari.

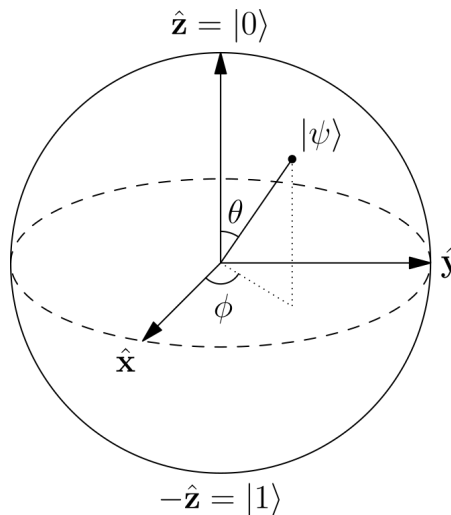


Figura 2.1: Esfera de Bloch, on es representa un estat arbitrari $|\psi\rangle$ amb els vectors \hat{x} , \hat{y} , \hat{z} representat els eixos ortogonals de la esfera.

i sur definits per els vectors $|0\rangle$ i $|1\rangle$, respectivament, tenim que cada punt de la seva superfície és un estat quàntic vàlid on les seves bases computacionals són $|0\rangle$ i $|1\rangle$.

2.2 Operacions per a només un qubit

Una vegada es té la informació representada, estaria bé poder operar amb aquella informació, aquest es justament lo que fa que els ordinadors siguin ordinadors, poder operar amb la informació. En computació quàntica els qubits al poder ser representats amb vectors que tenen els seus coeficients són operats per les anomenades portes lògiques quàntiques, que són matrius. Per exemple, si es vol passar de tindre un qubit en l'estat $|0\rangle$ al estat $|1\rangle$, s'utilitza la porta lògica X representada a continuació:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Podem veure com fa l'acció al multiplicar la matriu per el vector:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Que en notació de Dirac s'expressaria com:

$$X |0\rangle = |1\rangle$$

D'una manera més general:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 0 \times \alpha & 1 \times \beta \\ 1 \times \alpha & 0 \times \beta \end{bmatrix} = \begin{bmatrix} \beta \\ \alpha \end{bmatrix}$$

Es pot veure que aquesta matriu lo que fa és donar la volta als coeficients d'un vector, per tant:

$$X |0\rangle = |1\rangle \text{ i } X |1\rangle = |0\rangle$$

Aquesta porta lògica forma part d'un grup important, les matrius de Pauli. Hi han 3 d'aquestes la X, la Y i la Z, usualment representades per X , Y , Z o per $\sigma_x, \sigma_y, \sigma_z$. Aquestes matrius són les següents:

1. **Pauli-X:** $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

2. **Pauli-Y:** $\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$

3. **Pauli-Z:** $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

Aquestes matrius corresponen a rotacions

2.3 Operacions per a múltiples qubits

2.3 Mesurament quàntic

Tornant als fotons i la seva polarització, per predir en quina base col·lapsa, necessitem utilitzar els mesuraments quàntics. Aquests mesuraments ens donen la probabilitat que té un fotó de col·lapsar en una certa base. On el vector $|\rightarrow\rangle$ expressa la polarització d'un filtre i $|\uparrow\rangle$ és el seu complement perpendicular, un fotó polaritzat té el estat següent [18]:

$$|\nearrow\rangle = \alpha |\rightarrow\rangle + \beta |\uparrow\rangle$$

La probabilitat de ser absorbit és α^2 , cal notar que α és el coeficient que correspon a la base $|\rightarrow\rangle$, la base en la qual el filtre es orientat. En altres paraules, α^2 indica la probabilitat del col·lapse de la polarització en l'estat $|\rightarrow\rangle$.

Els coeficients α i β poden ser expressats com una funció del angle θ ⁹

$$|\nearrow\rangle = \cos \theta |\rightarrow\rangle + \sin \theta |\uparrow\rangle \quad (2.1)$$

I les probabilitats expectades es converteixen en $\cos^2 \theta$ i $\sin^2 \theta$, respectivament [18]. Ara podem mirar que els coeficients són la probabilitat, al veure el cas per $\theta = \pi/2$:

$$\begin{aligned} p(|\rightarrow\rangle) &= \cos^2 \theta = \cos^2 \frac{\pi}{2} = \left(\frac{1}{\sqrt{2}}\right)^2 \\ &= 0.5 \\ p(|\uparrow\rangle) &= \sin^2 \theta = \sin^2 \frac{\pi}{2} = \left(\frac{1}{\sqrt{2}}\right)^2 \\ &= 0.5 \end{aligned}$$

⁹Per més informació sobre la polarització dels fotons, veure l'apèndix C.

Part II

Part Experimental

Part III

Conclusions

Appendices

Apèndix A

More Linear Algebra

I have wrote many pages of linear algebra theory, but that wasn't enough. So here we go, I guess.

A.1 Gram–Schmidt Procedure

The Gram–Schmidt procedure is a method used to produce orthonormal basis for a vectors space [9]. For a d -dimensional inner product vector space V with a basis vectors set $|v_1\rangle, \dots, |v_d\rangle$, we can define a new orthonormal basis set $\{|u\rangle\}$. The first element of that set is $|u_1\rangle = |v_1\rangle / \| |v_1\rangle \|$, with the following element $|v_{k+1}\rangle$ being:

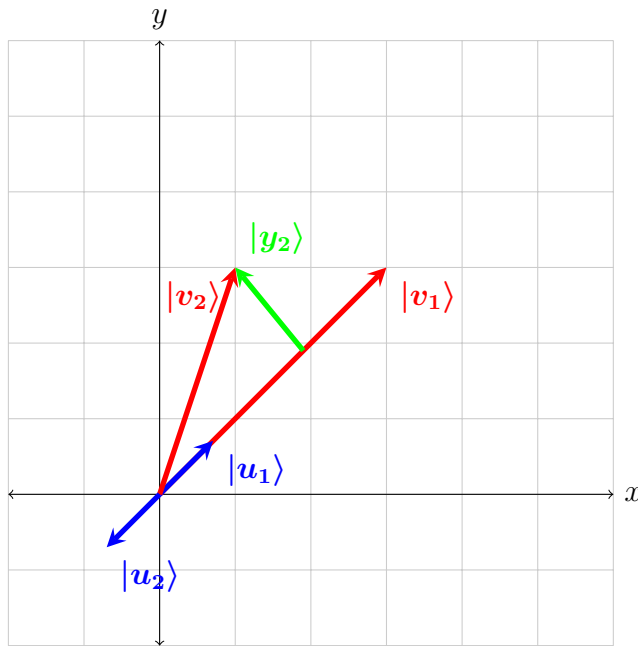
$$|u_{k+1}\rangle = \frac{|v_{k+1}\rangle - \sum_{i=1}^k \langle u_i | v_{k+1} \rangle |u_i\rangle}{\left\| |v_{k+1}\rangle - \sum_{i=1}^k \langle u_i | v_{k+1} \rangle |u_i\rangle \right\|}$$

For k in the interval $1 \leq k \leq d - 1$.

If we follow the above for each k in $1 \leq k \leq d - 1$, we obtain the new vector set $|u_1\rangle, \dots, |u_d\rangle$ that is a valid orthonormal basis for the space V . The created vector set must have the same span¹ as the previous one to be a valid basis for the space V :

$$\text{span}(\{|v\rangle\}) = \text{span}(\{|u\rangle\}) = V$$

¹The span of a set of vectors is all the possible linear combinations that come from those vectors.



Note that the span of the basis set is the definition of the space. In other words, every vector in V can be represented as a linear combination of the basis vectors.

To get at some intuition let's look at the case for $k = 1$:

$$|u_2\rangle = \frac{|v_2\rangle - \langle u_1|v_2\rangle |u_1\rangle}{\| |v_2\rangle - \langle u_1|v_2\rangle |u_1\rangle \|}$$

We can see that the vector $|u_2\rangle$ is defined by the subtraction of the vector $|v_2\rangle$ by the projection² of $|v_2\rangle$ onto $|v_1\rangle$. Which is equivalent for the projection of $|v_2\rangle$ onto $|u_1\rangle$, remember that $|u_1\rangle$ is just $|v_1\rangle$ normalized. Therefore we have:

finish the graph

The proof that is an orthonormal basis is quite simple: We can see immediately that the components of $\{|u\rangle\}$ are unit vectors because they are normalized (the vectors $|v_{k+1}\rangle - \sum_{i=1}^k \langle u_i|v_{k+1}\rangle |u_i\rangle$ are divided by their norm). And we can see that they are orthogonal by checking that the inner product of non-equal vectors in the set is 0:

For $k = 1$:

$$|u_2\rangle = \frac{|v_2\rangle - \langle u_1|v_2\rangle |u_1\rangle}{\| |v_2\rangle - \langle u_1|v_2\rangle |u_1\rangle \|}$$

²The way to project the vector $|a\rangle$ onto $|b\rangle$ is: $\langle b|a\rangle |b\rangle$.

Then the inner product with $|v_1\rangle$ is:

$$\begin{aligned}\langle u_1|u_2\rangle &= \langle u_1|\left(\frac{|v_2\rangle - \langle u_1|v_2\rangle |u_1\rangle}{\| |v_2\rangle - \langle u_1|v_2\rangle |u_1\rangle \|}\right) \\ &= \frac{\langle u_1|v_2\rangle - \langle u_1|v_2\rangle \langle u_1|u_1\rangle}{\| |v_2\rangle - \langle u_1|v_2\rangle |u_1\rangle \|} \\ &= 0\end{aligned}$$

By induction we can see that for $j \leq d$, with d being the dimension of the vector space:

$$\begin{aligned}\langle u_j|u_{n+1}\rangle &= \langle u_j|\left(\frac{|v_{n+1}\rangle - \sum_{i=1}^n \langle u_i|v_{n+1}\rangle |u_i\rangle}{\| |v_{n+1}\rangle - \sum_{i=1}^n \langle u_i|v_{n+1}\rangle |u_i\rangle \|}\right) \\ &= \frac{\langle u_j|v_{n+1}\rangle - \sum_{i=1}^n \langle u_i|v_{n+1}\rangle \langle u_j|u_i\rangle}{\| |v_{n+1}\rangle - \sum_{i=1}^n \langle u_i|v_{n+1}\rangle |u_i\rangle \|} \\ &= \frac{\langle u_j|v_{n+1}\rangle - \sum_{i=1}^n \langle u_i|v_{n+1}\rangle \delta_{ij}}{\| |v_{n+1}\rangle - \sum_{i=1}^n \langle u_i|v_{n+1}\rangle |u_i\rangle \|} \\ &= \frac{\langle u_j|v_{n+1}\rangle - \langle u_j|v_{n+1}\rangle}{\| |v_{n+1}\rangle - \sum_{i=1}^n \langle u_i|v_{n+1}\rangle |u_i\rangle \|} \\ &= 0\end{aligned}$$

All of this doesn't seem straightforward at first but remember that the inner product of two orthogonal vectors is zero, and the inner product between the same unit vector is one.

A.2 Dirac Notation Crash Course

On the following table there is a quick summary important mathematical concepts of linear algebra [19] expressed in Dirac Notation³.

³The notation used for the complex vector spaces and complex number space are not standard Dirac Notation, but I included them in the table to explain what they mean.

Notation	Description
z	Complex number
z^*	Complex conjugate of the a complex number z . $(a+bi)^* = (a-bi)$
$ \psi\rangle$	Vector with label ψ . Known as <i>ket</i>
$ \psi\rangle^T$	Transpose of vector $ \psi\rangle$
$ \psi\rangle^\dagger$	Hermitian conjugate of vector. $ \psi\rangle^\dagger = (\psi\rangle^T)^*$
$\langle\psi $	Dual vector to $ \psi\rangle$. $ \psi\rangle = \langle\psi ^\dagger$ and $\langle\psi = \psi\rangle^\dagger$. Known as <i>bra</i>
$\langle\varphi \psi\rangle$	Inner product of vectors $\langle\varphi $ and $ \psi\rangle$
$ \varphi\rangle\langle\psi $	Outer product of vectors $\langle\varphi $ and $ \psi\rangle$
$ \psi\rangle \otimes \varphi\rangle$	Tensor product of vectors $ \varphi\rangle$ and $ \psi\rangle$
0	Zero vector and zero operator
\mathbb{I}_n	Identity matrix of dimension n
\mathbb{C}_n	Complex vector space of dimension n
\mathbb{C}_1 or \mathbb{C}	Complex number space

A.3 More on the Partial Trace

Apèndix B

Quantum Computation vs Quantum Mechanics

On the introduction I mentioned that one of the reasons for with I started learning and researching quantum computing is because it is easy, on this appendix I am going to present why this is the case with a practical example.

On quantum mechanics the more general way to represent quantum state, like the orbitals of an atom of hydrogen, are wavefunctions, not statevectors. Wavefunctions are extremely useful, however, working with them adds a hole new level of complexity because they are continuous functions that depend on time. Compare them with vectors, which are discrete packets of information that evolve through discrete amounts of time.

B.1 Normalizing

Because the probabilistic interpretation of both statevectors and wavefunctions, these two objects have to be normalize upon measurement, thus making the total sum of probabilities 1. How to normalize these objects is a perfect example to illustrate the difference in complexity that I see between quantum mechanics and quantum computation.

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V\Psi$$

Figura B.1: **Schrödinger's Equation.** Where \hbar is $h/2\pi$, and V is the potential energy function.

For a wavefunction Ψ that represents a particle, the probability of finding the particle in a point x is $|\Psi(x, t)|^2$. Then, the wavefunction has to be normalized like the following:

$$\int_{-\infty}^{+\infty} |\Psi(x, t)|^2 dx = 1 \quad (\text{B.1})$$

Since the wavefunctions evolves over time with the Schrödinger equation, see Fig.B.1, any solution this equation has to be normalized. Luckily, if the wavefunction is normalized at time $t = 0$ it stays this way, in other words, the Schrödinger equations preserves the normalization of the wavefunction [20].

We can prove that the equation, preserves B.1, starting from the trivial equality:

$$\frac{d}{dt} \int_{-\infty}^{+\infty} |\Psi(x, t)|^2 dx = \frac{\partial}{\partial t} \int_{-\infty}^{+\infty} |\Psi(x, t)|^2 dx$$

By the product rule we have that¹:

$$\frac{\partial}{\partial t} |\Psi|^2 = \frac{\partial}{\partial t} (\Psi \Psi^*) = \Psi^* \frac{\partial \Psi}{\partial t} + \frac{\partial \Psi^*}{\partial t} \Psi$$

Now the Schrödinger equation says that

$$\frac{\partial \Psi}{\partial t} = \frac{i\hbar}{2m} \frac{\partial^2 \Psi}{\partial x^2} - \frac{i}{\hbar} V\Psi$$

then by taking the complex conjugate we have that

$$\frac{\partial \Psi^*}{\partial t} = -\frac{i\hbar}{2m} \frac{\partial^2 \Psi^*}{\partial x^2} + \frac{i}{\hbar} V\Psi^*$$

so

$$\frac{\partial}{\partial t} |\Psi|^2 = \frac{i\hbar}{2m} \left(\Psi^* \frac{\partial^2 \Psi}{\partial x^2} - \frac{\partial^2 \Psi^*}{\partial x^2} \Psi \right) = \frac{\partial}{\partial x} \left[\frac{i\hbar}{2m} \left(\Psi^* \frac{\partial \Psi}{\partial x} - \frac{\partial \Psi^*}{\partial x} \Psi \right) \right]$$

finally we can evaluate the integral from the start:

$$\frac{d}{dt} \int_{-\infty}^{+\infty} |\Psi(x, t)|^2 dx = \frac{i\hbar}{2m} \left(\Psi^* \frac{\partial \Psi}{\partial x} - \frac{\partial \Psi^*}{\partial x} \Psi \right) \Big|_{-\infty}^{+\infty}$$

¹From now on I am going to write $\Psi(x, t)$ as simply Ψ to not make the equations to complicated.

Because $\Psi(x, t)$ has to go to zero when x goes to either infinity, is true that:

$$\frac{d}{dt} \int_{-\infty}^{+\infty} |\Psi(x, t)|^2 dx = 0$$

Thus the integral is constant and went Ψ is normalized at $t = 0$, it stays that way for any t (positive of course).

falta la prueba para la normalización de un statevector

Apèndix C

Polarization of a photon

On equation 2.1 I have excluded the concept of phase, which determinates the type of polarization that the photon has. There are three types of polarization:

1. **Linear:** A photon is linear polarized when the phase angles α_x, α_y of both basis $|x\rangle, |y\rangle$ states are equal:

$$\begin{aligned} |\nearrow\rangle &= \cos(\theta)e^{i\alpha_x} |x\rangle + \cos(\theta)e^{i\alpha_y} |y\rangle \\ &= [\cos(\theta) |x\rangle + \sin(\theta) |y\rangle]e^{i\alpha} \end{aligned}$$

Where $\alpha = \alpha_x = \alpha_y$.

2. **Circular:** Where the angles α_x, α_y are apart exactly $\frac{\pi}{2}$ and the amplitude of both basis states is the same:

$$\begin{aligned} |\nearrow\rangle &= \frac{1}{\sqrt{2}} \cos(\theta)e^{i\alpha_x} |x\rangle \pm i \frac{1}{\sqrt{2}} \sin(\theta)e^{i\alpha_y} |y\rangle \\ &= [\cos(\theta)e^{i\alpha_x} |x\rangle \pm i \sin(\theta)e^{i\alpha_y} |y\rangle] \frac{1}{\sqrt{2}} \end{aligned}$$

Where the \pm sign indicates the difference between right and left circular polarization, + and $-$, respectively.

3. **Elliptical:** Where just the phase angles α_x, α_y differ in some amount:

$$|\nearrow\rangle = \cos(\theta)e^{i\alpha_x} |x\rangle + \sin(\theta)e^{i\alpha_y} |y\rangle$$

This is the more general case.

Bibliografia

- [1] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. The quest for a quantum neural network. *Quantum Information Processing*, 13:2567–2586, 2014. [Online; accedit en 19/09/2021: [Link](#)].
- [2] Michael Broughton, Guillaume Verdon, Trevor McCourt, Antonio J. Martinez, Jae Hyeon Yoo, Sergei V. Isakov, Philip Massey, Ramin Halavati, Murphy Yuezhen Niu, Alexander Zlokapa, Evan Peters, Owen Lockwood, Andrea Skolik, Sofiene Jerbi, Vedran Dunjko, Martin Leib, Michael Streif, David Von Dollen, Hongxiang Chen, Shuxiang Cao, Roeland Wiersema, Hsin-Yuan Huang, Jarrod R. McClean, Ryan Babbush, Sergio Boixo, Dave Bacon, Alan K. Ho, Hartmut Neven, and Masoud Mohseni. Tensorflow quantum: A software framework for quantum machine learning. 2021. [Online; accedit en 19/09/2021: [Link](#)].
- [3] IBM. Ibm quantum experience, 2021. [Online; accedit en 19/09/2021: [Link](#)].
- [4] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. 2014. [Online; accedit en 19/09/2021: [Link](#)].
- [5] He-Liang Huang, Yuxuan Du, Ming Gong, Youwei Zhao, Yulin Wu, Chaoyue Wang, Shaowei Li, Futian Liang, Jin Lin, Yu Xu, and et al. Experimental quantum generative adversarial networks for image generation. *Physical Review Applied*, 16(2), 2021. [Online; accedit en 19/09/2021: [Link](#)].
- [6] Gilbert Strang. Linear algebra mit ocw, 2011. [Online; accedit en 19/09/2021: [Link](#)].
- [7] Gilbert Strang. Matrix methods in data analysis, signal processing, and machine learning mit ocw, 2018. [Online; accedit en 19/09/2021: [Link](#)].

- [8] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 10 edition, 2010.
- [9] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*, chapter 2, page 66. Cambridge University Press, 10 edition, 2010.
- [10] Sheldon Axler. *Linear Algebra Done Right*, chapter 3, pages 37–58. Springer, 2 edition, 1997.
- [11] Sheldon Axler. *Linear Algebra Done Right*, chapter 3, pages 48–52. Springer, 2 edition, 1997.
- [12] Wolfram MathWorld. L2-norm, 2021. [Online; accedit en 19/09/2021: [Link](#)].
- [13] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*, chapter 2, page 74. Cambridge University Press, 10 edition, 2010.
- [14] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*, chapter 2, page 73. Cambridge University Press, 10 edition, 2010.
- [15] Wikipedia. Tensor product, 2021. [Online; accedit en 19/09/2021: [Link](#)].
- [16] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*, chapter 2, pages 80–96. Cambridge University Press, 10 edition, 2010.
- [17] Asher Peres. *Quantum Theory: Concepts and Methods*, chapter 2, pages 24–29. Kluwer Academic Publishers, 2002.
- [18] Eleanor Rieffel and Wolfgang Polak. *Quantum Computing: A Gentle Introduction*, chapter 2, pages 12–13. MIT Press, 1 edition, 2011.
- [19] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*, chapter 2, page 62. Cambridge University Press, 10 edition, 2010.
- [20] David J. Griffiths. *Introduction to Quantum Mechanics*, chapter 1, pages 11–12. Prentice Hall, 1995.