



Departamento de Matemáticas, Facultad de Ciencias
Universidad Autónoma de Madrid

Computación Cuántica Básica con Álgebra Lineal

TRABAJO DE FIN DE GRADO

Grado en Matemáticas

Autor: Claudia Mielgo

Tutor: Fernando Chamizo

Curso 2023-2024

Resumen

El propósito de esta memoria es introducir los fundamentos y los algoritmos básicos de la computación cuántica haciendo hincapié en su relación con el álgebra lineal. Mientras que en la mecánica cuántica habitual los operadores de evolución son operadores unitarios en espacios de Hilbert que pertenecen al ámbito del análisis funcional, la computación cuántica, al trabajar en dimensión finita, permite que los espacios y operadores relevantes pasen a ser objetos del álgebra lineal.

A pesar del álgebra subyacente en el modelo matemático, es necesario conocer la situación física que lo motiva y justifica, como se hace en el Capítulo 1 al introducir los postulados de la física cuántica básica. Los capítulos 2 y 3 se adentran en la computación cuántica propiamente dicha elaborando el diccionario con el álgebra lineal. Así, los qubits se combinan en registros que forman bases de espacios dados por productos tensoriales y los circuitos están compuestos por puertas cuánticas que son operadores unitarios. Temas de interés como el entrelazamiento y paradojas relacionadas tienen su interpretación en este ámbito. También se ilustra el potencial de la computación cuántica con el algoritmo puramente académico de Deutsch-Jozsa. El Capítulo 4 se dedica a teoremas de imposibilidad teórica y teleportación, además del algoritmo de búsqueda de Grover, más complicado y con mayor aplicabilidad. Finalmente, el Capítulo 5 estudia al algoritmo de Shor, quizá el más famoso. Es el más intrincado de los tres algoritmos presentados y combina temas del área, como la transformada cuántica de Fourier, con otros de teoría de números.

Abstract

The aim of this paper is to introduce the foundations and basic algorithms in Quantum Computing exploiting its connection with Linear Algebra. In Quantum Mechanics, propagators are unitary operators on a Hilbert space in the domain of Functional Analysis whilst in Quantum Computing, since the studied space is finite-dimensional, objects and operators belong to Linear Algebra.

In spite of the Algebra underlying the mathematical model, understanding the Physics that motivate and justify the modelling is imperative, as done in Chapter 1 by introducing the postulates of Quantum Mechanics. Chapters 3 and 4 delve into proper Quantum Computing establishing the dictionary via Linear Algebra. Thus, qubits are combined into registers that form the basis of vector spaces given by tensor products and circuits are formed of quantum gates —unitary operators—. Catchy themes such as entanglement and related paradoxes are interpreted on these terms. Furthermore, Quantum Computing's potential is illustrated on the purely academic Deutsch-Jozsa (quantum) algorithm. Chapter 4 deals with the so-called *no-go theorems* and quantum teleportation as well as Grover's algorithm for quantum search, which is more complicated and oriented to implementation. Finally, Chapter 5 focuses on Shor's algorithm, arguably the most renowned. It is probably the most intricate amongst the algorithms analysed in the paper and combines topics of the area, such as the Quantum Fourier Transform, with Number Theory.

Índice general

1	Principios de la Física Cuántica	1
1.1	El sistema cuántico es un espacio de Hilbert	3
1.2	El papel del tiempo: la probabilidad de encontrar la partícula y la evolución de los estados	4
1.3	La interpretación de Copenhague: la medición y el colapso de la función de onda	5
1.4	Sistemas cuánticos compuestos	7
2	Qubits y entrelazamientos	9
2.1	Las nociones de espín y qubit	9
2.2	Mediciones y distribución de claves cuánticas (QKD)	10
2.3	Producto tensorial, entrelazamientos y paradoja EPR	11
3	Circuitos y algoritmos cuánticos	15
3.1	Computaciones, puertas y circuitos cuánticos	15
3.2	Supremacía cuántica. El algoritmo de Deutsch-Jozsa	17
4	Algoritmo de Grover e Información cuántica	21
4.1	El algoritmo de búsqueda de Grover	21
4.2	Información cuántica	24
5	El Algoritmo de Shor	29
5.1	Motivación del problema y el algoritmo de Shor	29
5.2	Estimación de fase cuántica y conclusiones	30
A	La función de onda y la esfera de Bloch	37
A.1	Construcción de la función de onda por el principio de superposición	37
A.2	La esfera de Bloch	39
B	Detalles del algoritmo de Shor	41
B.1	Probabilidad de éxito al primer intento	41
B.2	Algoritmo de la fracción continua y aproximación de números racionales	42
	Bibliografía	47

CAPÍTULO 1

Principios de la Física Cuántica

En 1900, las contradicciones entre predicciones de la mecánica clásica y el estudio empírico de la interacción de ondas electromagnéticas con la materia llevaron a Max Planck a desarrollar un modelo heurístico que iniciaría la mecánica cuántica. Años después, Einstein aplicó esta teoría al efecto fotoeléctrico: cuando una radiación electromagnética incide sobre la materia y esta expulsa electrones, lo hace de manera proporcional a la frecuencia de dicha onda y la energía no se extiende uniformemente en todo el frente de onda (como suponía la teoría clásica), sino en *paquetes separados* o *cuantos de luz*, hoy *fotones*, cuya energía es $E = h\nu$, donde h es la *constante de Planck* y ν es la frecuencia (número de oscilaciones por unidad de tiempo). Algunos fenómenos a escala atómica se producían, por tanto, “a saltos”, comportamiento que M. Born, W. Heisenberg y P. Jordan, codificaron en matrices infinitas con información sobre los saltos. Poco después, E. Schrödinger desarrolló un modelo basado en ondas que se mostró mucho más poderoso y que ha prevalecido hasta nuestros días.

Una idea fundamental, recogiendo una hipótesis de L. de Broglie («*Toda la materia presenta características tanto ondulatorias como corpusculares, comportándose de uno u otro modo dependiendo del experimento específico*»), es que las partículas cuánticas tienen naturaleza ondulatoria y si hasta el siglo XX se trataron como partículas es porque sus ondas están muy concentradas. Combinando suposiciones y experimentos, se infirió que la onda básica correspondiente a momento lineal p y energía E es

$$(1.1) \quad \varphi(x, t) = e^{i(px - Et/\hbar)},$$

donde $\hbar = 1,05457 \cdot 10^{-34}$ es la *constante de Planck reducida* tal que $h = 2\pi\hbar$ y x y t indican posición y tiempo. Estas, sin embargo, no describen correctamente lo observado en los experimentos, lo que llevó a enunciar uno de los postulados de la mecánica cuántica. Lo que en él se afirma —aunque Born lo enunciase en 1926 con las hipótesis clásicas de un electrón con momento lineal y energía bien definidos en todo momento— es, en términos sencillos:

(P.B.): Para un sistema físico consistente en una partícula existe una función de onda asociada $\Psi(x, t)$ que codifica toda la información acerca

de este (todas las propiedades observables del sistema), que es una función compleja univariante que depende de la posición y el tiempo y cuyo módulo al cuadrado es proporcional a la densidad de probabilidad de encontrar la partícula en cierta región, para un tiempo fijado. La función de onda es además continua, con derivadas continuas (salvo en posibles puntos donde el potencial se hace infinito) y de cuadrado integrable Lebesgue (respecto a la posición).

Ahora bien, si hay una “onda” asociada a cada partícula parece que debiera existir una ecuación de ondas que rija su evolución: ecuación formulada por Schrödinger en 1926 y que recibe su nombre, resuelta con ayuda de H. Weyl. Ese mismo año, Schrödinger estableció la equivalencia entre la mecánica ondulatoria y la matricial, aunque la demostración de la equivalencia matemática de ambos formalismos llegase en 1932 con J. von Neumann. La mecánica ondulatoria ganó terreno al complicado tratamiento algebraico que requería la matricial. En efecto, el rápido éxito de la formulación de Schrödinger radicaba en que una ecuación que regula la evolución de las ondas permite matematizar bastante los problemas sin depender de intuiciones físicas. Gracias al impulso de J. von Neumann y de P. Dirac, se extremó este punto intentando dar una base muy matemática a la teoría a través de axiomas, pese a que hoy en día no hay acuerdo unánime entre los autores acerca de los postulados de la mecánica cuántica. Veremos aquí una versión abreviada (y menos ambigua que la manera en que suele formularse) de ellos, usando un ejemplo de motivación, que se tornará muy simple en el contexto de la computación cuántica porque el espacio de Hilbert será \mathbb{C}^n con el producto escalar usual. Así, podremos ver la computación cuántica a través del álgebra lineal.

En este apartado trabajaremos con un caso concreto (una partícula y una región donde se localiza) para estudiar el aspecto de la función de onda y entender los postulados, trabajando con probabilidades y espacios unitarios.

Supongamos que una partícula de masa m sobre la que no actúa ninguna fuerza está confinada al intervalo $[0, L]$. Las paredes $x = 0$ y $x = L$ son *pozos de potencial infinito* que la partícula no puede traspasar. Estudiaremos su sistema cuántico.

Como se discute en el Apéndice [(A.1)], $\Psi(x, t) = 0$ para todo tiempo t si $x \notin [0, L]$. Además, la energía y momento lineal de la partícula toman, por pares, una cantidad numerable de valores:

$$(1.2) \quad p_n = \frac{n\pi\hbar}{L} \quad \text{y} \quad E_n = \frac{(n\pi\hbar)^2}{2mL^2}, \quad \text{donde} \quad \pm p_n = \pm\sqrt{2mE_n}, \quad n \in \mathbb{N}.$$

Por otro lado, la función de onda Ψ de la partícula se escribe como superposición de ondas caracterizadas por momento lineal y energía discretos:

$$(1.3) \quad \Psi = \sum_{n \in \mathbb{N}} c_n \Psi_n, \quad \text{donde} \quad c_n \in \mathbb{C}, \quad \Psi_n(x, t) = \sqrt{\frac{2}{L}} e^{i\frac{-E_n t}{\hbar}} \text{sen} \left(\frac{\pi n x}{L} \right).$$

1.1. El sistema cuántico es un espacio de Hilbert

Consideremos el espacio unitario $L^2([0, L])$ con medida de Lebesgue, formado por clases de funciones (dos funciones f, g pertenecen a la misma clase de equivalencia si y solo si son iguales para casi todo punto respecto de la medida de Lebesgue) de cuadrado integrable, donde la norma es $\|f\| = \int_0^L |f|^2$, que deriva del producto escalar $\langle f, g \rangle = \int_0^L \bar{f}g$, y *normalizar* se reduce a dividir por la norma.¹ Los espacios L^2 sobre cualquier dominio son espacios de Hilbert: espacios vectoriales sobre el cuerpo \mathbb{C} con producto escalar y completos (una sucesión en el espacio es de Cauchy si y solo si converge en norma —inducida por el producto escalar— a un elemento del espacio). De esta forma, el sistema físico de una partícula es un espacio de Hilbert de dimensión infinita, donde las funciones Ψ_n conforman una base numerable del espacio y cada estado se corresponde con una función de onda Ψ que se puede suponer normalizada, tal como se enuncia en el primer postulado:

Postulado 1. Un sistema cuántico aislado viene representado por un espacio de Hilbert y cada estado del sistema por un vector unitario en él.²

Observación 1. Podemos suponer los estados Ψ normalizados, pues dividir entre su norma no afecta a los observables (energía y momento lineal aparecen multiplicando a x y t en el exponente) sino a los coeficientes de la exponencial compleja en cada $c_n \Psi_n$, luego no se pierde información sobre la onda. Además, consideramos que la norma es finita, como se sabe ocurre en la mayoría de estados. Cuando esto no es cierto, los físicos toman la onda en intervalos finitos arbitrariamente grandes $[-N, N]$ e interpretan la función de densidad —de que la partícula esté en la posición x en tiempo t — de manera que la probabilidad de encontrar la partícula en algún punto de $[-N, N]$ es $\|\Psi\| = 1$. Para mayor coherencia con la interpretación probabilista, imponiendo $\|\Psi\| = 1$, como veremos que $\|\Psi\|^2 = \sum |c_n|^2$, se tiene, por el Teorema de la Convergencia Dominada y porque $\Psi_n \in L^2([0, L])$, que $\Psi = \sum c_n \Psi_n \in L^2([0, L])$.

Observación 2. El espacio de Hilbert que es cada sistema cuántico admite una base ortonormal numerable. En efecto, hemos visto que para un tiempo fijo t la colección numerable de funciones $\{\Psi_n(x) = \Psi_n(x, t)\}_n$, que en nuestro espacio son los vectores $\{|\Psi_n\rangle\}_n$, genera la función de onda buscada $\Psi(x, t)$. Tomemos Ψ_n, Ψ_m con $n \neq m$ y veamos que son ortogonales, es decir, $\langle \Psi_n | \Psi_m \rangle = \int_0^L \bar{\Psi}_n \Psi_m = 0$ (notar que los extremos de la integral son 0 y L en nuestro ejemplo porque la partícula está confinada a dicho intervalo, luego los vectores Ψ_n se anulan fuera de este):

$$\langle \Psi_n | \Psi_m \rangle = \int_0^L \bar{\Psi}_n \Psi_m = \frac{2}{L} e^{\frac{i(E_n - E_m)t}{\hbar}} \int_0^L \sin\left(\frac{\pi n x}{L}\right) \sin\left(\frac{\pi m x}{L}\right) dx = 0.$$

¹El producto escalar definido tiene linealidad por la derecha y linealidad conjugada por la izquierda: $\langle \phi | c_1 \psi_1 + c_2 \psi_2 \rangle = \int \bar{\phi} (c_1 \psi_1 + c_2 \psi_2) dx = c_1 \int \bar{\phi} \psi_1 dx + c_2 \int \bar{\phi} \psi_2 dx = c_1 \langle \phi | \psi_1 \rangle + c_2 \langle \phi | \psi_2 \rangle$ y $\langle c_1 \phi_1 + c_2 \phi_2 | \psi \rangle = \int (c_1 \bar{\phi}_1 + c_2 \bar{\phi}_2) \psi dx = \bar{c}_1 \int \bar{\phi}_1 \psi dx + \bar{c}_2 \int \bar{\phi}_2 \psi dx = \bar{c}_1 \langle \phi_1 | \psi \rangle + \bar{c}_2 \langle \phi_2 | \psi \rangle$.

²Para hacer hincapié en que las funciones Ψ_n son vectores si consideramos el sistema como espacio vectorial, utilizaremos la notación física $|\Psi\rangle$ y $\langle \cdot | \cdot \rangle$ para el producto escalar $\langle \cdot, \cdot \rangle$. En particular, si M es un operador, $\langle \Phi | M | \Psi \rangle$ denota el producto escalar $\langle \Phi, M\Psi \rangle$.

Habiendo visto que los vectores $\{|\Psi_n\rangle\}_n$ generan las funciones de onda de una partícula en $[0, L]$, que están normalizados por construcción y que son ortogonales, debemos probar que son linealmente independientes para concluir que conforman una base ortonormal. Esto es directo por su ortonormalidad, pues

$$\begin{aligned} &\text{si } \sum c_k \Psi_k(x, t) = 0, \quad \text{con } c_k \in \mathbb{C}, \quad \text{entonces} \\ &\text{para cada } n \in \mathbb{N}, \quad 0 = \langle \Psi_n | \sum c_k \Psi_k(x, t) \rangle = c_n. \end{aligned}$$

1.2. El papel del tiempo: la probabilidad de encontrar la partícula y la evolución de los estados

En toda la discusión anterior, el tiempo t era fijo e identificábamos la función de onda con la densidad de probabilidad de encontrar la partícula asociada a ella en cierto espacio para ese tiempo. Claro está que dicha probabilidad puede variar con el tiempo.

Supongamos una partícula, confinada también al intervalo $[0, L]$, con función de onda $\Psi_1 + \Psi_2$. Queremos saber cuándo es más probable encontrarla en el intervalo $[0, L/4]$, si en tiempo $t = 0$ o $t = mL^2/(6\pi\hbar)$. Normalicemos primero el vector para que los cálculos den una probabilidad:

$$\int_0^L |\Psi_1 + \Psi_2|^2 dx = 1 + 1 + \frac{4}{L} \cos\left(\frac{(E_1 - E_2)t}{\hbar}\right) \int_0^L \sin\left(\frac{\pi x}{L}\right) \sin\left(\frac{2\pi x}{L}\right) dx = 2.$$

Ahora, fijado t , la probabilidad de encontrar la partícula en $[0, L/4]$ es:

$$\int_0^{L/4} \left| \frac{\Psi_1 + \Psi_2}{\|\Psi_1 + \Psi_2\|} \right|^2 dx = \frac{\pi - 1}{4\pi} + \cos\left(\frac{-3\hbar\pi^2 t}{2mL^2}\right) \frac{\sqrt{2}}{3\pi}.$$

Sustituyendo $t = 0$ y $t = mL^2/(6\pi\hbar)$, como $\cos(0) = 1 \geq \frac{\sqrt{2}}{2} = \cos(-\pi/4)$, la probabilidad de encontrar la partícula en $[0, L/4]$ es mayor $t = 0$ que en

$t = mL^2/(6\pi\hbar)$ y es $\frac{3\pi - 3 + 4\sqrt{2}}{12\pi} \approx 0,32$. Concluimos que la probabilidad de encontrar una misma partícula en una región fija sí que depende del tiempo, ¿pero hay algún invariante en el tiempo, algo que se conserve?

A modo de ejemplo, tomemos Ψ_1, Ψ_2 y el operador $P = -i\hbar \frac{d}{dx}$ para calcular $\langle \Psi_1 | P | \Psi_2 \rangle$. De nuevo, con la expresión (1.3) y la definición del producto escalar,

$$\begin{aligned} \frac{d\Psi_2}{dx} &= \sqrt{\frac{2}{L}} e^{i(-E_2 t/\hbar)} \frac{2\pi}{L} \cos\left(\frac{\pi 2x}{L}\right), \quad \text{por lo que} \\ \langle \Psi_1 | P | \Psi_2 \rangle &= \int_0^L \frac{-4i\hbar\pi}{L^2} \sin\left(\frac{\pi x}{L}\right) e^{i((E_1 - E_2)t/\hbar)} \cos\left(\frac{\pi 2x}{L}\right) dx = \frac{8\hbar}{3L} i e^{i((E_1 - E_2)t/\hbar)}. \end{aligned}$$

Fijado t , la probabilidad de detectar una partícula en algún punto es 1 y la intuición nos lleva a exigir que esto no debería cambiar con el tiempo. Si pasamos de

un estado a otro por medio de un operador, basta imponer que dicho operador sea unitario³ para que se satisfaga lo que deseamos. Podemos así enunciar el segundo postulado:

Postulado 2. La evolución de un estado en un tiempo inicial t_i a otro estado en un tiempo t viene dada por la acción de un operador unitario $U(t, t_i)$, esto es, $|\Psi(t)\rangle = U(t, t_i)|\Psi(t_i)\rangle$.

En la computación cuántica, este postulado se realiza en que los cálculos de un ordenador cuántico son multiplicaciones por matrices unitarias (matriz con entradas en \mathbb{C} cuya inversa es su traspuesta conjugada).

En nuestro ejemplo, el operador unitario $U(t, t_i)$ pasa de un estado genérico $\Psi(x, t_i) = \sum c_n \Psi_n(x, t_i)$ al estado $\tilde{\Psi}(x, t) = \sum c_n \Psi_n(x, t)$. Comprobar que el operador conserva la norma es directo con la notación empleada por los físicos cuánticos, interpretando los estados en cada tiempo t como vectores:

$$\begin{aligned} \text{Sean } |\Psi(t_i)\rangle &= \sum c_n |\Psi_n(t_i)\rangle, \quad |\tilde{\Psi}(t)\rangle = \sum c_n |\Psi_n(t)\rangle \text{ los estados en tiempos } t_i, t, \\ \text{por la ortonormalidad de la base y linealidad del producto escalar en todo tiempo} \\ \|\Psi(t_i)\|^2 &= \langle \sum c_m \Psi_m(t_i) | \sum c_n \Psi_n(t_i) \rangle = \sum c_n \sum \bar{c}_m \langle \Psi_m(t_i) | \Psi_n(t_i) \rangle = \sum c_n \bar{c}_n. \\ \text{y } \|\tilde{\Psi}(t)\|^2 &= \langle \sum c_m \Psi_m(t) | \sum c_n \Psi_n(t) \rangle = \sum c_n \sum \bar{c}_m \langle \Psi_m(t) | \Psi_n(t) \rangle = \sum c_n \bar{c}_n. \end{aligned}$$

Se concluye que la integral no depende del tiempo y la norma de un estado es $\|\Psi\|^2 = \sum |c_n|^2$ para todo tiempo t , una ley de conservación.

1.3. La interpretación de Copenhague: la medición y el colapso de la función de onda

En la física de ondas, la intensidad de la onda viene dada por el cuadrado de su amplitud, que es constante en el tiempo, motivo por el que E. Schrodinger, al introducir las funciones de onda y las ecuaciones que las regulan, interpretó $\|\Psi\|^2$ como la densidad de carga, de forma que la carga total se conserva a lo largo del tiempo. Sin embargo, la mecánica cuántica se aleja de la física de ondas en tanto que el hecho de que la intensidad $\|\Psi\|^2$ es el doble en una región que en otra no significa que la “cantidad de partícula” en esa región doble a la otra, sino que la probabilidad de que la partícula se encuentre en la primera región dobla la probabilidad de hallarla en la segunda⁴. Esto también derivó en la *dualidad onda-corpúsculo*, pues al medir estas ondas, al detectarlas en ciertos puntos, se obtenían partículas, esto es, que un

³Sea $U : H \rightarrow H$ un operador lineal U en un espacio de Hilbert H , U es unitario si $U^*U = UU^* = I$, donde I es $I : H \rightarrow H$, el operador identidad. Equivalentemente, U es unitario si conserva la norma.

⁴En una ola del mar, por ejemplo, si la amplitud es el doble en un lugar que en otro, la intensidad (amplitud al cuadrado) de la ola en el primer lugar se notará 4 veces más que en el segundo: aquí sí se tomaría esta intensidad como “cantidad de ola”.

electrón —de naturaleza ondulatoria— está o no: una detección radicalmente cuantizada.

Estudiaremos las *mediciones proyectivas*, aunque en computación cuántica se utilicen otras mediciones más abstractas. Sea Ψ una función de onda, expresada como superposición de ondas $c_n \Psi_n$ de intensidad $\|c_n \Psi_n\|^2 = |c_n|^2$ donde la energía de cada onda Ψ_n es $E_n = \frac{(n\pi\hbar)^2}{2mL^2}$ por (1.2) y (1.3). La interpretación de Copenhague supone que, de haber un instrumento que distinga las energías E_n de la función de onda Ψ , este devolvería el valor $c_n \Psi_n$ con la probabilidad dada por su intensidad $|c_n|^2$ y, tras la medición, la función Ψ colapsaría en Ψ_n . Es decir,

$$(1.4) \quad \Psi \xrightarrow{\text{medición}} c_n \Psi_n \quad \text{con probabilidad} \quad |c_n|^2 = |\langle \Psi_n | \Psi \rangle|^2.$$

Si tuviéramos un instrumento que comprobara si la energía es E_n e introducimos el estado $0,8\Psi_n + 0,2\Psi_k$ —que normalizado es $\frac{0,8}{\sqrt{0,8^2+0,2^2}}\Psi_n + \frac{0,2}{\sqrt{0,8^2+0,2^2}}\Psi_k$ —, por la relación entre la densidad de probabilidad y las funciones de onda, nos saldría que sí con probabilidad $\frac{0,64}{0,68} \approx 0,94$ y colapsaría en Ψ_n , y saldría que no un 6% de las veces aproximadamente, transformándose en Ψ_k tras la medición.

Consideremos la proyección ortogonal⁵ $P_m : L^2([0, L]) \rightarrow L^2([0, L])$, que devuelve el m -ésimo término del sumatorio $\Psi = \sum c_n \Psi_n$ para cada estado:

$$P_m(\Psi) = c_m \Psi_m.$$

Es una proyección ortogonal porque $P_m(P_m\Psi) = P_m(c_m\Psi_m) = c_m\Psi_m = P_m(\Psi)$; y sea $\tilde{\Psi} = \sum \tilde{c}_n \Psi_n$ otro estado, $\langle P_m\Psi | \tilde{\Psi} \rangle = \tilde{c}_m \langle \Psi_m | \tilde{\Psi} \rangle = \tilde{c}_m \tilde{c}_m = \tilde{c}_m \langle \Psi | \Psi_m \rangle = \langle \Psi | P_m | \tilde{\Psi} \rangle$. Notar, además, que $\sum_m P_m|\Psi\rangle = \sum c_m |\Psi_m\rangle = |\Psi\rangle$, y que $|\Psi\rangle$ colapsa con probabilidad $|c_m|^2$ en $\frac{P_m|\Psi\rangle}{c_m}$ tras la medición⁶. Esto se enuncia en uno de los postulados más controvertidos:

Postulado 3. Una medición es un conjunto de proyecciones ortogonales P_m con $\sum_m P_m = I$, donde m parametriza los resultados de la medición. La probabilidad de que al medir un estado Ψ se obtenga m es $p(m) = \langle \Psi | P_m | \Psi \rangle$ y al hacer la medición el estado se transforma inmediatamente (*colapsa*) en $\frac{P_m|\Psi\rangle}{\sqrt{p(m)}}$.

Es claro que el estado al que colapsa la función de onda está normalizado, por cómo construimos la base de nuestro espacio, pero también puede deducirse del postulado, pues, por ser P_m proyección ortogonal,

$$\left\| \frac{P_m|\Psi\rangle}{\sqrt{p(m)}} \right\|^2 = \frac{1}{p(m)} \langle P_m\Psi | P_m|\Psi \rangle = \frac{1}{p(m)} \langle \Psi | P_m^2 | \Psi \rangle = \frac{1}{p(m)} \langle \Psi | P_m | \Psi \rangle = 1.$$

Debemos interpretar que P_m proyecta sobre el espacio generado por los estados compatibles con la medición m . Así, si queremos medir con qué probabilidad la

⁵Sea V un espacio vectorial, el operador lineal $P : V \rightarrow V$ es una proyección si $P^2 = P$. Sea H un espacio de Hilbert, una proyección $P : H \rightarrow H$ es ortogonal si $\langle Px, y \rangle = \langle x, Py \rangle$ o, equivalentemente, la matriz del operador es autoadjunta: $P^2 = P = P^*$.

⁶Es legítimo dividir por c_m , pues si $c_m = 0$, entonces colapsa con probabilidad $|c_m|^2 = 0$: “casi nunca” veríamos esta posibilidad.

energía de la onda toma ciertos valores, deberemos proyectar sobre el espacio generado por los estados compatibles con energías de esos valores. Recordando que $\{|\Psi_n\rangle\}_n$ conforman la base de nuestro espacio, deberemos proyectar sobre los vectores $|\Psi_n\rangle$ cuya energía E_n sea compatible con m . Cuando el subespacio sobre el que se proyecta es de dimensión mayor que 1, entonces el estado colapsa a su “parte” contenida en dicho subespacio⁷.

Consideremos el estado (normalizado) $\Psi = \frac{4}{21}\Psi_5 + \frac{3}{21}\Psi_{20} + \frac{4}{21}\Psi_{23} + \frac{20}{21}\Psi_{50}$. Si tuviéramos un aparato que mide si la energía E de la onda es $E < E_{10}$, $E \in [E_{10}, E_{30})$ o $E \geq E_{30}$, al medir Ψ nos quedaría:

$E \in [E_{10}, E_{30})$ si y solo si $\Psi \xrightarrow{\text{medición}} \sum_{10 \leq n < 30} c_n \Psi_n = \frac{3}{21}\Psi_{20} + \frac{4}{21}\Psi_{23}$
con probabilidad $|\langle \sum_{10 \leq n < 30} c_n \Psi_n | \Psi \rangle|^2 = \sum_{10 \leq n < 30} |c_n|^2 = \frac{25}{441}$,
 $E < E_{10}$ si al medir Ψ colapsa en Ψ_5 y lo hace con probabilidad $\frac{16}{441}$,
o bien $E \geq E_{30}$ si al medir Ψ colapsa en Ψ_{50} y lo hace con probabilidad $\frac{400}{441}$.

Veamos, por último, que la interpretación probabilista de las funciones de onda es consecuencia de este tercer postulado.

Dado⁸ $\ell \in (0, L)$, consideremos los operadores $P_-, P_+ : L^2[0, L] \rightarrow L^2[0, L]$ como:

$$P_- \Psi(x) = \begin{cases} \Psi(x) & \text{si } x < \ell, \\ 0 & \text{si } x \geq \ell \end{cases} \quad \text{y} \quad P_+ \Psi(x) = \begin{cases} 0 & \text{si } x < \ell, \\ \Psi(x) & \text{si } x \geq \ell. \end{cases}$$

Es fácil comprobar que $P_- + P_+ = I$ y que $(P_\pm)^2 = P_\pm$. Además⁹, sean $\Psi, \Phi \in L^2[0, L]$, $x \in [0, L]$, si $x < \ell$ (respecto $x \geq \ell$), entonces $\langle P_- \Psi(x) | \Phi(x) \rangle = \langle \Psi(x) | \Phi(x) \rangle = \langle \Psi(x) | P_- \Phi(x) \rangle$ (respecto $\langle P_- \Psi(x) | \Phi(x) \rangle = 0 = \langle \Psi(x) | P_- \Phi(x) \rangle$). Análogamente se prueba que la proyección P_+ es ortogonal.

La primera es la medición asociada a que se detecte la partícula en $x < \ell$ y la segunda a hacerlo en $x \geq \ell$. Así, por el Postulado 3, la probabilidad de que la partícula esté en $x < \ell$ (resp. $x \geq \ell$) es la probabilidad con que Ψ colapsa a Ψ tras la medición P_- (respecto P_+), que coincide con la interpretación probabilista inicial:

$$p(x < \ell) = \langle \Psi | P_- | \Psi \rangle = \int_0^\ell |\Psi(x)|^2 dx, \quad p(x \geq \ell) = \langle \Psi | P_+ | \Psi \rangle = \int_\ell^L |\Psi(x)|^2 dx.$$

Como en los cálculos anteriores no hemos supuesto nada sobre el tiempo, se concluye que para todo t , la densidad de probabilidad de que la partícula asociada a la onda Ψ se encuentre en x en tiempo t es $|\Psi(x, t)|^2$, en concordancia con la proporcionalidad enunciada en el postulado de (P.B.): la constante es 1 por ser Ψ un estado normalizado.

1.4. Sistemas cuánticos compuestos

Para concluir este capítulo, consideremos dos partículas (distintas), una en $[0, L]$ y otra en $[2L, 3L]$. Buscamos funciones que dependan de la posición de la primera

⁷Como si en el plano, al hacer la proyección de E sobre un subespacio $S = \langle a_1, \dots, a_k \rangle$, cogemos la parte de E que es combinación lineal de $\{a_1, \dots, a_k\}$.

⁸Solo nos interesa un caso, pues si $\ell \geq L$, $\Psi(x) = 0, \forall x \geq \ell$; y si $\ell \leq 0$, $\Psi(x) = 0, \forall x \leq \ell$.

⁹De nuevo, nos interesa solo un caso porque si $x \notin [0, L]$, $\Psi(x) = 0$ para toda función de onda.

partícula y de la segunda, y del tiempo. Además, es natural pensar que, para t fijo, la probabilidad de que la primera partícula se halle en una región y la segunda partícula en otra sea el *producto* de probabilidades. Esto motiva que el espacio en el que trabajar sea el *producto tensorial* $L^2([0, L]) \otimes L^2([2L, 3L])$, generado por el producto de

funciones de onda de la forma: $\{\Psi(x_1)\tilde{\Psi}(x_2) : \Psi \in L^2([0, L]), \tilde{\Psi} \in L^2([2L, 3L])\}$,

donde se define el producto escalar: $\langle \Phi, \tilde{\Phi} \rangle = \int_0^L \int_{2L}^{3L} \overline{\Phi}(x_1, x_2)\tilde{\Phi}(x_1, x_2) dx_2 dx_1$.

Se deduce que la norma de $\Phi(x_1, x_2) = \Psi(x_1)\tilde{\Psi}(x_2)$ es $\|\Psi\|_{[0,L]}\|\tilde{\Psi}\|_{[2L,3L]}$, normalizado cuando las funciones que lo generan lo estén. Esto queda recogido en:

Postulado 4. *El espacio de Hilbert correspondiente a un sistema cuántico aislado compuesto es el producto tensorial de los espacios correspondientes a cada una de sus componentes.*

Para poner un caso concreto, supongamos que el par de partículas tiene la función de onda asociada, en $t = 0$ ¹⁰,

$$\Psi(x_1, x_2, 0) = \frac{5}{13}\Psi_1(x_1, 0)\Psi_2(x_2 + 2L, 0) + \frac{12i}{13}\Psi_2(x_1, 0)\Psi_1(x_2 + 2L, 0).$$

Como enunciamos en el Postulado 1, el estado está normalizado (aplicando Fubini):

$$\begin{aligned} \|\Psi\|^2 &= \int_0^L \int_{2L}^{3L} \left| \frac{5}{13}\Psi_1(x_1)\Psi_2(x_2 + 2L) \right|^2 + \left| \frac{12}{13}\Psi_2(x_1)\Psi_1(x_2 + 2L) \right|^2 dx_2 dx_1 = \\ &= \frac{5^2}{13^2} \int_0^L \Psi_1^2(x_1) dx_1 \int_{2L}^{3L} \Psi_2^2(x_2 + 2L) dx_2 + \frac{12^2}{13^2} \int_0^L \Psi_2^2(x_1) dx_1 \int_{2L}^{3L} \Psi_1^2(x_2 + 2L) dx_2 = 1. \end{aligned}$$

Por último, calculamos la probabilidad de que ambas partículas se encuentren en la mitad derecha del intervalo en $t = 0$. Haciendo uso de la interpretación probabilista:

$$\begin{aligned} \int_{L/2}^L \int_{5L/2}^{3L} |\Psi(x_1, x_2)|^2 dx_2 dx_1 &= \int_{L/2}^L \int_{5L/2}^{3L} \left| \frac{5}{13}\Psi_1(x_1)\Psi_2(x_2 + 2L) \right|^2 dx_2 dx_1 + \\ &+ \int_{L/2}^L \int_{5L/2}^{3L} \left| \frac{12}{13}\Psi_2(x_1)\Psi_1(x_2 + 2L) \right|^2 dx_2 dx_1 = \frac{5^2}{13^2} \cdot \frac{1}{4} + \frac{12^2}{13^2} \cdot \frac{1}{4} = \frac{1}{4}. \end{aligned}$$

Con esto, quedan enunciados los 4 postulados fundamentales de la mecánica cuántica y que se utilizarán a lo largo de toda la memoria. No obstante, quedarán formulados e interpretados en términos del álgebra lineal sobre el que formalizaremos la computación cuántica.

¹⁰ Así, por (1.3), $\bar{\Psi}_1(x) = \Psi_1(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{\pi x}{L}\right)$ y $\bar{\Psi}_2(x) = \Psi_2(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{2\pi x}{L}\right)$.

CAPÍTULO 2

Qubits y entrelazamientos

Como vimos en el anterior capítulo, los observables de momento lineal y energía están cuantizados, pero la base de la computación cuántica es que algunas magnitudes son, de hecho, finitas. Ahora dejaremos la posición x y el tiempo t fijados, y nuestro espacio L^2 se sustituirá por un espacio vectorial de dimensión finita. En este concepto se fundamenta la computación cuántica.

2.1. Las nociones de espín y qubit

Toda partícula tiene un momento angular intrínseco denominado *espín*¹ que toma dos posibles valores. Clásicamente, el momento angular se corresponde con un vector en \mathbb{R}^3 y puede tomar dos posibles sentidos para una misma dirección (“arriba” o “abajo”). Por ejemplo, tomando un eje de rotación para un electrón, se generaría un campo magnético y el momento angular, dependiendo del giro, sería en la dirección del eje hacia arriba o hacia abajo. Consideraríamos entonces una base $\mathfrak{B} = \{|\uparrow\rangle, |\downarrow\rangle\}$ para cada dirección y estudiaríamos todos los posibles momentos. Esto, extrapolado a mecánica cuántica, es lo que se descubrió tras la experimentación: todo estado Ψ es combinación lineal de los dos *autoestados* espín que, por ende, forman una base \mathfrak{B} :

$$(2.1) \quad |\Psi\rangle = a|\uparrow\rangle + b|\downarrow\rangle \quad a, b \in \mathbb{C}, \quad |\uparrow\rangle = |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathfrak{B}}, \quad |\downarrow\rangle = |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\mathfrak{B}}.$$

Atendiendo al Postulado 3, podríamos medir cuánta cantidad de cada autoestado tiene Ψ . Como los autoestados generan todos los posibles estados y las mediciones son proyecciones ortogonales, debe ser que $|0\rangle$ y $|1\rangle$ son ortogonales, luego, por el Postulado 1, el espacio de Hilbert a considerar es \mathbb{C}^2 con el producto escalar usual. Por esto y por la probabilidad de las mediciones, se observa que $|a|^2 + |b|^2 = 1$. La medición $\{P_0, P_1\}$ nos permitirá conocer el espín del estado y consistirá en proyectar sobre los subespacios generados por $|0\rangle$ y $|1\rangle$, donde:

$$p(0) = |a|^2, \quad P_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad y \quad p(1) = |b|^2, \quad P_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

¹Este no está asociado a giros ni rotaciones. Utilizamos el término recogido por la RAE de *spin*.

La expresión (2.1) es un *qubit*, un vector unitario en \mathbb{C}^2 que representa un espín. Toma valores en la circunferencia compleja de radio 1, mientras que el *bit* solo toma $a, b \in \{0, 1\}$. Además, al multiplicar un qubit aislado $|\Psi\rangle = a|0\rangle + b|1\rangle$ por un número complejo c de módulo 1, el estado no varía² y se tiene que $c|\Psi\rangle = ca|0\rangle + cb|1\rangle$, donde $|ca|^2 + |cb|^2 = 1$; es decir, que podemos suponer que $a \in \mathbb{R}_{\geq 0}$.

Cuando $S^2 = \{x \in \mathbb{R}^3 : \|x\|^2 = 1\}$ se utiliza para parametrizar qubits, se dice que es la *esfera de Bloch*, estudiada en el Apéndice [(A.2)]. Su motivación, viene de la física del espín, pues con esta identificación recuperamos el momento angular de tres coordenadas como en la mecánica clásica.

Supongamos que hacemos un experimento con el espín del electrón para ver en qué sentido apunta de la dirección $\vec{n} \in \mathbb{R}^3$, con $\|\vec{n}\| = 1$. La teoría dice que los dos posibles estados son los autovectores normalizados de la matriz $\vec{n} \cdot \vec{\sigma}$ definida como $n_1\sigma_1 + n_2\sigma_2 + n_3\sigma_3$ con σ_j las *matrices de Pauli* en la base \mathfrak{B} :

$$(2.2) \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{y} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Los autovalores son siempre $\lambda_+ = 1$ y $\lambda_- = -1$, pues $\det(\vec{n} \cdot \vec{\sigma} - \lambda I) = \lambda^2 - \|\vec{n}\|^2$. Además, notando que $\vec{n} \cdot \vec{\sigma}$ es unitaria³, los autovectores v_+ y v_- son ortogonales: $\langle v_+ | v_- \rangle = \langle \vec{n} \cdot \vec{\sigma} \cdot v_+ | \vec{n} \cdot \vec{\sigma} \cdot v_- \rangle = (v_+)^* (\vec{n} \cdot \vec{\sigma})^* (\vec{n} \cdot \vec{\sigma} \cdot v_-) = \langle v_+ | v_- \rangle$.

Sean $\varphi \in [0, 2\pi)$, $\theta \in [0, \pi]$ y $\vec{n} = (\cos \varphi \sin \theta, \sin \varphi \sin \theta, \cos \theta)$, el autoespacio de autovalor 1 de $\vec{n} \cdot \vec{\sigma}$ es $\{(x, y) \in \mathbb{C}^2 : (1 - \cos \theta)x = ye^{-i\varphi} \sin \theta\}$.

El caso $\theta = 0$ devuelve qubits en sentidos opuestos del eje Z porque $\vec{n} \cdot \vec{\sigma} = \sigma_z$ y el autovector asociado al autovalor 1 es trivialmente $|0\rangle$. En otro caso, para $\theta \neq 0$, el autoespacio es $\{(x, y) \in \mathbb{C}^2 : 2x \sin^2 \frac{\theta}{2} = ye^{-i\varphi} \sin \theta\} = \text{span}\{(\cos \frac{\theta}{2}, \sin \frac{\theta}{2} e^{-i\varphi})\}$. Nótese que en cualquier caso, el autovalor es $\cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle$, salvo multiplicar por constante. Por ejemplo, los qubits asociados a sentidos opuestos del eje X son los autovectores de σ_x , es decir, de $\vec{n} \cdot \vec{\sigma}$, donde $\theta = \frac{\pi}{2}$, $\varphi = 0$ en \vec{n} . Por lo anterior, el autovector de autovalor 1 es $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ y como el asociado a -1 ha de ser ortogonal, podemos escoger $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$.

2.2. Mediciones y distribución de claves cuánticas (QKD)

Incluso sin más desarrollo, el concepto de qubit y medición ya dan lugar a una manera segura de distribuir claves criptográficas, como estudiamos en esta sección.

Dada una base ortonormal $\mathfrak{B} = \{v_1, v_2\}$ de \mathbb{C}^2 , podemos proyectar ortogonalmente sobre el subespacio generado por cada vector y como $\{v_1\} \oplus \{v_2\} = \mathbb{C}^2$, tenemos por el postulado 3 mediciones de qubits que colapsarán con probabilidad $\langle \Psi | P_{v_j} | \Psi \rangle = \langle \Psi | \langle v_j | \Psi \rangle v_j \rangle = |\langle \Psi | v_j \rangle|^2$ en cada autoestado v_j , $j = 1, 2$.

Por ejemplo, al medir el qubit $\Psi = \frac{1}{\sqrt{3}}|0\rangle + \frac{1+i}{\sqrt{3}}|1\rangle$ usando la base canónica

²Dos estados que difieren en multiplicar por un número complejo de módulo 1 se dice que difieren en una *fase*.

³ $(\vec{n} \cdot \vec{\sigma})^* (\vec{n} \cdot \vec{\sigma}) = I$.

$\mathfrak{B} = \{|0\rangle, |1\rangle\}$, este colapsará en $|0\rangle$ con probabilidad $1/3$ y en $|1\rangle$ con $|1+i|^2/3 = 2/3$ de probabilidad. Obsérvese que dichas probabilidades dependen de la base escogida. En efecto, sea $\mathfrak{B} = \{\vec{v}_1, \vec{v}_2\}$ donde $\vec{v}_1 = \frac{2+i}{3}|0\rangle + \frac{2}{3}|1\rangle$, $\vec{v}_2 = -\frac{2i}{3}|0\rangle + \frac{1+2i}{3}|1\rangle$, la base ortonormal⁴, la probabilidad de colapsar en \vec{v}_2 es $|\langle\Psi|\vec{v}_2\rangle|^2 = \left|\frac{-2i+(1-i)(1+2i)}{3\sqrt{3}}\right|^2 = \frac{10}{27}$.

Quantum Key Distribution se basa en la idea fundamental de que para conocer la información de un qubit se deben realizar mediciones y, con ello, el qubit colapsa en un autoestado. Sean A y B dos sujetos que desean intercambiar información utilizando el siguiente código binario: se elige una clave y codificar/decodificar mensajes consiste en sumar bit a bit módulo 2, de forma que si la clave es 10001 y A quiere enviar 01101, B recibe 11100. Para que el sistema sea seguro, basta que lo sea el intercambio de claves.

Supongamos que tenemos un canal simétrico tradicional y uno cuántico unidireccional, de A a B. Disponemos también de dos sistemas de código: a partir de dos bases de \mathbb{C}^2 , se define un sistema para cada una, asignando los valores 0 y 1 a cada vector de la base. Asignar el valor 0 o 1 a un qubit consiste en elegir un sistema, medir el qubit en esa base y apuntar el valor correspondiente al vector de la base en que haya colapsado.

A genera $4N$ qubits. Para cada uno, elige un sistema y hace una medición, apuntando 0 o 1 según el resultado. Envía los $4N$ qubits, ya colapsados, a B. Para cada uno recibido, B elige una base y hace la medición, de nuevo apuntando el resultado 0 o 1. Si para un qubit A y B utilizaron el mismo sistema, el resultado que apuntaron tras la medición es el mismo con probabilidad 1. En caso contrario, la probabilidad de que apunten el mismo valor 0 o 1 será menor que 1. Ahora bien, si E está espionando y recibe los qubits, para saber qué información contienen deberá medirlos y, al colapsar y enviarlos de nuevo a B, todo lo anterior deja de ser cierto. Ahora por el canal tradicional, A y B intercambian la secuencia de sistemas utilizados, en el que coinciden $\sim 2N$ veces, y descartan los qubits de sistemas distintos. De los $\sim 2N$ qubits que se quedan, eligen N e intercambian los valores 0 o 1 que apuntaron. Si E no estuvo espionando, estos valores deberían coincidir, luego si no coinciden, buscarán otro canal porque el utilizado no es seguro. Si coinciden, saben que su canal es seguro y que los $\sim 2N$ bits que se quedaron son los mismos para A y para B. Así, pueden utilizar los N bits restantes sin necesidad de intercambiarlos y que solamente ellos conocen, como clave de cifrado/descifrado para el código que queríamos.

2.3. Producto tensorial, entrelazamientos y paradoja EPR

La computación cuántica aspira a operar con muchos qubits simultáneamente. Si cada uno se asocia con un sistema cuántico, se pretende operar con varios sistemas de partículas a la vez, cuyo modelo matemático es, por el Postulado 4, el producto tensorial. Por tanto, estudiaremos $\mathbb{C}^m \otimes \mathbb{C}^n \cong \mathbb{C}^{mn}$, definido por el isomorfismo:

$$(v_1, v_2, \dots, v_m)^t \otimes (w_1, w_2, \dots, w_n)^t = (v_1w_1, v_1w_2, \dots, v_1w_n, v_2w_1, v_2w_2, \dots, v_mw_n)^t.$$

⁴ $\langle\vec{v}_1|\vec{v}_2\rangle = \frac{-2i(2-i)}{9} + \frac{2(1+2i)}{9} = 0$ y $\|\vec{v}_j\|^2 = \frac{4+1+4}{9} = 1$.

Notar que si $\vec{v} \otimes \vec{w}, \vec{a} \otimes \vec{b} \in \mathbb{C}^m \otimes \mathbb{C}^n$, como hereda el producto escalar de \mathbb{C}^{mn} :

$$(2.3) \quad \begin{aligned} \langle \vec{v} \otimes \vec{w} | \vec{a} \otimes \vec{b} \rangle &= (\overline{v_1 w_1}, \dots, \overline{v_1 w_n}, \overline{v_2 w_1}, \dots, \overline{v_m w_n}) (a_1 b_1, \dots, a_1 b_n, a_2 b_1, \dots, a_m b_n)^t = \\ &= \sum_{i=1}^m \sum_{j=1}^n \overline{v_i w_j} a_i b_j = \sum_{i=1}^m \overline{v_i} a_i \sum_{j=1}^n \overline{w_j} b_j = \langle \vec{v} | \vec{a} \rangle_{\mathbb{C}^m} \langle \vec{w} | \vec{b} \rangle_{\mathbb{C}^n}. \end{aligned}$$

Dadas las bases $\mathfrak{B}_X, \mathfrak{B}_Y$ de los espacios vectoriales X e Y , la base del producto tensorial $X \otimes Y$ es $\mathfrak{B}_{\otimes} = \{\vec{v} \otimes \vec{w} : \vec{v} \in \mathfrak{B}_X, \vec{w} \in \mathfrak{B}_Y\}$. Empleando la notación de computación cuántica⁵, el espacio de Hilbert para dos qubits es \mathbb{C}^4 y su base es \mathcal{B}_2 :

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |01\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |10\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |11\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

El estado de un sistema de dos qubits será, por el postulado 1, en la base \mathcal{B}_2 :

$$(2.4) \quad |\Psi\rangle = a_{00}|00\rangle + a_{10}|10\rangle + a_{01}|01\rangle + a_{11}|11\rangle \quad \text{con} \quad |a_{00}|^2 + |a_{10}|^2 + |a_{01}|^2 + |a_{11}|^2 = 1.$$

Inductivamente, si nuestro sistema es el generado por n qubits, el espacio de Hilbert será \mathbb{C}^{2^n} y tendremos \mathcal{B}_n como base, que son las 2^n cadenas de longitud n de ceros y unos.⁶ Un *registro* de n qubits es un estado de dicho sistema, denotado por:

$$|\Psi\rangle = \sum_{b \in \mathcal{B}_n} a_b |b\rangle \quad \text{con} \quad \sum_{b \in \mathcal{B}_n} |a_b|^2 = 1. \quad \text{Se suele escribir } |\Psi\rangle = \sum_{b=0}^{2^n-1} a_b |b\rangle.$$

Con esto, veamos cómo queda el producto tensorial $\mathbb{C}^m \otimes \mathbb{C}^n$. Sean $|\Psi\rangle = \sum_{b=0}^{2^m-1} a_b |b\rangle$, $|\Phi\rangle = \sum_{\beta=0}^{2^n-1} \alpha_\beta |\beta\rangle$ registros de m y n qubits respectivamente. Dados $b \in \{0, \dots, 2^m-1\}$, $\beta \in \{0, \dots, 2^n-1\}$, con la notación cuántica se tiene $|b\rangle \otimes |\beta\rangle = |b\beta\rangle = |b2^n + \beta\rangle$, pues concatenar las dos cadenas de binario será poner la primera cadena desplazada n posiciones y añadir la segunda. Por último, por la propiedad distributiva respecto a la suma y la multiplicación por escalares del producto tensorial⁷, se tiene que:

$$|\Psi\rangle \otimes |\Phi\rangle = \sum_{b=0}^{2^m-1} \sum_{\beta=0}^{2^n-1} a_b |b\rangle \otimes \alpha_\beta |\beta\rangle = \sum_{b=0}^{2^m-1} \sum_{\beta=0}^{2^n-1} a_b \alpha_\beta |2^n b + \beta\rangle.$$

Sean V y W \mathbb{C} -espacios vectoriales, se define el producto tensorial $T \otimes S$ de dos endomorfismos $T : V \rightarrow V$ y $S : W \rightarrow W$ como: $(T \otimes S)(\vec{v} \otimes \vec{w}) = T(\vec{v}) \otimes S(\vec{w})$, para cada $\vec{v} \in V$ y $\vec{w} \in W$ y se extiende por linealidad.

En \mathbb{C}^4 tomemos el operador $\sigma_1 \otimes \sigma_2$. Calculando $\sigma_1|0\rangle = |1\rangle$, $\sigma_2|0\rangle = i|1\rangle$, y

⁵En computación cuántica se suele abreviar $|m\rangle \otimes |n\rangle$ como $|mn\rangle$.

⁶Cada cadena se identifica con el entero $j \in [0, 2^n)$ al que corresponde en binario (añadiendo ceros al principio cuando sea necesario) y el vector e_j de la base canónica de \mathbb{C}^{2^n} se corresponde con el $j+1$ -ésimo vector de la base \mathcal{B}_n .

⁷Dados $\lambda \in \mathbb{C}$, $\vec{v} \otimes \vec{w} \in \mathbb{C}^{mn}$, se cumple $(\lambda \vec{v}) \otimes \vec{w} = \vec{v} \otimes (\lambda \vec{w}) = \lambda(\vec{v} \otimes \vec{w})$

$\sigma_1|1\rangle = |0\rangle$, $\sigma_2|1\rangle = -i|0\rangle$, obtenemos la imagen para cada elemento de \mathcal{B}_2 , luego la matriz de $\sigma_1 \otimes \sigma_2$ en \mathcal{B}_2 y la imagen de $\frac{1}{3}|00\rangle - \frac{2i}{3}|01\rangle + \frac{2}{3}|11\rangle$ por dicho operador:

$$\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}; \quad \left(\frac{1}{3}, -\frac{2i}{3}, 0, \frac{2}{3}\right)_{\mathcal{B}_2} \mapsto \left(-\frac{2i}{3}, 0, \frac{2}{3}, \frac{i}{3}\right)_{\mathcal{B}_2} = -\frac{2i}{3}|00\rangle + \frac{2}{3}|10\rangle + \frac{i}{3}|11\rangle.$$

Después de la base canónica \mathcal{B}_2 , la base más importante de $\mathbb{C}^2 \otimes \mathbb{C}^2$ en computación cuántica es la base $\{|\Phi_j\rangle\}_{j=0}^3$ formada por los llamados *estados de Bell* donde⁸

$$\begin{aligned} |\Phi_0\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \left(\frac{1}{\sqrt{2}}, 0, 0, \frac{1}{\sqrt{2}}\right)_{\mathcal{B}_2}; & |\Phi_1\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) = \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)_{\mathcal{B}_2}; \\ |\Phi_2\rangle &= \frac{i}{\sqrt{2}}(|01\rangle - |10\rangle) = \left(0, \frac{-i}{\sqrt{2}}, \frac{i}{\sqrt{2}}, 0\right)_{\mathcal{B}_2}; & |\Phi_3\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) = \left(\frac{1}{\sqrt{2}}, 0, 0, \frac{-1}{\sqrt{2}}\right)_{\mathcal{B}_2}. \end{aligned}$$

Se dice que un registro de n qubits está *entrelazado* si no es producto tensorial de n qubits. Para $n \geq 2$, casi todos los registros de n qubits están entrelazados, aunque solo hay un criterio simple para decidirlo cuando $n = 2$. En efecto, sea $|\Psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ como en (2.4), $|\Psi\rangle$ no está entrelazado si y solo si existen $a, b, c, d \in \mathbb{C}$ tales que $|\Psi\rangle = (a|0\rangle + b|1\rangle) \otimes (c|0\rangle + d|1\rangle) = ac|00\rangle + bc|10\rangle + ad|01\rangle + bd|11\rangle$, que equivale a $a_{00}a_{11} = a_{10}a_{01}$. Luego $|\Psi\rangle$ entrelazado si y solo si $a_{00}a_{11} - a_{10}a_{01} \neq 0$.

Si la mayoría de registros de n qubits están entrelazados, podemos obtener información sobre alguno de los espines de los factores del tensor sin tener que medirlo, pues en cierta forma guardan la “relación” entre las partículas. En esto se basa la *paradoja de Einstein-Podolsky-Rosen*. La versión clásica es que si tenemos un par de guantes y nos llevamos uno pero el otro lo dejamos en casa, sabremos cuál nos hemos olvidado tan solo viendo cuál tenemos; mientras que para la versión cuántica se enunciaría algo como «*Si una partícula se desintegra en dos partículas y medimos el espín de una de ellas, conocemos automáticamente el espín de la otra*». La paradoja se manifiesta en que podemos conocer el espín de una partícula sin realizar mediciones sobre ella⁹ y, además, con independencia de la distancia entre una y otra. En \mathbb{C}^4 , supongamos que $|\Psi\rangle$ está entrelazado. Podemos realizar mediciones sobre el primer qubit y determinar con ello el segundo qubit. Midamos solo el primer qubit del estado de Bell $|\Phi_0\rangle$. Hacemos las proyecciones sobre los subespacios generados por $|0\rangle$ y $|1\rangle$ (que vimos son P_0 y P_1) dejando el segundo qubit invariante. Así, consideramos la medición $\{M_0, M_1\}$, donde $M_0 = P_0 \otimes I$ y $M_1 = P_1 \otimes I$, de forma que $|\Phi_0\rangle$ colapsa en $M_0(|\Phi_0\rangle) = |00\rangle$ con probabilidad $(1/\sqrt{2})^2$ y en $|11\rangle$ con probabilidad $1/2$. En cualquier caso, midiendo solo el primer qubit (no proyectamos sobre el segundo factor del producto tensorial) sabemos con probabilidad 1 que el segundo qubit coincide con el primero y podemos determinarlo con la medición.

Hagamos otros dos ejemplos con la base ortonormal $\{|\Psi_\alpha\rangle, |\Psi_\alpha^\perp\rangle\}$ y las proyecciones ortogonales sobre sendos subespacios que generan, denotadas por P_α y P_α^\perp :

$$|\Psi_\alpha\rangle = \cos \alpha |0\rangle + \sin \alpha |1\rangle \quad \text{y} \quad |\Psi_\alpha^\perp\rangle = -\sin \alpha |0\rangle + \cos \alpha |1\rangle.$$

⁸ I es la identidad en \mathbb{C}^2 y se define $|\Phi_j\rangle = (I \otimes \sigma_j)|\Phi_0\rangle$, para $j = 1, 2, 3$. Algunos autores no consideran el factor i en Φ_2 , pero ambas definiciones de dicho estado de Bell son equivalentes por lo comentado anteriormente, pues difieren en multiplicar por $e^{i\frac{\pi}{2}}$.

⁹Para conocer ciertas propiedades de la función de onda se realizan mediciones sobre esta y los resultados obtenidos son probabilísticos, no deterministas, como se enuncia en el Postulado 4.

Consideremos la medición $\{M_0, M_1\}$, donde $M_0 = P_\alpha \otimes I$ y $M_1 = P_\alpha^\perp \otimes I$ son operadores de medición de acuerdo con el postulado 3. En efecto, por ser $\{|\Psi_\alpha\rangle, |\Psi_\alpha^\perp\rangle\}$ base ortonormal de \mathbb{C}^2 , este es suma directa de los subespacios generados, luego, por linealidad del producto tensorial de endomorfismos, $P_\alpha + P_\alpha^\perp = I$ implica que $P_\alpha \otimes I + P_\alpha^\perp \otimes I = I \otimes I$, la identidad en \mathbb{C}^4 . Veamos que M_0 es proyección ortogonal (para M_1 es análogo): por ser P_α e I proyecciones ortogonales, se sigue de la definición de producto tensorial de endomorfismos que $(P_\alpha \otimes I)^2 = P_\alpha^* \otimes I^* = P_\alpha \otimes I$, y que $(P_\alpha \otimes I)^2 = P_\alpha^2 \otimes I^2 = P_\alpha \otimes I$.

Estudiemos la medición sobre el *par EPR*: $|\Phi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle)$.

Sea $|\Psi\rangle$ como en (2.1), tenemos $P_\alpha|\Psi\rangle = \langle\Psi_\alpha|\Psi\rangle|\Psi_\alpha\rangle = (a \cos \alpha + b \sin \alpha)|\Psi_\alpha\rangle$ y $P_\alpha^\perp|\Psi\rangle = \langle\Psi_\alpha^\perp|\Psi\rangle|\Psi_\alpha^\perp\rangle = (-a \sin \alpha + b \cos \alpha)|\Psi_\alpha^\perp\rangle$, luego, con unos cálculos rápidos:

$$P_\alpha|0\rangle = \cos \alpha|\Psi_\alpha\rangle, \quad P_\alpha|1\rangle = \sin \alpha|\Psi_\alpha\rangle, \quad P_\alpha^\perp|0\rangle = -\sin \alpha|\Psi_\alpha^\perp\rangle, \quad P_\alpha^\perp|1\rangle = \cos \alpha|\Psi_\alpha^\perp\rangle,$$

$$M_0|\Phi_0\rangle = \frac{1}{\sqrt{2}}(P_\alpha|0\rangle \otimes |0\rangle + P_\alpha|1\rangle \otimes |1\rangle) = \frac{1}{\sqrt{2}}|\Psi_\alpha\rangle \otimes (\cos \alpha|0\rangle + \sin \alpha|1\rangle) = \frac{1}{\sqrt{2}}|\Psi_\alpha\rangle \otimes |\Psi_\alpha\rangle,$$

$$M_1|\Phi_0\rangle = \frac{1}{\sqrt{2}}(P_\alpha^\perp|0\rangle \otimes |0\rangle + P_\alpha^\perp|1\rangle \otimes |1\rangle) = \frac{1}{\sqrt{2}}|\Psi_\alpha^\perp\rangle \otimes (-\sin \alpha|0\rangle + \cos \alpha|1\rangle) = \frac{1}{\sqrt{2}}|\Psi_\alpha^\perp\rangle \otimes |\Psi_\alpha^\perp\rangle.$$

De nuevo por el postulado 3, $|\Phi_0\rangle$ colapsa con probabilidad $\langle\Phi_0|M_0|\Phi_0\rangle = \frac{1}{2}$ en $\frac{1}{\sqrt{1/2}}M_0|\Phi_0\rangle = |\Psi_\alpha\rangle \otimes |\Psi_\alpha\rangle$ tras la medición, y en $|\Psi_\alpha^\perp\rangle \otimes |\Psi_\alpha^\perp\rangle$ con probabilidad $\frac{1}{2}$.

Si tomamos otra base ortonormal $\{|\Psi_\beta\rangle, |\Psi_\beta^\perp\rangle\}$ de \mathbb{C}^2 construida análogamente y las proyecciones ortogonales sobre los subespacios que generan, podríamos estudiar la medición¹⁰ $M = \{P_\alpha \otimes P_\beta, P_\alpha^\perp \otimes P_\beta, P_\alpha \otimes P_\beta^\perp, P_\alpha^\perp \otimes P_\beta^\perp\}$ sobre el estado de Bell $|\Phi_0\rangle$. Con esta medición no podemos conocer el estado en que colapsa uno de los qubits a partir del colapso/resultado del otro. Fijémonos en el primer operador:

$$P_\alpha|0\rangle = \cos \alpha|\Psi_\alpha\rangle, \quad P_\alpha|1\rangle = \sin \alpha|\Psi_\alpha\rangle, \quad P_\beta|0\rangle = \cos \beta|\Psi_\beta\rangle, \quad P_\beta|1\rangle = \sin \beta|\Psi_\beta\rangle.$$

Aplicando de nuevo que $|\Phi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle)$ y propiedades de $\langle\cdot|\cdot\rangle_{\mathbb{C}^2 \otimes \mathbb{C}^2}$:

$$(P_\alpha \otimes P_\beta)|\Phi_0\rangle = \frac{1}{\sqrt{2}}(\cos \alpha \cos \beta|\Psi_\alpha\rangle \otimes |\Psi_\beta\rangle + \sin \alpha \sin \beta|\Psi_\alpha\rangle \otimes |\Psi_\beta\rangle) = \frac{\cos(\alpha-\beta)}{\sqrt{2}}|\Psi_\alpha\rangle \otimes |\Psi_\beta\rangle,$$

$$\langle\Phi_0|P_\alpha \otimes P_\beta|\Phi_0\rangle = \frac{\cos(\alpha-\beta)}{\sqrt{2}} \frac{1}{\sqrt{2}} \langle(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle) || \Psi_\alpha\rangle \otimes |\Psi_\beta\rangle\rangle = \frac{\cos^2(\alpha-\beta)}{2}.$$

Se concluye, por el postulado 3, que $|\Phi_0\rangle$ colapsa tras la medición en $|\Psi_\alpha\rangle \otimes |\Psi_\beta\rangle$ con probabilidad $\frac{1}{2} \cos^2(\alpha - \beta)$.

El entrelazamiento supone una gran ventaja en la computación cuántica, en este caso ilustrada en la posibilidad de conocer qubits con certeza sin tener que medirlos e incluso situados a grandes distancias. Sin embargo, como hemos visto, deben tomarse bases y mediciones adecuadas para el propósito.

¹⁰Se comprueba que es una medición con un razonamiento similar al del párrafo anterior.

CAPÍTULO 3

Circuitos y algoritmos cuánticos

3.1. Computaciones, puertas y circuitos cuánticos

Se define una *computación cuántica* como el conjunto $\{\mathcal{H}, U, \{M_m\}\}$, donde $\mathcal{H} = \mathbb{C}^{2^n}$ es el espacio de Hilbert de n -registros definido en el capítulo anterior, U es una matriz unitaria que implementa un algoritmo (cuántico) y $\{M_m\}$ es un conjunto de mediciones. Un *circuito cuántico* es un modelo suyo, por lo general representado con un esquema donde se indica el orden de las transformaciones y mediciones (cajas) aplicadas a los qubits afectados (cables) de izquierda a derecha. Su expresión más simple son las *puertas (lógicas) cuánticas*, que operan sobre un número pequeño de qubits. Según el postulado 2, la evolución de un sistema cuántico viene dada por un operador unitario que conserva la norma y, según el postulado 3, el resultado de una medición es una proyección ortogonal que colapsa el estado con cierta probabilidad, ambas representadas en el circuito con transformaciones y mediciones, respectivamente.

Un resultado importante, el teorema de universalidad, asegura que todo circuito cuántico (matriz unitaria¹) puede construirse a partir de un conjunto de puertas cuánticas: las simples, que actúan sobre un solo qubit, y la conocida como *controlled NOT* (CNOT). Es más, cualquier conjunto de puertas lógicas clásicas sobre un ordenador clásico puede construirse con su contrapartida cuántica.

En lo que sigue, las matrices asociadas a circuitos en \mathbb{C}^{2^n} serán de dimensión 2^n , escritas en la base canónica para n -registros: $\mathcal{B}_n = \{|k\rangle : k \in [0, 2^n)$ en forma binaria}.

Algunos ejemplos de puertas simples son las correspondientes a las matrices (unitarias) de Pauli vistas en (2.2) y que representamos:

$$\text{---}\boxed{X}\text{---} \quad \text{---}\boxed{Y}\text{---} \quad \text{---}\boxed{Z}\text{---} \quad . \text{ Notar que } |0\rangle \text{---}\boxed{X}\text{---} |1\rangle \text{ y que } |1\rangle \text{---}\boxed{X}\text{---} |0\rangle ,$$

motivo por el que la puerta X se conoce como NOT y se representa con $\text{---}\oplus\text{---}$.

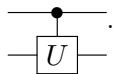
¹En lo que sigue, utilizaremos que las matrices unitarias se identifican con aplicaciones en \mathbb{C}^N que conservan la norma y que además son necesariamente lineales. Por ello, bastará conocer la imagen de cada elemento de la base para definir las.

La *puerta de Hadamard* y la *puerta* $\sqrt{\neg}$ tienen como matrices asociadas en \mathcal{B}_1 :

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad y \quad \sqrt{\neg} = \frac{1}{2} \begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix}, \quad \text{respectivamente.}$$

Es fácil comprobar que ambas son unitarias y, por tanto, conservan la norma de los estados. Además, $|0\rangle \xrightarrow{\sqrt{\neg}} \sqrt{\neg} |0\rangle = |1\rangle$ y $|1\rangle \xrightarrow{\sqrt{\neg}} \sqrt{\neg} |1\rangle = |0\rangle$, de ahí la notación: aplicar dos veces $\sqrt{\neg}$ es aplicar NOT.

Por otro lado, CNOT es un ejemplo de *puerta controlada*. En $\mathbb{C}^{2^{n+1}}$, se dice que el primer qubit *controla* la puerta cuántica asociada a la matriz U (unitaria y de dimensión 2^n) si deja el $(n+1)$ -registro invariante cuando el primer qubit es 0 y aplica a los n últimos qubits la matriz U si el primer qubit es 1. Como ya hemos visto, $|0\rangle$ y $|1\rangle$ son ortogonales en \mathbb{C}^2 , luego la proyección de uno sobre el espacio generado por el otro es el vector $\vec{0} = (0, 0)_{\mathcal{B}_1}$. Además, por la construcción del producto tensorial en [§2.3] de $\mathbb{C}^{2^{n+1}} = \mathbb{C}^2 \otimes \mathbb{C}^n$, se tiene que $\vec{0} \otimes \vec{v} = (0, \dots, 0)_{\mathcal{B}_{n+1}}$. Así, la puerta controlada asociada a U se corresponde con el operador $P_{|0\rangle} \otimes I + P_{|1\rangle} \otimes U$, donde

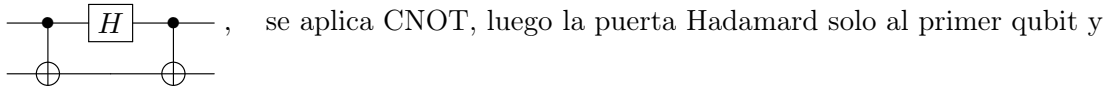
$P_{|j\rangle}$ es la proyección ortogonal sobre $|j\rangle$ ($j = 0, 1$), y su esquema es²: 

CNOT es una puerta en \mathbb{C}^4 que aplica NOT al segundo qubit si el primero es 1: deja invariantes $|00\rangle$ y $|01\rangle$ e intercambia $|11\rangle$ y $|10\rangle$. Su matriz y la transformación de $\frac{3}{5}|10\rangle + \frac{4}{5}|11\rangle$ en diagrama³:

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad y \quad \left. \begin{array}{l} \frac{3}{5}|0\rangle + \frac{4}{5}|1\rangle \\ \frac{3}{5}|10\rangle + \frac{4}{5}|11\rangle \end{array} \right\} \begin{array}{l} \text{---} \bullet \text{---} \\ \text{---} \oplus \text{---} \end{array}$$

Notemos por último que CNOT no puede ser producto tensorial de dos operadores unitarios (en \mathbb{C}^2) debido al control del primer qubit: sea $\text{CNOT} = U \otimes V$, no puede darse $\text{CNOT}|01\rangle = |01\rangle = U|0\rangle \otimes V|1\rangle$ y $\text{CNOT}|11\rangle = |10\rangle = U|1\rangle \otimes V|1\rangle$ a la vez, porque en tal caso V no estaría bien definida.

Veamos ahora un circuito que genera los estados de Bell, vistos en [§2.3]. Al interpretar un diagrama, se aplican los operadores de izquierda a derecha a los qubits correspondientes y atendiendo a los que controlan. En el circuito a tratar, que es



de nuevo CNOT pero con el primer qubit cambiado. Tomando $|00\rangle$ (que en notación binaria es $|0\rangle$), el primer paso lo deja invariante y tras el segundo obtenemos $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ en el primer qubit, luego en el último paso nos queda $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, que es el estado de Bell Φ_0 . Para $|01\rangle$ (en binario $|1\rangle$), el primer paso no afecta, después el primer qubit queda $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ y al final obtenemos $\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) = \Phi_1$. Repitiendo el proceso, aplicando el circuito al estado $|2\rangle = |10\rangle$ obtenemos

²Con tantos cables de entrada en U como número de qubits a los que afecta (n).

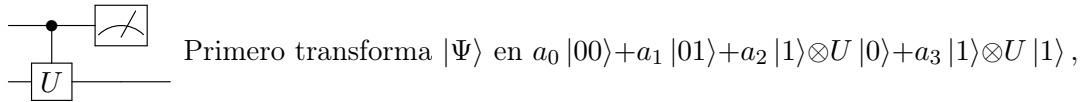
³Las matrices unitarias conservan la norma y, por tanto, son lineales, por lo que aplicamos linealidad para calcular las imágenes. En el diagrama suele separarse el qubit de control del resto.

$\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) = -i\Phi_2$ y con $|3\rangle = |11\rangle$ sale $\frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) = \Phi_3$. Así, el circuito aplica $|b\rangle$ en Φ_b , $b = 0, 1, 3$. En el caso $b = 2$, el estado viene multiplicado por una constante global.

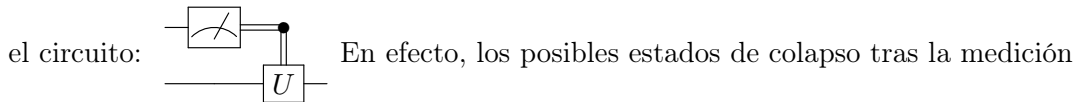
El último elemento de los esquemas más básicos de circuitos cuánticos es el de los medidores, representados con un cable de entrada, una especie de dial con aguja en una caja y ningún cable de salida o uno doble. Un cable doble indica que se trata de un bit clásico ($|0\rangle$ o $|1\rangle$, no una superposición de autoestados) y se debe a que, por lo general, se mide si un qubit es $|0\rangle$ o $|1\rangle$; es decir, atendiendo al postulado 3, los medidores hacen la medición $\{P_{|0\rangle}, P_{|1\rangle}\}$ conformada por las proyecciones ortogonales de un qubit sobre los espacios generados por $|0\rangle$ y $|1\rangle$, respectivamente. Tras la medición, el qubit colapsa en un bit clásico y, por tanto, futuras mediciones (como estas) no afectarán.

En \mathbb{C}^4 ($n = 2$), por ejemplo, de la misma forma que la medición empleada en la paradoja EPR, dado el estado $|\Psi\rangle = \sum_{b=0}^3 a_b |b\rangle$, el postulado 3 supone que al hacer la medición $\{P_{|0\rangle} \otimes I, P_{|1\rangle} \otimes I\}$, el estado $|\Psi\rangle$ colapsa en $\frac{a_0|0\rangle+a_1|1\rangle}{\sqrt{a_0^2+a_1^2}}$ con probabilidad $a_0^2 + a_1^2$ o en $\frac{a_2|2\rangle+a_3|3\rangle}{\sqrt{a_2^2+a_3^2}}$ con probabilidad $a_2^2 + a_3^2$.

Considérese el circuito donde la matriz unitaria U en \mathbb{C}^2 está controlada por el primer qubit sobre el que posteriormente se realiza una medición:



que colapsa en $\begin{cases} \frac{a_0|00\rangle+a_1|01\rangle}{\sqrt{a_0^2+a_1^2}}, prob = a_0^2 + a_1^2 \\ \frac{a_2|10\rangle\otimes U|0\rangle+a_3|11\rangle\otimes U|1\rangle}{\sqrt{a_2^2+a_3^2}}, prob = a_2^2 + a_3^2 \end{cases}$ tras la medición. Coincide con



son $\begin{cases} \frac{a_0|00\rangle+a_1|01\rangle}{\sqrt{a_0^2+a_1^2}}, prob = a_0^2 + a_1^2 \\ \frac{a_2|10\rangle+a_3|11\rangle}{\sqrt{a_2^2+a_3^2}}, prob = a_2^2 + a_3^2. \end{cases}$ Aplicando la puerta controlada asociada a U en cada posibilidad, se obtienen los mismos resultados con mismas probabilidades.

3.2. Supremacía cuántica. El algoritmo de Deutsch-Jozsa

La *supremacía cuántica*⁴ se refiere al momento en que la computación cuántica suponga una ventaja diferencial respecto a la clásica por realizar alguna tarea imposible para esta última en un tiempo razonable. En principio, ambas pueden resolver el mismo problema computacional, por lo que la única ventaja de una respecto a

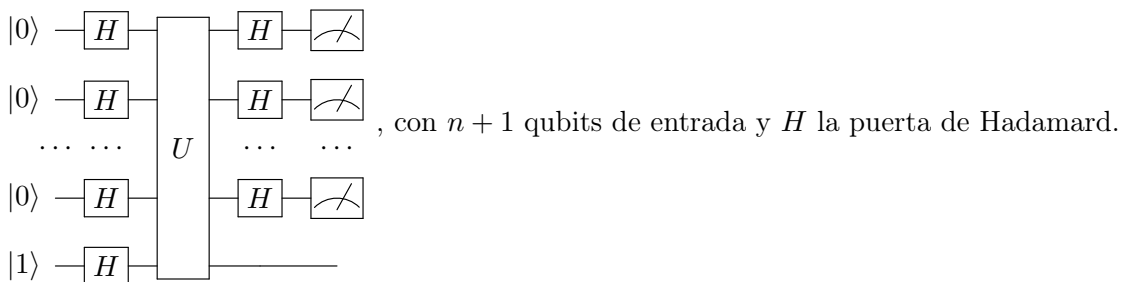
⁴Término acuñado por John Preskill en 2012 y que algunos autores prefieren llamar *primacía* o *ventaja cuántica* para evitar la comparación peyorativa que sugiere la primera.

otra sería la rapidez en la ejecución (*paralelismo cuántico*). Es un tema muy discutido en la actualidad por el que las tecnológicas compiten para hacerse con el liderazgo, pues aunque dicha supremacía solo pueda demostrarse para ciertas tareas y a nivel teórico, sin aplicaciones prácticas, el futuro de la computación cuántica es prometedor en tanto a optimización de algoritmos, estrategias financieras, criptoanálisis y simulación de sistemas cuánticos.

Por el momento, los ordenadores cuánticos pueden implementar ciertos algoritmos pero con un tamaño de datos “pequeño” porque aún no soportan suficientes qubits, y se encuentran con otros obstáculos para desarrollar todo el potencial de la computación cuántica. Por ejemplo, los estados cuánticos se mantienen durante un tiempo limitado, en el que han de realizarse todas las operaciones antes de que desaparezcan las condiciones necesarias para mantener el entrelazamiento del sistema (*decoherencia cuántica*), y se buscan sistemas de corrección de errores y nuevas herramientas para controlar los qubits con precisión, cuya manipulación en circuitos se vuelve más compleja a medida que aumentan en número.

A continuación, exponemos un ejemplo teórico, sin aplicación práctica por el momento, de cómo la computación cuántica mejora la complejidad exponencial de un algoritmo clásico a polinómica, mediante el *algoritmo de Deutsch-Jozsa*.

Dada una función f que asigna a cada cadena de n bits clásicos un valor en $\{0, 1\}$, $f : \mathcal{B}_n \rightarrow \{0, 1\}$ queremos ver si f es constante o *equilibrada*⁵, suponiendo que solo puede ser una de las dos. Podríamos considerar una lista de longitud 2^n de unos y ceros, conformada por los valores de f en cada cadena. El ordenador clásico extraería $2^{n-1} + 1$ valores (la mitad más 1) en el peor caso para concluir con certeza sobre f . En la práctica, sin embargo, se podría establecer que f es constante si y solo si 11 valores tomados al azar son iguales, y la probabilidad de error con este método es menor que 0,001. Con el algoritmo de Deutsch-Jozsa bastan 4 pasos y el circuito correspondiente es el siguiente:



El operador $H^{\otimes n} = H \otimes \dots \otimes H$ es la *transformada de (Walsh-)Hadamard* y aparece en varios algoritmos. Notar que $H |0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ y $H |1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, luego $H^{\otimes n} |0\rangle^{\otimes n} = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \dots \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Puesto que los elementos de \mathcal{B}_k son $\{|b_0\rangle, |b_1\rangle : b \in \mathcal{B}_{k-1}\}$, se tiene que $(\sum_{b \in \mathcal{B}_{k-1}} |b\rangle) \otimes (|0\rangle + |1\rangle) = \sum_{b \in \mathcal{B}_k} |b\rangle$

⁵En inglés, *balanced*, si f asigna 0 a la misma cantidad de cadenas que 1: $\#f^{-1}(\{0\}) = \#f^{-1}(\{1\})$.

y la expresión anterior queda como:

$$(3.1) \quad H^{\otimes n} |0\rangle^{\otimes n} = 2^{-n/2} \sum_{b \in \mathcal{B}_n} |b\rangle.$$

Con ello, obtenemos la primera columna del algoritmo:

$$|\psi_1\rangle = H^{\otimes n} |0\rangle^{\otimes n} \otimes H |1\rangle = 2^{-n/2} \sum_{b \in \mathcal{B}_n} |b\rangle \otimes \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) = 2^{-\frac{n+1}{2}} \sum_{b \in \mathcal{B}_{n+1}} (-1)^b |b\rangle.$$

La segunda columna del algoritmo es $|\Psi_2\rangle = U |\Psi_1\rangle$. Dados $b \in \mathcal{B}_n$, $c \in \{0, 1\}$, se define $U |bc\rangle = |b\rangle \otimes |c \oplus f(b)\rangle$, donde \oplus es la suma módulo 2. Así, si $f(b) = 0$ entonces $U |bc\rangle = |bc\rangle$, mientras que si $f(b) = 1$ entonces $U |bc\rangle = |b\rangle \otimes |1 \oplus c\rangle$. Por ello, en la base \mathcal{B}_{n+1} , U permuta, quizás (dependiendo de $f(b)$), $|b0\rangle$ y $|b1\rangle$; es decir, la matriz de U en dicha base es simétrica de ceros y unos. Además, como $U(U |bc\rangle) = |bc\rangle$, definiendo su extensión por linealidad, U es unitaria y, por ende, se asocia a una puerta cuántica⁶.

Con lo anterior, $U(|b\rangle \otimes H |1\rangle) = \frac{1}{\sqrt{2}} (U |b0\rangle - U |b1\rangle)$ es

$$\begin{cases} \frac{1}{\sqrt{2}} (|b0\rangle - |b1\rangle) = |b\rangle \otimes H |1\rangle, & \text{si } f(b) = 0 \implies (-1)^{f(b)} |b\rangle \otimes H |1\rangle, \\ \frac{1}{\sqrt{2}} (|b1\rangle - |b0\rangle) = -|b\rangle \otimes H |1\rangle, & \text{si } f(b) = 1 \implies (-1)^{f(b)} |b\rangle \otimes H |1\rangle. \end{cases}$$

Concluimos que $|\Psi_2\rangle = U(2^{-\frac{n}{2}} \sum_{b \in \mathcal{B}_n} |b\rangle \otimes H |1\rangle) = 2^{-\frac{n}{2}} \sum_{b \in \mathcal{B}_n} (-1)^{f(b)} |b\rangle \otimes H |1\rangle$.

Si pudiéramos conocer el estado $|\Psi_2\rangle$, sabríamos a partir de $n + 1$ qubits, aunque nos bastan los n primeros, el valor de f para las 2^n cadenas de \mathcal{B}_n , algo que la computación clásica no permite con n bits de entrada, pues solo codifica uno de los valores de la función, solo determina un número entre 0 y $2^n - 1$. Ahora bien, ¿cómo recuperar la información sobre f de $|\Psi_2\rangle$? Podríamos realizar la medición $\{P_{|b\rangle} \otimes I : b \in \mathcal{B}_n\}$ de forma que $|\Psi_2\rangle$ colapsaría, por el postulado 3, en $(-1)^{f(b)} |b\rangle \otimes H |1\rangle$ con probabilidad 2^{-n} . Por tanto, la ventaja cuántica sería ilusoria en tanto quipor mucho que se recuperen los valores específicos de f sobre cada cadena, el resultado es probabilista.

Sin embargo, el algoritmo aplica una transformación que acumula toda la probabilidad de ser equilibrada o constante en un solo coeficiente, obteniendo una condición necesaria y suficiente para concluir sobre f , como demostramos a continuación.

En la tercera columna, se pasa al estado $|\psi_3\rangle = (H^{\otimes n} \otimes I) |\psi_2\rangle$. Por lo anterior, $|\psi_3\rangle = 2^{-\frac{n}{2}} \sum_{b \in \mathcal{B}_n} (-1)^{f(b)} H^{\otimes n} |b\rangle \otimes H |1\rangle$. Como vimos, $H |c\rangle = \frac{1}{\sqrt{2}} (|0\rangle + (-1)^c |1\rangle)$

⁶En la literatura de la computación cuántica se dice que U es un *oráculo*, una caja negra que cumple cierta tarea, sin especificar su implementación. En particular, U se conoce como el oráculo Deutsch-Jozsa o *f(x)-controlled-NOT*, que intercambia 0 y 1 en el $(n + 1)$ -ésimo qubit si $f(x) = 1$ para el input x .

con $c \in \{0, 1\}$, por lo que dado $|b\rangle = |b_{n-1} \cdots b_0\rangle \in \mathcal{B}_n$,

$$\begin{aligned} H^{\otimes n} |b\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + (-1)^{b_{n-1}}|1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + (-1)^{b_{n-2}}|1\rangle) \otimes \cdots \otimes \frac{1}{\sqrt{2}}(|0\rangle + (-1)^{b_0}|1\rangle) \\ &= 2^{-\frac{n}{2}}(|00\rangle + (-1)^{b_{n-2}}|01\rangle + (-1)^{b_{n-1}}|10\rangle + (-1)^{b_{n-1}+b_{n-2}}|11\rangle) \otimes \cdots \otimes (|0\rangle + (-1)^{b_0}|1\rangle) \\ &= \sum_{d \in \mathcal{B}_2} 2^{-\frac{n}{2}} (-1)^{\langle b_{n-1} b_{n-2} | d \rangle} |d\rangle \otimes \cdots \otimes (|0\rangle + (-1)^{b_0}|1\rangle) \stackrel{\text{inductivamente}}{=} \\ &= \sum_{d \in \mathcal{B}_{n-1}} 2^{-\frac{n}{2}} (-1)^{\langle b_{n-1} \cdots b_1 | d \rangle} |d\rangle \otimes (|0\rangle + (-1)^{b_0}|1\rangle) = \sum_{d \in \mathcal{B}_n} 2^{-\frac{n}{2}} (-1)^{\langle b_{n-1} \cdots b_1 b_0 | d \rangle} |d\rangle. \end{aligned}$$

$$\text{Por tanto, tras el tercer paso, } |\Psi_3\rangle = 2^{-n} \sum_{b \in \mathcal{B}_n} \sum_{d \in \mathcal{B}_n} (-1)^{\langle b | d \rangle + f(b)} |d\rangle \otimes H|1\rangle.$$

El último operador del circuito es una medición sobre los primeros n qubits, medición que habíamos denotado $\{P_{|d\rangle} \otimes I : d \in \mathcal{B}_n\}$. Por el postulado 3, puesto que H preserva la norma, $|\Psi_3\rangle$ colapsa en el estado $|d\rangle \otimes H|1\rangle$ con probabilidad

$$p_d = \left| 2^{-n} \sum_{b \in \mathcal{B}_n} (-1)^{\langle b | d \rangle + f(b)} \right|^2. \quad \text{Analicemos } p_d \text{ en función de } f:$$

Si f es constante, $p_d = 2^{-2n} \left| \sum_{b \in \mathcal{B}_n} (-1)^{\langle b | d \rangle} \right|^2$. Sea $|d\rangle = |d_{n-1} \cdots d_0\rangle$. Por un lado, si $|d\rangle \neq |0\rangle$ el sumatorio es 0 y por tanto $p_d = 0$. En efecto, para el caso $n = 1$, solo puede ser $|d\rangle = |1\rangle$ y queda $(-1)^{\langle 0 | d \rangle} + (-1)^{\langle 1 | d \rangle} = 0$. Por inducción en n , suponiendo cierto para $k < n$,

$$\sum_{b \in \mathcal{B}_n} (-1)^{\langle b | d \rangle} = \sum_{b \in \mathcal{B}_{n-1}} (-1)^{\langle b | d_{n-1} \cdots d_1 \rangle + 0 \cdot d_0} + \sum_{b \in \mathcal{B}_{n-1}} (-1)^{\langle b | d_{n-1} \cdots d_1 \rangle + 1 \cdot d_0} = 0 + 0 = 0.$$

Por otro lado, si $|d\rangle = |0\rangle$, el sumatorio es 2^n y por tanto $p_d = 1$.

Así, cuando f es constante, $|\Psi_3\rangle$ colapsa en $|0\rangle \otimes H|1\rangle$ con probabilidad 1 y con probabilidad 0 en el resto de casos.

Si, por el contrario, f es equilibrada, para $|d\rangle = |0\rangle$ tenemos que:

$$p_0 = 2^{-2n} \left| \sum_{b \in \mathcal{B}_n} (-1)^{\langle b | 0 \rangle + f(b)} \right|^2 = 2^{-2n} \left| \sum_{b \in f^{-1}\{0\}} (-1)^{f(b)} + \sum_{b \in f^{-1}\{1\}} (-1)^{f(b)} \right|^2 = 0$$

porque $\#f^{-1}\{0\} = \#f^{-1}\{1\}$.

Se concluye que $p_0 = 1$ si es f constante y $p_0 = 0$ si es equilibrada y, como por hipótesis f solo puede ser constante o equilibrada, establecemos que f es constante si y solo si el resultado tras la medición es $|0\rangle^{\otimes n}$ (fijándonos solo en los primeros n qubits). De este modo, clasificamos la función con certeza absoluta y realizando solo 4 pasos.

El algoritmo de Deutsch-Jozsa mejora significativamente la tarea propuesta, aunque es una demostración meramente teórica del paralelismo cuántico. En efecto, la ventaja la proporciona esencialmente el oráculo U , que, además de que podría ser difícil de construir, depende de los valores de f y, por ende, debería conocerse “bien” f para desarrollar el algoritmo, pero en tal caso sería innecesario plantearse el problema.

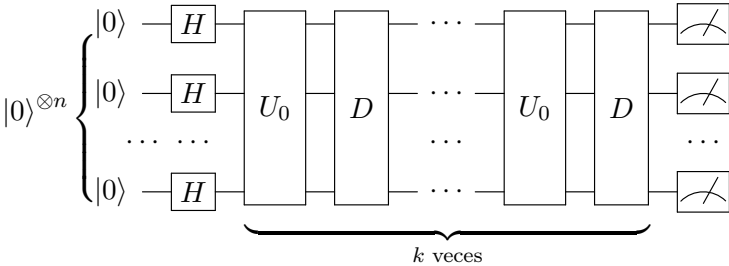
Algoritmo de Grover e Información cuántica

4.1. El algoritmo de búsqueda de Grover

Como vimos con el algoritmo de Deutsch-Jozsa, podemos aprovechar la superposición de estados de los registros de qubits para mejorar computacionalmente algunas soluciones clásicas. El algoritmo de Grover es otra muestra del paralelismo cuántico donde, a pesar de ofrecer un resultado probabilista, el error cometido es despreciable y realiza menos llamadas a la base de datos que las que haría un ordenador clásico.

Supongamos que tenemos $N = 2^n$ datos –de no ser una potencia de dos podríamos rellenar la base de datos con unos artificiales– ordenados de 0 a $2^n - 1$ y queremos estudiar la posición que ocupa el único dato que cumple cierta propiedad. Identificándolos con \mathcal{B}_n , se trata de encontrar $|b_0\rangle = f^{-1}\{1\}$ donde se define $f : \mathcal{B}_n \mapsto \{0, 1\}$ tal que $f|b\rangle = 0$ si $|b\rangle \neq |b_0\rangle$, y $f|b_0\rangle = 1$. Si quisiéramos encontrar un dato escogido al azar en un ordenador convencional, empezariamos examinando en orden la imagen por f de cada uno y parariamos la búsqueda al encontrar el primer (y único) $f^{-1}\{1\}$, un modelo de distribución binomial de parámetros $1/2, 2^n$, luego haríamos una media de 2^{n-1} llamadas a la base de datos. Veremos que con el algoritmo de Grover bastan unas $2^{n/2}$ para obtener el output correcto con una probabilidad cercana a 1, con lo que “ganamos” una raíz cuadrada.

El circuito que corresponde al algoritmo es:



Denotamos $|\phi_0\rangle = |b_0\rangle$. Por lo visto de la transformada de Hadamard en (3.1), el primer paso queda como $|\phi_1\rangle = H^{\otimes n}|0\rangle^{\otimes n} = 2^{-n/2} \sum_{b \in \mathcal{B}_n} |b\rangle$. El algoritmo de Grover devuelve como output la medición de $(DU_0)^k |\phi_1\rangle$, donde se elige el k que maximice la probabilidad de colapso en $|\phi_0\rangle$, como querriamos.

Definamos los operadores U_0 y D . Ambos tienen asociadas *matrices de Householder* extendidas a \mathbb{C} -espacios vectoriales: dado $\vec{v} \in \mathbb{C}^N$ unitario, se define la matriz de Householder asociada a él como¹ $H_{\vec{v}} = I - 2\vec{v}\vec{v}^*$ y es unitaria porque $\vec{v}^*\vec{v} = 1$ y $H_{\vec{v}}^* = I - 2(\vec{v}\vec{v}^*)^* = H_{\vec{v}}$ implican $H_{\vec{v}}H_{\vec{v}}^* = H_{\vec{v}}^2 = I + 4\vec{v}\vec{v}^*\vec{v}\vec{v}^* - 4\vec{v}\vec{v}^* = I$. Además, para todo $\vec{u} \in \mathbb{C}^N$, $H_{\vec{v}}\vec{u} = \vec{u} - 2\vec{v}\vec{v}^*\vec{u} = I - 2\langle\vec{v}|\vec{u}\rangle\vec{v}$, representado una simetría respecto al plano perpendicular a \vec{v} en tanto que $H_{\vec{v}}\vec{v} = -\vec{v}$ y $H_{\vec{v}}\vec{w} = \vec{w}$ si $\vec{v} \perp \vec{w}$. Con esta notación, puesto que los registros son unitarios,

$$U_0 = H_{|\phi_0\rangle} \quad y \quad D = -H_{|\phi_1\rangle} \quad \text{quedan bien definidas como puertas cuánticas.}$$

Nótese que U_0 es un oráculo para consultar la base de datos y depende de $|\phi_0\rangle$ (y de f). De hecho, veamos que se relaciona con el oráculo de Deutsch-Jozsa U definido para f mediante:

$$\left. \begin{array}{c} |b\rangle \text{---} \boxed{U} \text{---} \\ |1\rangle \text{---} \boxed{H} \text{---} \boxed{U} \text{---} \boxed{H} \text{---} \end{array} \right\} (U_0 |b\rangle) \otimes |1\rangle$$

Por la ortogonalidad de \mathcal{B}_n y la simetría que representan las matrices de Householder, $U_0|\phi_0\rangle = -|\phi_0\rangle$ y $U_0|b\rangle = |b\rangle$ si $b \in \mathcal{B}_n - \{b_0\}$. Por otro lado, como $f(b) = 1$ si y solo si $|b\rangle = |\phi_0\rangle$, dado $c \in \{0, 1\}$, $U|\phi_0c\rangle = |\phi_0\rangle \otimes |1 \oplus c\rangle$ y $U|bc\rangle = |bc\rangle$ si $b \in \mathcal{B}_n - \{b_0\}$.

Recordando que $H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, la acción del circuito sobre $|b\rangle \otimes |1\rangle$ es:

$$\begin{aligned} \text{Si } |b\rangle = |\phi_0\rangle, \quad & \frac{1}{\sqrt{2}}(|\phi_00\rangle - |\phi_01\rangle) \rightarrow \frac{1}{\sqrt{2}}(U|\phi_00\rangle - U|\phi_01\rangle) = \frac{1}{\sqrt{2}}(|\phi_01\rangle - |\phi_00\rangle) = \\ & -|\phi_0\rangle \otimes H|1\rangle \rightarrow I \otimes H(-|\phi_0\rangle \otimes H|1\rangle) = -|\phi_0\rangle \otimes H^2|1\rangle = -|\phi_0\rangle \otimes |1\rangle = U_0|\phi_0\rangle \otimes |1\rangle, \\ \text{Si } b \in \mathcal{B}_n - \{b_0\}, \quad & \frac{1}{\sqrt{2}}(|b0\rangle - |b1\rangle) \rightarrow \frac{1}{\sqrt{2}}(U|b0\rangle - U|b1\rangle) = \frac{1}{\sqrt{2}}(|b0\rangle - |b1\rangle) = \\ & = |b\rangle \otimes H|1\rangle \rightarrow I \otimes H(|b\rangle \otimes H|1\rangle) = |b\rangle \otimes H^2|1\rangle = |b\rangle \otimes |1\rangle = U_0|b\rangle \otimes |1\rangle. \end{aligned}$$

En lo que sigue, estudiamos el comportamiento de D (*operador de difusión de Grover*) y U_0 conjuntamente eligiendo convenientemente las bases de \mathbb{C}^N para expresar sus matrices.

Como $|\phi_1\rangle = 2^{-n/2} \sum_{b \in \mathcal{B}_n} |b\rangle$ y $|\phi_0\rangle \in \mathcal{B}_n$, $\mathfrak{B}_2 = \{|\phi_0\rangle, |\phi_1\rangle\}$ es base de un subespacio de dimensión 2, cuyo complemento ortogonal es de dimensión $2^n - 2$. Así, \mathbb{C}^N es suma directa de ambos y $\mathfrak{B} = \{|\phi_0\rangle, |\phi_1\rangle, |\phi_2\rangle, \dots, |\phi_{2^n-1}\rangle\}$ es una base de todo

¹ I denota la matriz identidad y $*$ la traspuesta conjugada, de forma que $\|\vec{v}\|^2 = \langle\vec{v}|\vec{v}\rangle = \vec{v}^*\vec{v}$.

el espacio donde $|\phi_j\rangle \perp \{|\phi_0\rangle, |\phi_1\rangle\}$, $j = 2, \dots, 2^n - 1$. Sus imágenes por DU_0 son:

$$\begin{aligned} \text{Si } j \geq 2, \quad U_0 |\phi_j\rangle &= |\phi_j\rangle \Rightarrow DU_0 |\phi_j\rangle = -|\phi_j\rangle = (0, 0, \dots, \overset{j}{-1}, 0, \dots, 0)_{\mathfrak{B}}, \\ \text{Si } j = 1, \text{ para todo } b \in \mathcal{B}_n, \quad D |b\rangle &= -|b\rangle + 2\langle b|\phi_1\rangle |\phi_1\rangle = -|b\rangle + 2 \cdot 2^{-n/2} |\phi_1\rangle, \text{ y} \\ U_0 |\phi_1\rangle &= 2^{-\frac{n}{2}} (-|\phi_0\rangle + \sum_{b \neq b_0} |b\rangle) \Rightarrow DU_0 |\phi_1\rangle = 2^{-\frac{n}{2}} (|\phi_0\rangle - \sum_{b \neq b_0} |b\rangle + 2 \cdot 2^{-\frac{n}{2}} (2^n - 2) |\phi_1\rangle) \\ &= 2 \cdot 2^{-\frac{n}{2}} |\phi_0\rangle + (-1 + 2 - 2^{2-n}) |\phi_1\rangle = (2^{1-\frac{n}{2}}, 1 - 2^{2-n}, 0, \dots, 0)_{\mathfrak{B}}, \\ \text{Si } j = 0, \quad U_0 |\phi_0\rangle &= -|\phi_0\rangle \Rightarrow DU_0 |\phi_0\rangle = |\phi_0\rangle - 2 \cdot 2^{-\frac{n}{2}} |\phi_1\rangle = (1, -2^{1-\frac{n}{2}}, 0, \dots, 0)_{\mathfrak{B}}. \end{aligned}$$

Expresado en forma matricial, el operador DU_0 en la base \mathfrak{B} es²:

$$\left(\begin{array}{c|c} G & O_{2, 2^{n-2}} \\ \hline O_{2^{n-2}, 2} & -I_{2^{n-2}} \end{array} \right) \quad \text{donde} \quad G = \begin{pmatrix} 1 & 2^{1-n/2} \\ -2^{1-n/2} & 1 - 2^{2-n} \end{pmatrix}.$$

Expresemos DU_0 en términos de una base que simplifique el cálculo de las k composiciones. Tomemos $\mathfrak{B}' = \{|\psi_0\rangle, \dots, |\psi_{2^n-1}\rangle\}$ la base \mathfrak{B} ortonormalizada por el proceso de Gram-Schmidt. El conjunto $\mathfrak{B}'_2 = \{|\psi_0\rangle, |\psi_1\rangle\}$ correspondiente a \mathfrak{B}_2 ortonormalizado genera el mismo subespacio que este: por la estructura del algoritmo, como $|\phi_0\rangle = |b_0\rangle$ normal y $|\psi_1\rangle$ es la proyección (normalizada) de $|\phi_1\rangle$ sobre el subespacio ortogonal a $|\psi_0\rangle$, tenemos:

$$(4.1) \quad |b_0\rangle = |\phi_0\rangle = |\psi_0\rangle; \quad |\psi_1\rangle = \frac{2^{-\frac{n}{2}} \sum_{b \in \mathcal{B}_n - \{b_0\}} |b\rangle}{\|2^{-\frac{n}{2}} \sum_{b \in \mathcal{B}_n - \{b_0\}} |b\rangle\|} = \frac{\sum_{b \in \mathcal{B}_n - \{b_0\}} |b\rangle}{\sqrt{2^n - 1}} = \frac{2^{\frac{n}{2}} |\phi_1\rangle - |\phi_0\rangle}{\sqrt{2^n - 1}}.$$

Así, $DU_0 |\psi_j\rangle = D |\psi_j\rangle = -|\psi_j\rangle$ por su ortogonalidad respecto a \mathfrak{B}'_2 . Como $|\psi_0\rangle, |\psi_1\rangle$ dependen únicamente de $|\phi_0\rangle, |\phi_1\rangle$, basta calcular la matriz G en la base \mathfrak{B}'_2 para obtener DU_0 en la base \mathfrak{B}' . Sea C^{-1} (resp. C) la matriz de cambio de base de \mathfrak{B} a \mathfrak{B}' (resp. de \mathfrak{B}' a \mathfrak{B}), calculadas a partir de (4.1), denotando $'$ a las matrices en \mathfrak{B}' :

$$G' = C^{-1}GC = \begin{pmatrix} 1 - 2^{1-n} & \frac{\sqrt{2^n-1}}{2^{n-1}} \\ \frac{-\sqrt{2^n-1}}{2^{n-1}} & 1 - 2^{1-n} \end{pmatrix} \Rightarrow (DU_0)' = \left(\begin{array}{c|c} G' & O_{2, 2^{n-2}} \\ \hline O_{2^{n-2}, 2} & (-1)^k I_{2^{n-2}} \end{array} \right).$$

En cualquiera de las bases, tomar potencias de DU_0 (aplicarlo tantas veces como indique el exponente) solo afecta a I y a G , luego $((DU_0)')^k = \left(\begin{array}{c|c} (G')^k & O_{2, 2^{n-2}} \\ \hline O_{2^{n-2}, 2} & (-1)^k I_{2^{n-2}} \end{array} \right)$.

Notando que si $\alpha = -2 \arcsen(2^{-\frac{n}{2}}) \Rightarrow \cos \alpha = 1 - 2 \sen^2(\arcsen(2^{-\frac{n}{2}})) = 1 - 2^{1-n}$ y $\sen \alpha = \sqrt{1 - \cos^2 \alpha} = -\frac{\sqrt{2^n-1}}{2^{n-1}}$, concluimos que G' es un giro de ángulo α en la base \mathfrak{B}'_2 .

Es decir, $G' = \begin{pmatrix} \cos \alpha & -\sen \alpha \\ \sen \alpha & \cos \alpha \end{pmatrix}$ y por tanto $(G')^k = \begin{pmatrix} \cos(k\alpha) & -\sen(k\alpha) \\ \sen(k\alpha) & \cos(k\alpha) \end{pmatrix}$.

Como DU_0 es la misma aplicación que DU_0' , obtenemos:

$$(4.2) \quad (DU_0)^k |\psi_0\rangle = \cos(k\alpha) |\psi_0\rangle + \sen(k\alpha) |\psi_1\rangle; \quad (DU_0)^k |\psi_1\rangle = -\sen(k\alpha) |\psi_0\rangle + \cos(k\alpha) |\psi_1\rangle.$$

² I denota la matriz identidad de dimensión m y $0_{m,k}$ la matriz nula de dimensiones $m \times k$.

Esperamos que el output del algoritmo sea $|b_0\rangle = |\phi_0\rangle$ con una probabilidad alta (conseguida eligiendo k). Esto se materializa en el último paso del circuito, que consiste en realizar sobre $(DU_0)^k |\phi_1\rangle$ la medición M sobre los registros de la base \mathcal{B}_n , es decir, $M = \{P_{|b_j\rangle}\}_{j=0}^{2^n-1}$. Por el Postulado 3, la probabilidad de colapso en $|\phi_0\rangle$ es

$p = \langle (DU_0)^k |\phi_1\rangle | P_{|\phi_0\rangle} | (DU_0)^k |\phi_1\rangle \rangle$. Por (4.1) y notando que $\sin(\frac{\alpha}{2}) = -2^{-\frac{n}{2}}$ y $\cos(\frac{\alpha}{2}) = 2^{-\frac{n}{2}} \sqrt{2^n - 1}$, nos queda que $|\phi_1\rangle = \cos(\frac{\alpha}{2}) |\psi_1\rangle - \sin(\frac{\alpha}{2}) |\phi_0\rangle$.

Por linealidad, $(DU_0)^k |\phi_1\rangle = \cos(\frac{\alpha}{2})(DU_0)^k |\psi_1\rangle - \sin(\frac{\alpha}{2})(DU_0)^k |\phi_0\rangle$, y por (4.2) esto es $-\sin(k\alpha + \frac{\alpha}{2}) |\phi_0\rangle + \cos(k\alpha + \frac{\alpha}{2}) |\psi_1\rangle$. Por la ortonormalidad de \mathfrak{B}'_2 , se tiene que $P_{|\phi_0\rangle}(DU_0)^k |\phi_1\rangle = \langle (DU_0)^k |\phi_1\rangle | \phi_0 \rangle |\phi_0\rangle = -\sin(k\alpha + \frac{\alpha}{2}) |\phi_0\rangle$ y entonces por el Postulado 3 obtenemos $p = \sin^2((k + \frac{1}{2})\alpha)$.

Conociendo el valor de p , justifiquemos el valor óptimo de k de manera heurística. Buscamos $p \approx 1$, o equivalentemente, $(k + \frac{1}{2})\alpha \approx \pm\frac{\pi}{2}$. Puesto que $\alpha < 0$ y $k \in \mathbb{N}$, queremos que $k \approx -\frac{\pi}{2\alpha} - \frac{1}{2}$. Cuantas menos iteraciones, más eficiente es el algoritmo, luego k es el menor entero cercano a $-\frac{\pi}{2\alpha} - \frac{1}{2}$, es decir, $k = \lfloor -\frac{\pi}{2\alpha} \rfloor$. Con esto la probabilidad de obtener $|\phi_0\rangle$ en el output es $p \geq 1 - 2^{-n}$ ($p \approx 1$). En efecto, sea $\delta = -k - \frac{\pi}{2\alpha} \in [0, \frac{1}{2}]$, $p = \sin^2((-\delta - \frac{\pi}{2\alpha} + \frac{1}{2})\alpha) = \sin^2(-\frac{\pi}{2} + (\frac{1}{2} - \delta)\alpha) = \cos^2((\frac{1}{2} - \delta)\alpha)$. Como $|(\frac{1}{2} - \delta)\alpha| \leq |\frac{\alpha}{2}|$ y $\alpha \approx 0$, tenemos que $\cos((\frac{1}{2} - \delta)\alpha) \geq \cos(\frac{\pi}{2})$ y por tanto $p = \cos^2((\frac{1}{2} - \delta)\alpha) \geq \cos^2(\frac{\alpha}{2}) = 2^{-n}(2^n - 1) = 1 - 2^{-n}$.

Por último, $\arcsin(x) \geq x$ si $x \in [0, 1]$ implica que $\alpha = -2 \arcsin(2^{-\frac{n}{2}}) \leq -2 \cdot 2^{-\frac{n}{2}}$, por lo que $k \leq \frac{\pi}{-2\alpha} \leq \frac{\pi}{4 \cdot 2^{-n/2}} \leq 2^{n/2}$. Esto respalda el comentario inicial de que en el algoritmo de Grover bastan $2^{n/2}$ llamadas a la base de datos, a diferencia de las 2^{n-1} del algoritmo clásico.

4.2. Información cuántica

La *información cuántica* es la información del estado de un sistema cuántico, aunque el término suele referirse al uso de principios de la mecánica cuántica para procesar y almacenar información. En lo que sigue estudiaremos la imposibilidad de que exista una máquina que copie perfectamente estados cuánticos arbitrarios pero que, en aparente contradicción, existe un protocolo para teletransportar estados cuánticos.

En 1980, después del desarrollo de técnicas para manipular sistemas de un solo estado, creció el interés por aprovecharse de estados entrelazados para contradecir la Teoría de la Relatividad de Einstein, pues de poder clonar estados cuánticos se podrían usar estados entrelazados (compuestos por partículas lejanas) para transmitir información más rápido que la luz. Dicha posibilidad se eliminó con el *teorema de no-clonación* (*no-cloning theorem* en inglés). Su enunciado: «Un estado cuántico desconocido no puede clonarse con transformaciones unitarias»; es decir, dados dos sistemas cuánticos de la misma “dimensión”, no hay forma de clonar el primero en el segundo mediante puertas cuánticas. En términos matemáticos,

sea $|\phi_0\rangle \in \mathbb{C}^{2^n}$ un estado (normalizado), no existe un operador unitario U tal que $U(|\phi\rangle \otimes |\phi_0\rangle) = |\phi\rangle \otimes |\phi\rangle$ para todo $|\phi\rangle \in \mathbb{C}^{2^n}$.³

Demostración:

Sea $|\phi_0\rangle \in \mathbb{C}^{2^n}$ un estado normalizado, supongamos que existe un operador unitario (y lineal) U tal que $U(|\phi\rangle \otimes |\phi_0\rangle) = |\phi\rangle \otimes |\phi\rangle, \forall |\phi\rangle \in \mathbb{C}^{2^n}$. Las matrices unitarias preservan el producto escalar, que junto con la clonación de U implica que dados $|\psi\rangle, |\phi\rangle \in \mathbb{C}^{2^n}$ estados normalizados,

$$\langle |\phi\rangle \otimes |\phi\rangle | |\psi\rangle \otimes |\psi\rangle \rangle = \langle U(|\phi\rangle \otimes |\phi_0\rangle) | U(|\psi\rangle \otimes |\phi_0\rangle) \rangle = \langle |\phi\rangle \otimes |\phi_0\rangle | |\psi\rangle \otimes |\phi_0\rangle \rangle.$$

Por lo visto en [§2.3] del producto escalar en el producto tensorial, como $\|\phi_0\| = 1$,

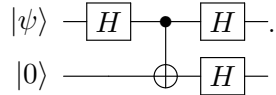
$$\langle \phi | \psi \rangle^2 = \langle |\phi\rangle \otimes |\phi\rangle | |\psi\rangle \otimes |\psi\rangle \rangle = \langle |\phi\rangle \otimes |\phi_0\rangle | |\psi\rangle \otimes |\phi_0\rangle \rangle = \langle \phi | \psi \rangle \langle \phi_0 | \phi_0 \rangle = \langle \phi | \psi \rangle.$$

$$\text{Por lo que } \langle \phi | \psi \rangle (\langle \phi | \psi \rangle - 1) = 0 \Leftrightarrow \langle \phi | \psi \rangle = 0 \text{ o } \langle \phi | \psi \rangle = 1.$$

En el primer caso, los estados son ortogonales. En el segundo caso, por la desigualdad de Cauchy-Schwartz, $1 = |\langle \phi | \psi \rangle| \leq \|\phi\| \|\psi\| = 1$, pero la igualdad se alcanza si los estados son linealmente dependientes: existe $c \in \mathbb{C}$ con $\|c\| = 1$ tal que $c|\psi\rangle = |\phi\rangle$. En este último caso, por generalidades de la mecánica cuántica ya mencionadas en [§2.1], los estados serían indistinguibles (si no interactúan con otros estados).

Así, $|\phi\rangle, |\psi\rangle$ no son arbitrarios sino que deben ser ortogonales o indistinguibles, pero no todos los estados de \mathbb{C}^{2^n} satisfacen dichas relaciones: no puede existir tal U . \square

Observar que en la demostración se concluye que la puerta U no impide clonar estados ortogonales entre sí. Es más, se puede probar que hay puertas cuánticas que clonan estados ortogonales entre sí. Un ejemplo sencillo es el siguiente. Sean $|\phi\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle, |\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$ dos qubits ortogonales⁴, definimos el circuito (operador unitario por ser composición de unitarios):



En el primer paso obtenemos $H^2|0\rangle \otimes |0\rangle = |0\rangle \otimes |0\rangle$. En el segundo no varía y en el último se tiene $H|0\rangle \otimes H|0\rangle = |\phi\rangle \otimes |\phi\rangle$, consiguiendo clonarlo. Si en vez de $|\phi\rangle$ pusiéramos $|\psi\rangle$, tendríamos de forma similar $|1\rangle \otimes |0\rangle$ en el primer paso, $|1\rangle \otimes |1\rangle$ en el segundo y $H|1\rangle \otimes H|1\rangle = |\psi\rangle \otimes |\psi\rangle$ en el output, también clonado. Este resultado explica por qué los bits clásicos sí que pueden clonarse. En particular, la puerta CNOT clona los bits tomando $|0\rangle$ como prefijado:

$$\begin{array}{ccc} |0\rangle & \text{---} \bullet & |0\rangle \\ |0\rangle & \text{---} \oplus & |0\rangle \end{array} ; \quad \begin{array}{ccc} |1\rangle & \text{---} \bullet & |1\rangle \\ |0\rangle & \text{---} \oplus & |1\rangle \end{array}.$$

Una pregunta natural es la de permitir clonaciones “imperfectas”. Puesto que estados que difieren en multiplicar por una constante de módulo 1 son indistinguibles (salvo interacciones con otros estados), podríamos plantearnos un operador que clone los estados módulo multiplicación por cierta constante $c, \|c\| = 1$, permitiendo

³Observar que los estados a clonar y el prefijado han de ser ambos n -registros, pues U opera sobre espacios de llegada y salida con misma dimensión y en el de llegada se hace la clonación ($\mathbb{C}^{2^n} \otimes \mathbb{C}^{2^n}$).

⁴Notar que $|\phi\rangle = H|0\rangle, |\psi\rangle = H|1\rangle$ y $H^2 = I$ por lo ya visto de la puerta de Hadamard.

que dependa del estado a clonar e incluso del prefijado. La respuesta sigue siendo negativa: sea $|\phi_0\rangle \in \mathbb{C}^{2^n}$ un estado (normalizado), no existe un operador unitario U tal que $U(|\phi\rangle \otimes |\phi_0\rangle) = c|\phi\rangle \otimes |\phi\rangle$ para todo $|\phi\rangle \in \mathbb{C}^{2^n}$, con $\|c\| = 1$ una constante no necesariamente independiente de $|\phi_0\rangle, |\phi\rangle$.

Demostración:

Como antes, sea $|\phi_0\rangle \in \mathbb{C}^{2^n}$ un estado normalizado y U un operador unitario tal que dados $|\phi\rangle, |\psi\rangle \in \mathbb{C}^{2^n}$, existen constantes c y d con $\|c\| = \|d\| = 1$ tales que $U(|\phi\rangle \otimes |\phi_0\rangle) = c|\phi\rangle \otimes |\phi\rangle$, $U(|\psi\rangle \otimes |\phi_0\rangle) = d|\psi\rangle \otimes |\psi\rangle$. U preserva el producto escalar:

$$\langle c|\phi\rangle \otimes |\phi| d|\psi\rangle \otimes |\psi\rangle\rangle = \langle U(|\phi\rangle \otimes |\phi_0\rangle) | U(|\psi\rangle \otimes |\phi_0\rangle) \rangle = \langle |\phi\rangle \otimes |\phi_0\rangle | |\psi\rangle \otimes |\phi_0\rangle \rangle \Leftrightarrow \\ c\bar{d}\langle\phi|\psi\rangle^2 = \langle |\phi\rangle \otimes |\phi\rangle | |\psi\rangle \otimes |\psi\rangle \rangle = \langle |\phi\rangle \otimes |\phi_0\rangle | |\psi\rangle \otimes |\phi_0\rangle \rangle = \langle\phi|\psi\rangle\langle\phi_0|\phi_0\rangle = \langle\phi|\psi\rangle.$$

$$\text{Tomando módulos, } |c\bar{d}\langle\phi|\psi\rangle^2| = |\langle\phi|\psi\rangle|^2 = |\langle\phi|\psi\rangle| \Leftrightarrow |\langle\phi|\psi\rangle|(|\langle\phi|\psi\rangle| - 1) = 0.$$

Y llegamos a la misma conclusión: $|\langle\phi|\psi\rangle| = \langle\phi|\psi\rangle = 0$ o $|\langle\phi|\psi\rangle| = 1$. Es decir, los estados son ortogonales o indistinguibles, de nuevo por Cauchy-Schwartz. \square

El teorema de no clonación forma parte de los llamados *no-go theorems*, que establecen imposibilidades físicas, en este caso relacionadas con la mecánica cuántica. El *no-teleportation theorem* prohíbe transformar estados cuánticos arbitrarios en secuencias de bits o reconstruir un qubit a partir de bits (i.e. la teletransportación), teorema que se deduce de la no-clonación pues, de ser posible transformar qubits en bits, estos podrían clonarse y después recuperar el qubit. El *no-deleting theorem* impide eliminar una copia de entre dos creadas de un qubit y el *no-broadcasting theorem* generaliza el de no-clonación al impedir enviar copias de un mismo qubit a distintas localizaciones (aunque se pueda llevar el estado de un lugar a otro por teletransporte cuántico).

La *teletransportación cuántica* “rompe” aparentemente algunos principios físicos, como la imposibilidad de viajar más rápido que la luz o el teorema de no-clonación, pero es algo ilusorio, como quedará explicado en lo que sigue. A grandes rasgos, el teletransporte clásico separa cada componente de un objeto en un lugar (desarma en átomos), midiendo posición y tamaño, y envía la información a otro lugar para recomponer el objeto allí (rearma los átomos). Podría ser que el objeto inicial desapareciera, o podría ser que permaneciera, teniendo dos copias de lo mismo. El teletransporte cuántico, en cambio, no permite que existan dos copias simultáneamente (no-clonación) ni destruye para reconstruir en el nuevo lugar (no-teletransporte), sino que utiliza entrelazamientos y colapsos de mediciones para *enviar* información del primer lugar al segundo, sin conocerla per sé, y requiere la presencia de alguna partícula en el segundo lugar, además de un canal de comunicación.

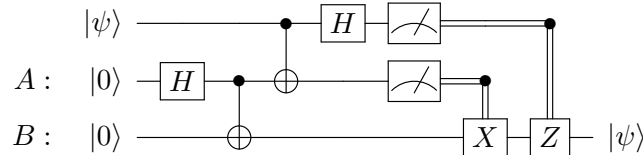
Supongamos que tenemos dos sujetos A y B en dos lugares distintos A y B ⁵. Supongamos que cada sujeto tiene un qubit de un par de qubits entrelazados. Esto no debería traer dificultades: una opción es que A y B , estando juntos, hubieran generado

⁵En teoría, podrían estar arbitrariamente lejos. A nivel experimental, se han superado los 100 km.

un par de qubits entrelazados y después se alejase a los lugares A y B; y otra opción sería que un tercer sujeto crease un par entrelazado y enviara un qubit a cada sujeto en cada lugar por un canal cuántico. Aquí tomaremos la primera opción para recalcar el hecho de que la teletransportación cuántica puede llevarse a cabo usando únicamente un canal de comunicación convencional, para que A pueda comunicar a B el resultado de una medición y con ello concluir.

La idea del proceso es la siguiente. A quiere enviar a B la información del qubit $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. Para ello, entrelaza su qubit con el nuevo y realiza una medición sobre dicho par entrelazado. Como el qubit de B está entrelazado con el de A, la medición también cambiará su estado. Así, lo que A obtenga tras la medición será comunicado a B (llamada, mensaje, etc) y este implementará alguna transformación al suyo con la que obtendrá $|\psi\rangle$ con total seguridad.

Matemáticamente podemos representar el algoritmo con un circuito. Aquí suponemos que el par entrelazado que A y B comparten es el estado de Bell $|\Phi_0\rangle$, definido en [§2.4]. El circuito es:



En la primera parte, A y B entrelazan los qubits para obtener $|\Phi_0\rangle$: como vimos en [§3.1], la puerta de Hadamard y la CNOT generan los estados de Bell y puesto que la CNOT deja invariante el registro $|00\rangle$, podemos omitir dicha puerta en la primera columna. Por tanto, la tercera columna es $|\psi\rangle \otimes |\Phi_0\rangle = (\alpha|0\rangle + \beta|1\rangle) \otimes \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Aplicando la puerta CNOT sobre los dos primeros qubits, obtenemos en la cuarta columna $\frac{1}{\sqrt{2}}[\alpha|0\rangle(|00\rangle + |11\rangle) + \beta|1\rangle(|10\rangle + |01\rangle)]$. Aplicando lo visto en [§3.2] de la puerta de Hadamard, al actuar sobre el primer qubit deja el registro como $\frac{1}{2}[\alpha(|0\rangle + |1\rangle)(|00\rangle + |11\rangle) + \beta(|0\rangle - |1\rangle)(|10\rangle + |01\rangle)]$. Por la construcción de los registros, esto es $\alpha|000\rangle + \beta|001\rangle + \alpha|011\rangle + \beta|010\rangle + \alpha|100\rangle - \beta|101\rangle + \alpha|111\rangle - \beta|110\rangle$, y reagrupando términos, queda en la quinta columna:

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

En el penúltimo paso hacemos una medición sobre la base \mathcal{B}_2 , reflejado, como en otros circuitos, con las mediciones de cada uno de los dos primeros qubits sobre los qubits $|0\rangle$ y $|1\rangle$. Es claro que, tras la medición, por el Postulado 3 y la ortonormalidad de los qubits, el registro colapsa en uno de los sumandos módulo una constante no nula, específicamente en⁶:

$$\frac{\frac{1}{2} |b\rangle (\alpha|c\rangle \pm \beta|1 \oplus c\rangle)}{\| \frac{1}{2} |b\rangle (\alpha|c\rangle \pm \beta|1 \oplus c\rangle) \|} = |b\rangle (\alpha|c\rangle \pm \beta|1 \oplus c\rangle), \quad \text{donde } b \in \mathcal{B}_2, c \in \{0, 1\}.$$

Tras la medición, el tercer qubit –que tiene B– será $\alpha|c\rangle \pm \beta|1 \oplus c\rangle$ con los valores determinados por el resultado. Establecemos que los resultados de la medición son 0, 1, 2, 3, cada uno correspondiendo a que los dos primeros qubits colapsan en dicho valor escrito en binario. Sabiendo el resultado de A, B conoce la disposición de su

⁶En particular, los resultados posibles de la medición son equiprobables: $p(b) = \frac{1}{4}$, $b \in \mathcal{B}_2$.

qubit e , implementando cierta transformación, recupera el estado $|\psi\rangle$. Ahora, por un canal de comunicación convencional, A envía $b \in \{0, 1, 2, 3\}$ a B .

Si A envía 0 entonces B tiene el qubit $|\psi\rangle$. Si envía 1, tiene $\alpha|1\rangle + \beta|0\rangle$, que es aplicar la puerta NOT= X a $|\psi\rangle$. Si envía 2, tiene el coeficiente de $|1\rangle$ cambiado de signo, resultado de aplicar la puerta Z a $|\psi\rangle$. Si envía 3, tiene el resultado de cambiar el signo del coeficiente $|1\rangle$ en $|\psi\rangle$ y aplicar NOT, es decir, $XZ|\psi\rangle$ (§3.1). Para recuperar $|\psi\rangle$, como las puertas vienen de operadores unitarios, B deberá aplicarlas (en el orden correcto) de nuevo sobre el resultado. Esto viene codificado en la última parte del circuito: después de cada caja de medición hay una puerta controlada (§3.1) y así,

A envía 0 \Rightarrow se obtiene $|0\rangle$ en el primer qubit y $|0\rangle$ en el segundo, así que B aplica la identidad I . Notar que $I = Z^0 X^0$, $00 = 0$ en binario.

A envía 1 \Rightarrow se obtiene $|0\rangle$ en el primer qubit y $|1\rangle$ en el segundo, así que B aplica la puerta X . Notar que $X = Z^0 X^1$, $01 = 1$ en binario.

A envía 2 \Rightarrow se obtiene $|1\rangle$ en el primer qubit y $|0\rangle$ en el segundo, así que B aplica la puerta Z . Notar que $Z = Z^1 X^0$, $10 = 2$ en binario.

A envía 3 \Rightarrow se obtiene $|1\rangle$ en el primer qubit y $|1\rangle$ en el segundo, así que B aplica la puerta XZ . Notar que $ZX = Z^1 X^1$, $11 = 3$ en binario.

Finalmente, el qubit de B es $|\Psi\rangle$, como buscábamos.

Conviene hacer algunas observaciones del algoritmo:

- 1) El qubit $|\psi\rangle$ es arbitrario, del que ni A ni B saben nada, pues para ello deberían hacer mediciones sobre él: su información se envía de A en B sin conocerse.
- 2) No hay un teletransporte de A a B del qubit como ocurre en la contrapartida clásica, pues no se envía una partícula, no hay masa, sino que se transmite la información necesaria a B para que reproduzca en su qubit el estado de $|\psi\rangle$.
- 3) No se realiza en B una copia del qubit $|\psi\rangle$, que está en A porque al final del proceso en A queda un registro de \mathcal{B}_2 , donde se ha perdido cualquier información de $|\psi\rangle$. No se contradice el teorema de no-clonación.
- 4) La única información que se envía es por un canal clásico, que si fuera un mensaje requeriría de dos bits. Además, A y B no tienen por qué establecer qué debe hacer B en función del mensaje de A , pues conociendo el resultado de la medición B sabe la disposición de su qubit (dando por hecho que conoce la teoría). Cabe destacar que a partir de información probabilista de la medición se “reconstruye” con total precisión $|\psi\rangle$ y que si alguien espicara el canal entre A y B sabría el proceso de B para recuperar $|\psi\rangle$, pero no podría conocer $|\psi\rangle$ si no puede interactuar con la partícula de B .
- 5) Parece que el teletransporte cuántico permite transmitir información más rápido que la luz, aplicando puertas cuánticas y realizando mediciones, pero esto es solo una aparente paradoja porque el resultado de la medición debe comunicarse por un canal convencional⁷: la ilusoria contradicción con la Teoría de la Relatividad queda resuelta.

⁷Se puede probar teóricamente que el teletransporte sin un canal clásico no transmite información alguna y por tanto A necesita comunicar a B el resultado con uno de este tipo.

CAPÍTULO 5

El Algoritmo de Shor

En esta sección se estudia el *algoritmo de Shor*, introducido por P.W. Shor en 1977 y en especial su componente cuántico. De poder implementarse, el algoritmo permitiría descomponer cualquier entero N en sus factores primos en tiempo polinomial ($\log N$) y romper la base de la mayoría de sistemas criptográficos de hoy en día. Hasta la fecha, el mayor número factorizado con el algoritmo es 21: el reto sigue siendo conseguir qubits menos ruidosos que permitan algoritmos que detecten y corrijan errores, algo que requiere una inmensa cantidad de qubits.

5.1. Motivación del problema y el algoritmo de Shor

Un criptosistema de clave pública es aquel en el que, en vez de generar una clave para cada par de interlocutores en el foro (2^n claves, n individuos), se generan (n) claves públicas de tal forma que si A quiere enviar un mensaje a B, deberá cifrarlo conforme a la clave pública de B y este utilizará su clave privada para descifrarlo. El criptosistema RSA es un ejemplo, en el que la clave pública de un individuo B consiste en el par (e_B, N_B) donde N_B es el módulo y es producto de dos primos distintos (cuanto mayor sean y mayor distancia entre ellos, más difícil su factorización) y e_B es coprimo con $\phi(N)$ (función *phi de Euler*). Si el alfabeto es de M letras, se elige k entero tal que $M^k < N_B < M^{k+1}$ y se cifra el mensaje en k -tuplas. Si A quiere enviar a B la k -tupla m (como mensaje numérico), A calcula $c \equiv m^{e_B} \pmod{N_B}$. Para descifrarlo, B conoce los factores primos $N_B = pq$ y por tanto conoce $\phi(N_B) = (p-1)(q-1)$. Con el algoritmo de Euclides, B halla (tiempo polinomial) d tal que $e_B d_B \equiv 1 \pmod{\phi(N_B)}$. Gracias a la congruencia de Euler, si pq y m son coprimos, $m^{\phi(pq)} \equiv 1 \pmod{pq}$ y por tanto $m^{e_B d_B} \equiv m \pmod{pq}$, también en el caso $pq \mid m$. Así, B debe calcular $c^{d_B} \pmod{N_B}$ para descifrar el mensaje. El criptosistema RSA es muy seguro imponiendo esas dos condiciones sobre los factores primos, pues para encontrarlos solo se conocen algoritmos de tiempo exponencial. Sin embargo, el algoritmo de Shor, haciendo uso de la computación cuántica, podría *romper* el criptosistema en tiempo polinomial. El mayor impedimento, a diferencia de los algoritmos tratados con anterioridad, no es teórico (no hay oráculos no descritos) sino práctico, pues se requieren muchos qubits entrelazados y puertas cuánticas que funcionen libres de errores y ruido durante cierto tiempo.

En lo que sigue, utilizamos la notación \gcd para el máximo común divisor y el cálculo de potencias (mód N) con exponenciación modular, eficiente con computación clásica. El algoritmo de Shor para encontrar un factor del entero N es el siguiente:

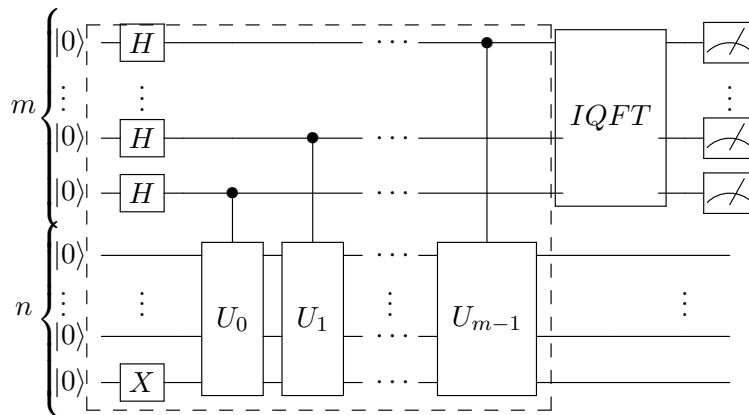
1. Se elige un entero $a \in [2, N)$. Si $\gcd(a, N) \neq 1$ obtenemos un factor no trivial de N y el algoritmo termina.
2. Se calcula el orden r de a en $(\mathbb{Z}/N\mathbb{Z})^*$: $r = \min\{s \in \mathbb{Z}^+ : a^s \equiv 1 \pmod{N}\}$.
3. Si r es impar o r es par con $a^{r/2} \equiv -1 \pmod{N}$, el algoritmo no funciona y empezamos de nuevo. En caso contrario, r par tal que $a^r \equiv 1 \pmod{N}$, es decir, $N \mid (a^r - 1) = (a^{r/2} - 1)(a^{r/2} + 1)$ y puesto que $N \nmid (a^{r/2} + 1)$, el algoritmo termina habiendo obtenido $\gcd(a^{r/2} + 1, N)$ como factor no trivial de N .

Por ejemplo, con datos $N = 15$, $a = 7$, se calcula $a^2 \equiv 4 \pmod{N}$, $a^3 \equiv -2 \pmod{N}$ y $a^4 \equiv 1 \pmod{N}$. Entonces $r = 4$ es orden par tal que $a^{r/2} \equiv 4 \not\equiv -1 \pmod{N}$, por lo que $\gcd(N, a^{r/2} + 1) = \gcd(15, 50) = 5$ es un factor no trivial de N .

5.2. Estimación de fase cuántica y conclusiones

El segundo paso del algoritmo es el único que requiere de la computación cuántica para resolverse de manera eficiente. Se trata de calcular el orden r módulo N del entero escogido a , coprimo con N . Esto se lleva a cabo mediante un circuito cuántico basado en el algoritmo de estimación de fase cuántica (en inglés, *Quantum phase estimation: QPE*). El QPE devuelve, dada una puerta U y un autovector $|\phi\rangle$, el valor exacto de θ donde $e^{2\pi i\theta}$ es el autovalor¹. En el algoritmo de Shor, en vez de un autovector $|\phi\rangle$ tenemos $|1\rangle$ como una suma de autovectores de una matriz U , por lo que el QPE devuelve el valor θ de uno de los autovalores.

El circuito es el siguiente:



¹El autovalor es de esta forma porque U es una matriz unitaria, luego todos los autovalores son de módulo 1. de $|\phi\rangle$

Las puertas X y H son las ya empleadas NOT y Hadamard, y los pasos finales son mediciones sobre los primeros m qubits. Se eligen n tal que $2^n > N$ y $m = 2n + 1$. En el QPE, el input sería $|0\rangle^{\otimes m} \otimes |\phi\rangle$ y no habría puerta X . La matriz U se usaría para construir las puertas U_s e $IQFT$ es la *Inverse Quantum Fourier Transform*.

En el circuito y su análisis utilizaremos una variante del *núcleo de Dirichlet*. Dado $M \in \mathbb{N}$, se define para cada $x \in \mathbb{R}$:

$$(5.1) \quad D_M(x) = \frac{1}{M} \sum_{k=0}^{M-1} e^{2\pi i k x}. \quad \text{En particular, si } x \notin \mathbb{Z} \text{ entonces } e^{i2\pi x} \neq 1 \text{ y}$$

$$D_M(x) = \frac{1 - e^{i2\pi Mx}}{M(1 - e^{i2\pi x})} \implies |D_M(x)|^2 = \frac{2 - 2\cos(2\pi Mx)}{M^2(2 - 2\cos(2\pi x))} = \frac{\text{sen}^2(\pi Mx)}{M^2 \text{sen}^2(\pi x)}.$$

Sean $j, \ell, r \in \mathbb{N}$, $D_{2^m} \left(\frac{j}{r} - \frac{\ell}{2^m} \right) = \frac{1}{2^m} \sum_{k=0}^{2^m-1} e^{\frac{2\pi i k j}{r}} e^{-\frac{2\pi i k \ell}{2^m}}$; $D_r \left(\frac{-\ell}{r} \right) = 0$, $0 \leq \ell < r$.

En particular, si $|x| \leq \frac{1}{2}$ entonces por la *desigualdad de Jordan*, $\text{sen}^2(\pi x) \geq 4x^2$ y

$$(5.2) \quad |D_M(x)|^2 \leq \frac{1}{4M^2x^2}, \quad \text{si } |x| \leq \frac{1}{2}.$$

Se define $\|x\|$ como la menor distancia de x a un entero. Así, dado $\epsilon \in [0, \frac{1}{2}]$, se tiene que $\|x\| \leq \epsilon$ si y solo si $|x| \leq \epsilon$ extendida de forma 1-periódica. Notando que las funciones $\text{sen}^2(\pi x)$, $\text{sen}^2(\pi Mx)$ son 1-periódicas, si $\|x\| \leq \frac{1}{2}$ entonces $|x| \leq \frac{1}{2}$ extendida con período 1 y sus valores $|D_M(x)|^2$ coinciden. Esto, junto con (5.2), implica que, como $\|x\| = |x|$ extendido de forma 1-periódica, $|D_M(x)|^2 \leq \frac{1}{4M^2\|x\|^2}$.

Tras el primer paso, como en (3.1), denotando el n -registro $|1\rangle = |0\dots 01\rangle$:

$$|\psi_1\rangle = H^{\otimes m} |0\rangle^{\otimes m} \otimes |0\rangle^{\otimes n-1} \otimes X|0\rangle = 2^{-m/2} \sum_{b \in \mathcal{B}_m} |b\rangle \otimes |0\rangle^{\otimes n-1} \otimes |1\rangle = \frac{1}{2^{m/2}} \sum_{k=0}^{2^m-1} |k\rangle \otimes |1\rangle.$$

Utilizamos $|a^k\rangle$, $k \in \mathbb{Z}^+$ para denotar el n -registro $|R\rangle$ tal que a^k es el resto de dividir R por N . A pesar de no conocer r , podemos expresar estados cuya expresión dependa de r . Se define para cada $j = 0, \dots, r-1$ el estado

$$|u_j\rangle = \frac{1}{\sqrt{r}} \sum_{\ell=0}^{r-1} e^{-2\pi i \ell j / r} |a^\ell\rangle, \quad \text{que está normalizado porque } a^\ell \text{ entero} \implies \langle a^\ell | a^k \rangle = \delta_{\ell, k}.$$

Definimos la aplicación U : dado $\ell \in \mathcal{B}_n$, $U|\ell\rangle = \begin{cases} |a^\ell \pmod{N}\rangle & \text{si } 0 \leq \ell < N, \\ |\ell\rangle & \text{en otro caso,} \end{cases}$ y se

extiende a n -registros por linealidad.

Nótese que a coprimo con N implica que $a\ell \equiv ak \pmod{N} \Leftrightarrow \ell \equiv k \pmod{N}$; en particular, $\ell = k$ si $0 \leq \ell, k < N$. Por lo tanto, $\langle U\ell | Uk \rangle = \delta_{\ell, k} = \langle \ell | k \rangle$ si $0 \leq \ell, k < N$ por ortonormalidad de \mathcal{B}_n y también en el resto de casos, de manera trivial. Se concluye que U es una puerta.

Los estados $|u_j\rangle$ son autovectores² de U y $\frac{1}{\sqrt{r}} \sum_{j=0}^{r-1} |u_j\rangle = |1\rangle$. Como $a^r \equiv a^0$:

$$(5.3) \quad \begin{aligned} U |u_j\rangle &= \frac{1}{\sqrt{r}} \sum_{\ell=0}^{r-1} e^{-2\pi i \ell j / r} |a \cdot a^\ell \pmod{N}\rangle = \frac{1}{\sqrt{r}} \sum_{\ell=1}^r e^{-2\pi i (\ell-1) j / r} |a^\ell \pmod{N}\rangle = \\ &= \frac{e^{2\pi i j / r}}{\sqrt{r}} \left(\sum_{\ell=1}^{r-1} e^{-2\pi i \ell j / r} |a^\ell \pmod{N}\rangle + e^{-2\pi i j} |a^r \pmod{N}\rangle \right) = \\ &= \frac{e^{2\pi i j / r}}{\sqrt{r}} \left(\sum_{\ell=0}^{r-1} e^{-2\pi i \ell j / r} |a^\ell\rangle + 1 |a^0\rangle \right) = e^{2\pi i j / r} |u_j\rangle; \end{aligned}$$

$$\text{además, por (5.1)} \quad \frac{1}{\sqrt{r}} \sum_{j=0}^{r-1} |u_j\rangle = \frac{1}{r} \sum_{j=0}^{r-1} \sum_{\ell=0}^{r-1} e^{-2\pi i \ell j / r} |a^\ell\rangle = \sum_{\ell=0}^{r-1} |a^\ell\rangle D_r \left(-\frac{\ell}{r} \right) = |1\rangle.$$

De esta forma, u_j es autovector de autovalor $e^{2\pi i j / r}$ de U y $|1\rangle$ es una combinación lineal de autovectores de la matriz unitaria U . La intuición es adaptar el QPE a $|1\rangle$ y obtener el valor $\theta_j = j/r$ de alguno de los autovalores $e^{2\pi i j / r}$. Con el algoritmo de la fracción continua podremos aproximar j/r por fracciones irreducibles para encontrar r , como se discute en el Apéndice [(B.2)].

Sigamos entonces con el algoritmo, ahora ya la parte del QPE. Se definen los operadores unitarios $U_s = U^{2^s}$. Nótese que $|k\rangle = U_0 |\ell\rangle = \begin{cases} |a^\ell \pmod{N}\rangle & \text{si } 0 \leq \ell < N, \\ |\ell\rangle & \text{en otro caso.} \end{cases}$

cumple $0 \leq k < N$ si $0 \leq \ell < N$, por lo que $U_1 |\ell\rangle = |ak \pmod{N}\rangle = |a^2 \ell \pmod{N}\rangle$ si $0 \leq \ell < N$ y $U_2 |\ell\rangle = U |\ell\rangle = |\ell\rangle$ en otro caso. Inductivamente se tiene que

$$U_s |\ell\rangle = \begin{cases} |a^{2^s} \ell \pmod{N}\rangle & \text{si } 0 \leq \ell < N, \\ |\ell\rangle & \text{en otro caso.} \end{cases} \quad \text{En el circuito, son puertas controladas}$$

por el $(m-s)$ -ésimo qubit. Así, dado el m -registro $k = b_{m-1}2^{m-1} + \dots + b_12^1 + b_0$, la puerta U_s está controlada por b_s como se explicó en el capítulo 3:

$$U_s |\ell\rangle = \begin{cases} U_s |\ell\rangle & \text{si } b_j = 1 \\ |\ell\rangle & \text{si } b_s = 0 \end{cases} = \begin{cases} |a^{2^j} \ell \pmod{N}\rangle & \text{si } 0 \leq \ell < N \text{ y } b_s = 1, \\ |\ell\rangle & \text{en otro caso.} \end{cases}$$

Como u_j es autovector de autovalor $e^{2\pi i j / r}$ de U , sobre el que actúa de forma no trivial, u_j es también autovector de la puerta controlada U_s :

$$(5.4) \quad U_s |u_j\rangle = \begin{cases} U^{2^s} |u_j\rangle = e^{2\pi i j 2^s / r} |u_j\rangle & \text{si } b_s = 1, \\ |u_j\rangle & \text{si } b_s = 0 \end{cases} \implies U_s |u_j\rangle = e^{2\pi i j 2^s b_s / r} |u_j\rangle.$$

Por lo calculado en el primer paso y en (5.3),

$$|\psi_1\rangle = \frac{1}{2^{m/2}} \sum_{k=0}^{2^m-1} |k\rangle \otimes |1\rangle = \frac{1}{2^{m/2}} \sum_{k=0}^{2^m-1} |k\rangle \otimes \frac{1}{\sqrt{r}} \sum_{j=0}^{r-1} |u_j\rangle = \frac{1}{2^{m/2} \sqrt{r}} \sum_{k=0}^{2^m-1} \sum_{j=0}^{r-1} |k\rangle \otimes |u_j\rangle.$$

²La puerta U no actúa trivialmente sobre $|u_j\rangle$ porque $|a^\ell\rangle$ son restos al dividir por N : $a^\ell < N$.

Podemos concluir, por (5.4), que la parte del circuito encerrada en el rectángulo de puntos da lugar al estado $|\psi_2\rangle = \frac{1}{2^{m/2}\sqrt{r}} \sum_{k=0}^{2^m-1} \sum_{j=0}^{r-1} e^{2\pi i k j / r} |k\rangle \otimes |u_j\rangle$:

(5.5)

$$\begin{aligned} & \text{Dado } k \in \mathcal{B}_m, k = b_{m-1}2^{m-1} + \dots + b_1 2 + b_0, \quad U_{m-1} \cdots U_1 U_0 \left(|k\rangle \otimes \sum_{j=0}^{r-1} |u_j\rangle \right) = \\ & = U_{m-1} \cdots U_1 \left(|k\rangle \otimes \sum_{j=0}^{r-1} e^{\frac{2\pi i j b_0}{r}} |u_j\rangle \right) = U_{m-1} \cdots U_2 \left(|k\rangle \otimes \sum_{j=0}^{r-1} e^{\frac{2\pi i j b_1 2}{r}} e^{\frac{2\pi i j b_0}{r}} |u_j\rangle \right) \\ & = \dots = U_{m-1} \left(|k\rangle \otimes \sum_{j=0}^{r-1} e^{\frac{2\pi i j (b_{m-2}2^{m-2} + \dots + b_1 2 + b_0)}{r}} |u_j\rangle \right) = |k\rangle \otimes \sum_{j=0}^{r-1} e^{\frac{2\pi i j k}{r}} |u_j\rangle. \end{aligned}$$

$$\begin{aligned} \text{Por linealidad, } |\psi_2\rangle & = U_{m-1} \cdots U_0 |\psi_1\rangle = \frac{1}{2^{m/2}\sqrt{r}} \sum_{k=0}^{2^m-1} U_{m-1} \cdots U_0 \left(\sum_{j=0}^{r-1} |k\rangle \otimes |u_j\rangle \right) \\ & = \frac{1}{2^{m/2}\sqrt{r}} \sum_{k=0}^{2^m-1} |k\rangle \otimes \sum_{j=0}^{r-1} e^{\frac{2\pi i j k}{r}} |u_j\rangle. \end{aligned}$$

Definamos la puerta *IQFT*. La *Transformada cuántica de Fourier* es el análogo a la DFT del análisis de Fourier discreto. La secuencia de números en la versión discreta es la base B_m en la versión cuántica, y el cardinal M es en este caso $M = 2^m$. Así, la transformada cuántica inversa de Fourier se define para cada $k \in \mathcal{B}_m$:

$$IQFT = F : |k\rangle \mapsto \frac{1}{2^{m/2}} \sum_{\ell=0}^{2^m-1} e^{-\frac{2\pi i k \ell}{2^m}} |\ell\rangle, \text{ y se extiende a } m\text{-registros por linealidad.}$$

F es unitaria porque preserva el producto escalar. Por linealidad, basta verlo en \mathcal{B}_m :

$$\langle Fj | Fk \rangle = \frac{1}{2^m} \sum_{\ell=0}^{2^m-1} \sum_{t=0}^{2^m-1} e^{\frac{2\pi i (kt-j\ell)}{2^m}} \langle \ell | t \rangle = \frac{1}{2^m} \sum_{\ell=0}^{2^m-1} \left(e^{\frac{2\pi i (k-j)\ell}{2^m}} \right)^\ell = \delta_{j,k} = \langle j | k \rangle.$$

Con esto, calculemos el estado previo a la medición: $(F \otimes I^{\otimes n}) |\psi_2\rangle$. Por (5.1) y (5.5)

$$\begin{aligned} |\psi\rangle & = (F \otimes I^{\otimes n}) |\psi_2\rangle = \frac{1}{2^{m/2}\sqrt{2^m r}} \sum_{k=0}^{2^m-1} \sum_{j=0}^{r-1} e^{\frac{2\pi i j k}{r}} (F \otimes I^{\otimes n}) (|k\rangle \otimes |u_j\rangle) = \\ & = \frac{1}{2^m \sqrt{r}} \sum_{k=0}^{2^m-1} \sum_{j=0}^{r-1} e^{\frac{2\pi i j k}{r}} F |k\rangle \otimes |u_j\rangle = \frac{1}{\sqrt{r}} \sum_{\ell=0}^{2^m-1} \sum_{j=0}^{r-1} \left(\sum_{k=0}^{2^m-1} \frac{1}{2^m} e^{\frac{2\pi i j k}{r}} e^{-\frac{2\pi i k \ell}{2^m}} \right) |\ell\rangle \otimes |u_j\rangle = \\ & = \frac{1}{\sqrt{r}} \sum_{\ell=0}^{2^m-1} \sum_{j=0}^{r-1} D_{2^m} \left(\frac{j}{r} - \frac{\ell}{2^m} \right) |\ell\rangle \otimes |u_j\rangle. \end{aligned}$$

En el último paso, el circuito realiza una medición sobre los primeros m qubits. Por el postulado 3, si el resultado de la medición es $k_0 \in \mathcal{B}_m$, como $\{|u_j\rangle\}_{j=0}^{r-1}$ son

ortogonales por ser autovectores de U unitaria y estar normalizados, $|\psi\rangle$ colapsa en:

$$\frac{1}{\sqrt{r \cdot p_{k_0}}} |k_0\rangle \otimes \sum_{j=0}^{r-1} D_{2^m} \left(\frac{j}{r} - \frac{k_0}{2^m} \right) |u_j\rangle, \quad \text{con probabilidad } p_{k_0} = \frac{1}{r} \sum_{j=0}^{r-1} \left| D_{2^m} \left(\frac{j}{r} - \frac{k_0}{2^m} \right) \right|^2.$$

En efecto, con la notación ya empleada, por (2.3) y la ortonormalidad de \mathcal{B}_m y $\{|u_j\rangle\}$,

$$\begin{aligned} P_{k_0} |\psi\rangle &= \frac{1}{\sqrt{r}} |k_0\rangle \otimes \sum_{j=0}^{r-1} D_{2^m} \left(\frac{j}{r} - \frac{k_0}{2^m} \right) |u_j\rangle, \quad \text{con probabilidad } p_{k_0} = \langle \psi | P_{k_0} | \psi \rangle = \\ &= \frac{1}{r} \sum_{\ell=0}^{2^m-1} \sum_{j=0}^{r-1} D_{2^m} \left(\frac{j}{r} - \frac{\ell}{2^m} \right) \sum_{k=0}^{r-1} D_{2^m} \left(\frac{k}{r} - \frac{k_0}{2^m} \right) \langle \ell | k_0 \rangle \langle u_j | u_k \rangle = \frac{1}{r} \sum_{j=0}^{r-1} \left| D_{2^m} \left(\frac{j}{r} - \frac{k_0}{2^m} \right) \right|^2. \end{aligned}$$

En definitiva, tras la medición el estado colapsa en un estado proporcional a $|k_0\rangle \otimes \sum_{j=0}^{r-1} D_{2^m} \left(\frac{j}{r} - \frac{k_0}{2^m} \right) |u_j\rangle$. En lo que sigue se demuestra que la probabilidad de que los sumandos satisfagan $\|j/r - k_0/2^m\| \geq \frac{1}{2}r^{-2}$ es pequeña.

Si el resultado de la medición es k_0 y colapsa en $|k_0\rangle \otimes \sum_{j \in J_{k_0}} D_{2^m} \left(\frac{j}{r} - \frac{k_0}{2^m} \right) |u_j\rangle$, donde $J_{k_0} = \{\|j/r - k_0/2^m\| \geq \frac{1}{2}r^{-2}\}$, con probabilidad $\frac{1}{r} \sum_{j \in J_{k_0}} \left| D_{2^m} \left(\frac{j}{r} - \frac{k_0}{2^m} \right) \right|^2$. Así, la probabilidad total de colapsar en un estado cuyos sumandos satisfagan esa condición está acotada por:

$$p = \sum_{k=0}^{2^m-1} \frac{1}{r} \sum_{j \in J_k} \left| D_{2^m} \left(\frac{j}{r} - \frac{k}{2^m} \right) \right|^2 = \frac{1}{r} \sum_{k,j} \left| D_{2^m} \left(\frac{j}{r} - \frac{k}{2^m} \right) \right|^2 \leq \frac{1}{4r2^{2m}} \sum_{k,j} \left\| \frac{j}{r} - \frac{k}{2^m} \right\|^{-2},$$

donde la prima indica la restricción de la suma a j, k tales que $\|j/r - k/2^m\| \geq \frac{1}{2}r^{-2}$ y la última desigualdad se sigue de (5.2), pues $\|x\| \leq \frac{1}{2}$ y $|D_{2^m}(x)|^2 \leq \frac{1}{4(2^m)^2 \|x\|}$.

Como $j < r$ y $k < 2^m$, se tiene que $j/r, k/2^m \in [0, 1)$, por lo que $\|j/r - k/2^m\|$ es $|j/r - k/2^m|$ o $|j/r \pm 1 - k/2^m|$. Por ello, suponemos sin pérdida de generalidad que $\|j/r - k/2^m\| = |j/r - k/2^m|$, pues en caso contrario bastaría trasladar k para cada j . Con esto y dividiendo el sumatorio en los casos $j/r - k/2^m < 0$ y $j/r - k/2^m \geq 0$,

$$p \leq \frac{1}{4r2^{2m}} \sum_{k,j} \left| \frac{j}{r} - \frac{k}{2^m} \right|^{-2} = \sum_j \sum_{k > Mj/r} \frac{\left(\frac{k}{2^m} - \frac{j}{r} \right)^{-2}}{4r2^{2m}} + \sum_j \sum_{k \leq 2^m j/r} \frac{\left(\frac{j}{r} - \frac{k}{2^m} \right)^{-2}}{4r2^{2m}}.$$

Por hipótesis, $|j/r - k/2^m|^{-2} = \|j/r - k/2^m\|^{-2} \leq (\frac{1}{2}r^{-2})^{-2}$ y $\|j/r - k/2^m\|$ varía en $1/2^m$ cada vez que k varía en una unidad, por tanto,

$$p \leq \frac{2r}{4r2^{2m}} \left(\left(\frac{1}{2}r^{-2} \right)^{-2} + \left(\frac{1}{2}r^{-2} + \frac{1}{2^m} \right)^{-2} + \left(\frac{1}{2}r^{-2} + \frac{2}{2^m} \right)^{-2} + \left(\frac{1}{2}r^{-2} + \frac{3}{2^m} \right)^{-2} + \dots \right),$$

donde el factor $2r$ del numerador en la primera fracción proviene de $0 \leq j < r$, así que j toma, en el peor caso, r valores en cada suma. Sacando factor común $(\frac{1}{2}r^{-2})^{-2}$ en lo anterior, como $r < 2^n$ y $m = 2n + 1$ entonces $2r^2 < 2^{2m}$ y queda:

$$p \leq \frac{2r}{4r2^{2m}} \left(\frac{1}{2}r^{-2} \right)^{-2} \sum_{k \geq 0} \frac{1}{(1 + k \frac{2r^2}{2^m})^2} < \frac{2r}{4r2^{2m}} \left(\frac{1}{2}r^{-2} \right)^{-2} \frac{2^m}{2r^2} \frac{1}{1 - \frac{2r^2}{2^m}} = \frac{1}{2^m/r^2 - 2},$$

donde hemos aplicado que si $0 < \delta < 1$, entonces por sumas de Riemann para la partición $\bigcup_{n=0}^{\infty} [(n-1)\delta, n\delta]$, se obtiene que $\sum_{n=0}^{\infty} \frac{\delta}{(1+n\delta)^2} < \int_{-\delta}^{\infty} \frac{dx}{(1+x)^2} = \frac{1}{1-\delta}$. Se concluye que $p < \frac{1}{2^{m/r^2-2}} < \frac{1}{2^{m/2^{2n}-2}}$, así que la probabilidad disminuye exponencialmente con la diferencia $m - 2n$, lo que justifica también la elección de m . Por ejemplo, para $m > 2n + 10$ la cota hallada ya asegura que la probabilidad de fallo es menor que media milésima.

Puesto que la probabilidad de que la medición resulte k_0 y colapse en un estado con sumandos en J_{k_0} es menor que p y esta es despreciable, el colapso se produce en un estado proporcional a $|k_0\rangle \otimes \sum_{j=0}^{r-1} D_{2^m} \left(\frac{j}{r} - \frac{k_0}{2^m} \right) |u_j\rangle$ donde hay una alta probabilidad de que $\left\| \frac{j}{r} - \frac{k_0}{2^m} \right\| < 1/2r^2$.

El valor $k_0/2^m$ es conocido gracias al resultado de la medición. Por lo observado, $\left\| \frac{j}{r} - \frac{k_0}{2^m} \right\|$ es $\left| \frac{j}{r} - \frac{k_0}{2^m} \right|$ o $\left| \frac{j}{r} \pm 1 - \frac{k_0}{2^m} \right|$, así que *aproximan bien* el número $k_0/2^m$, i.e.:

$$\left| \frac{j}{r} - \frac{k_0}{2^m} \right| < \frac{1}{2r^2} \quad \text{o} \quad \left| \frac{j \pm r}{r} - \frac{k_0}{2^m} \right| < \frac{1}{2r^2}.$$

Un resultado de teoría de números establece que si dichas fracciones son irreducibles, esto implica que son fracciones convergentes de $k_0/2^m$. Por el algoritmo de la fracción continua, sabiendo que $r < N$ es denominador de una convergente, podemos calcular posibles valores de r [Apéndice (B.2)]. Nótese que para aplicar dicho resultado es necesario que las fracciones j/r y $j/r \pm 1$ sean irreducibles. Esto no supone una desventaja puesto que si b/c es j/r o $j/r \pm 1$ reducida, entonces b/c es una convergente porque *aproxima bien*: $|b/c - k_0/2^m| < 1/2r^2 < 1/2c^2$. Repitiendo el QPT otras l veces se obtienen c_1, \dots, c_l y r será muy probablemente el mínimo común múltiplo de esos valores –más probable cuantas más repeticiones del algoritmo–.

Con este candidato a r (puede ser el orden o un divisor suyo), por construcción del circuito, basta probar que $a^r \equiv 1 \pmod{N}$ para concluir que es el orden de a . Este y los cálculos del paso 3, se llevan a cabo eficientemente con ordenadores convencionales mediante la exponenciación modular.

Por último, en caso de que r sea par y $a^{r/2} \not\equiv -1 \pmod{N}$, podemos obtener un divisor de N calculando $\gcd(a^{r/2} + 1, N)$, un proceso también asequible por el algoritmo de Euclides.

En este sentido, una pregunta natural es cuan probable es que el algoritmo funcione en una primera elección de a , esto es, sin tener que probar de nuevo. La probabilidad de éxito es mayor del 50% e incrementa con el número de factores primos entre sí de N , como se discute en el Apéndice [(B.1)].

Conclusión: dado $a \in \mathbb{N}$ tal que $2 \leq a < N$, $\gcd(a, N) \neq 1$, la parte cuántica del algoritmo de Shor permite aplicar el algoritmo de la fracción continua para calcular el orden r de a con una alta probabilidad de acierto y este, a su vez, satisface con probabilidad mayor que 1/2 las condiciones necesarias para calcular un divisor no trivial de N .

APÉNDICE A

La función de onda y la esfera de Bloch

A.1. Construcción de la función de onda por el principio de superposición

La onda básica correspondiente a momento lineal p y energía E es

$$(A.1) \quad \varphi(x, t) = e^{i(px - Et/\hbar)},$$

donde $\hbar = 1,05457 \cdot 10^{-34}$ es la *constante de Planck reducida* tal que $h = 2\pi\hbar$ y x y t indican posición y tiempo.

Por otro lado, como se recoge en uno de los postulados iniciales de la mecánica cuántica, función de onda $\Psi(x, t)$ asociada al sistema físico de una partícula codifica toda la información acerca de este (todas las propiedades observables del sistema), que es una función compleja univariante que depende de la posición y el tiempo y cuyo módulo al cuadrado es proporcional a la densidad de probabilidad de encontrar la partícula en cierta región, para un tiempo fijado. La función de onda es además continua, con derivadas continuas (salvo en posibles puntos donde el potencial se hace infinito) y de cuadrado integrable Lebesgue (respecto a la posición).

Tratemos de deducir la forma de la función de onda Ψ de una partícula de masa m sobre la que no actúa ninguna fuerza y que está confinada al intervalo $[0, L]$. Las paredes $x = 0$ y $x = L$ son *pozos de potencial infinito* que la partícula no puede traspasar, de forma que, en realidad, las paredes sí que ejercen una fuerza sobre la partícula para contenerla en dicho intervalo, aunque aquí tomaremos la situación idealizada.

Por el postulado comentado antes, es natural imponer que $\Psi(x, t) = 0$ para todo tiempo t si $x \notin [0, L]$, pues la probabilidad de encontrarla fuera de dicho intervalo es nula. Además, por la continuidad de la función de onda se sigue que para todo tiempo $\Psi(0, t) = \Psi(L, t) = 0$. Como no hay fuerzas debido a un potencial, la energía total coincide con la energía cinética, $E = \frac{1}{2}mv^2$. Puesto que el momento lineal es $p = mv$ tenemos que $E = p^2/2m$. De esta forma, dada una energía E tenemos dos

momentos posibles, $\pm p_E = \pm\sqrt{2mE}$ según el sentido de la velocidad v a derecha o izquierda, respectivamente.

Notar que las ondas como (A.1) no se corresponden con las funciones de onda vistas en la experimentación, puesto que $|\varphi(x, t)|^2$ no es integrable, en contradicción con el postulado enunciado. Son más bien superposiciones de este tipo de ondas las que generan funciones de onda asociadas a las partículas.

La manera habitual de proceder en problemas de mecánica cuántica es resolver el problema obteniendo soluciones particulares para *casos aislados* de los que conocemos ciertos observables y tomar la solución final al problema como combinación lineal de las particulares. En nuestro caso, sabemos que los observables E y p vienen en parejas con la relación $\pm p_E = \pm\sqrt{2mE}$, luego buscamos como soluciones particulares las ondas con estos momentos y energías. Así, por el principio de superposición (aplicado a la mecánica cuántica), la función de onda Ψ será combinación lineal de ondas como en (1.1) para cada pareja de energía y momento lineal:

$$(A.2) \quad Ae^{i(p_E x - Et/\hbar)} + Be^{i(-p_E x - Et/\hbar)}, \quad A, B \in \mathbb{C}.$$

Es claro que estas ondas deben valer 0 en los extremos del intervalo, para todo t . Sustituyendo $x = 0$ y $x = L$, se obtiene el sistema:

$$(A.3) \quad \left\{ \begin{array}{l} e^{i(-Et/\hbar)} A + e^{i(-Et/\hbar)} B = 0, \quad \forall t, \\ e^{i(p_E L - Et/\hbar)} A + e^{i(-p_E L - Et/\hbar)} B = 0, \quad \forall t. \end{array} \right\} \Rightarrow \begin{bmatrix} 1 & 1 \\ e^{i(p_E L/\hbar)} & e^{i(-p_E L/\hbar)} \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

que tiene solución no trivial si y solo si el determinante de la matriz es nulo, esto es:

$$(A.4) \quad e^{i(p_E L/\hbar)} - e^{i(-p_E L/\hbar)} = 2i \operatorname{sen}(p_E L/\hbar) = 0.$$

De esta forma, se obtiene un conjunto numerable y discreto de posibles momentos y energías¹, pues el seno se anula en los múltiplos enteros de π y para cada p_E tenemos el de sentido contrario $-p_E$ y la energía E . En conclusión e introduciendo una notación más conveniente,

$$(A.5) \quad A_n = -B_n \in \mathbb{C}, \quad p_n = \frac{n\pi\hbar}{L} \quad \text{y} \quad E_n = \frac{(n\pi\hbar)^2}{2mL^2}, \quad n \in \mathbb{N}.$$

Sabiendo esto, reescribimos (A.2), denotando cada onda con su correspondiente momento lineal y energía:

$$(A.6) \quad \Psi_n(x, t) = A_n e^{i(-E_n t/\hbar)} (e^{i(p_n x/\hbar)} - e^{i(-p_n x/\hbar)}) = A_n e^{i(-E_n t/\hbar)} 2i \operatorname{sen}(p_n x/\hbar).$$

Fijado t , tratemos de “normalizar” dichas ondas y tomarlas con coeficientes reales para obtener una expresión de $\Psi_n = C e^{i(-E_n t/\hbar)} \operatorname{sen}(\dots)$. En el capítulo 2 queda justificada la elección de coeficientes reales sin pérdida de generalidad, pues multiplicar un estado —una función de onda— por una constante de módulo 1 no cambia el estado. Por lo comentado en el postulado, $|\Psi(x, t)|^2$ es proporcional a la densidad de probabilidad de que la partícula se encuentre en la posición x en el momento t , luego $\int_R |\Psi(x, t)|^2$ es proporcional a la probabilidad de que la partícula

¹Aquí queda justificado matemáticamente la *cuantización* observada por Einstein en su experimentación.

esté en la región R para tiempo t . Para que sea una probabilidad, queremos que $\int_{\mathbb{R}} |\Psi_n(x, t)|^2 dx = \int_0^L |\Psi_n(x, t)|^2 dx = 1$, para todo n . Por (A.5) y (A.6),

$$\begin{aligned} \int_0^L |\Psi_n(x, t)|^2 dx &= \int_0^L |A_n e^{-\frac{iE_n t}{\hbar}} 2i|^2 \sin^2(p_n x / \hbar) dx = \\ &= \int_0^L |A_n|^2 4 \sin^2(p_n x / \hbar) dx = 2|A_n|^2 \int_0^L 1 - \cos\left(\frac{\pi 2nx}{L}\right) dx = 2L|A_n|^2, \end{aligned}$$

luego $\int_0^L |\Psi_n(x, t)|^2 dx = 1$ si y solo si $|A_n|^2 = \frac{1}{2L}$. Como buscamos coeficientes reales, esto ocurre si $A_n = \pm \frac{1}{i\sqrt{2L}}$. Sin pérdida de generalidad, tomamos los coeficientes positivos para concluir que

$$(A.7) \quad \Psi_n(x, t) = \frac{2i}{i\sqrt{2L}} e^{i(-E_n t / \hbar)} \sin(p_n x / \hbar) = \sqrt{\frac{2}{L}} e^{i(-E_n t / \hbar)} \sin\left(\frac{\pi n x}{L}\right), \quad n \in \mathbb{N}.$$

Es claro que, para cada t fijo, Ψ_n no está concentrada en un punto del espacio, como se esperarí de las partículas clásicas, luego deberíamos superponer muchas de las ondas para obtener una con una forma de “pico estrecho”, una función de onda Ψ concentrada en el espacio: $\Psi(x, t) = \sum c_n \Psi_n(x, t)$, $c_n \in \mathbb{C}$.

A.2. La esfera de Bloch

Consideramos $S^2 = \{x \in \mathbb{R}^3 : \|x\|^2 = 1\}$ y $Q = \{\text{qubits}\} / \sim$, donde \sim es la relación de equivalencia que identifica dos qubits si difieren en multiplicar por una constante de módulo 1. Sean $\theta \in [0, \pi]$, $\varphi \in [0, 2\pi)$, tenemos una biyección entre S^2 y Q :

$$(\cos \varphi \sin \theta, \sin \varphi \sin \theta, \cos \theta) \longmapsto \cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle.$$

Para la inyectividad usamos la ortonormalidad de $|0\rangle$ y $|1\rangle$ para notar que si $\cos \frac{\theta_1}{2} = \cos \frac{\theta_2}{2}$ entonces $\theta_1 = \theta_2$. Si $\theta_1 = \theta_2 = 0$, ambas preimágenes son $(0, 0, 1)$, mientras que si no es 0, $e^{i\varphi_1} \sin \frac{\theta_1}{2} = e^{i\varphi_2} \sin \frac{\theta_2}{2} \Rightarrow e^{i\varphi_1} = e^{i\varphi_2} \Rightarrow \varphi_1 = \varphi_2$. Para la sobreyectividad, dado $r e^{iA} |0\rangle + s e^{iB} |1\rangle$ con $r^2 + s^2 = 1$, tomamos $\theta \in [0, \pi]$ tal que $\cos \frac{\theta}{2} = r$, luego también $\sin \frac{\theta}{2} = s$. Este qubit en Q se identifica con $r e^{-iA} e^{iA} |0\rangle + s e^{-iA} e^{iB} |1\rangle = \cos \frac{\theta}{2} |0\rangle + \sin \frac{\theta}{2} e^{i\varphi} |1\rangle$, donde $\varphi = A - B$ (módulo 2π).

APÉNDICE B

Detalles del algoritmo de Shor

B.1. Probabilidad de éxito al primer intento

El algoritmo de Shor devuelve un factor no trivial de N para cada entero a elegido al azar, según su orden r . Hay dos situaciones en las que el algoritmo no funciona: si r es impar o si r es par y $a^{r/2} \equiv -1 \pmod{N}$. Cabe por tanto preguntarse si esta situación es muy probable o no.

Teorema. Supongamos que N factoriza en k factores coprimos, esto es, $N = \prod_{i=1}^k p_i^{\alpha_i}$, con p_i primos distintos y $\alpha_i \in \mathbb{Z}_{\geq 0}$. Sea $a \in [2, N)$ un entero donde r el orden de a en $(\mathbb{Z}/N\mathbb{Z})^*$. Sea $p \equiv \text{prob}(\{r \text{ impar}\} \cup \{r \text{ par y } a^{r/2} \equiv -1 \pmod{N}\})$. Entonces $p \leq \frac{1}{2^{k-1}}$.

Demostración:

Sea $a \in [2, N)$ un entero y r su orden en $(\mathbb{Z}/N\mathbb{Z})^*$. Por el Teorema Chino del Resto (TCR),

$$a^r \equiv 1 \pmod{N} \iff a^r \equiv 1 \pmod{p_i^{\alpha_i}}, \quad 1 \leq i \leq k,$$

luego $r_i | r$ para cada i , donde r_i es el orden de a en $(\mathbb{Z}/p_i^{\alpha_i}\mathbb{Z})^*$.

Sea m el mínimo común múltiplo de los r_i , tenemos por lo anterior que $m | r$ y también que $a^r \equiv 1 \pmod{p_i^{\alpha_i}}$, $1 \leq i \leq k$. De nuevo por el TCR, $a^m \equiv 1 \pmod{N}$, luego $r | m$ por ser r el orden de a . Se concluye que $r = m$.

Sea 2^d la mayor potencia de 2 que divide a r y sea 2^{d_i} la mayor potencia de 2 que divide a r_i , para cada $1 \leq i \leq k$.

Si r es impar, no puede tener divisores pares y por tanto r_i también es impar, de lo que se obtiene $d = d_1 = \dots = d_k = 0$.

Si r es par y $a^{r/2} \equiv -1 \pmod{N}$ entonces por el TCR $a^{r_i/2} \equiv -1 \pmod{p_i^{\alpha_i}}$, que implica que $r_i \nmid (r/2)$. Equivalentemente, para cada i se tiene que la mayor potencia de 2 que dividen a r_i es también la mayor que divide a r : $d = d_1 = \dots = d_k$.

Así, basta calcular la probabilidad de que las mayores potencias de 2 coincidan.

Por el TCR, elegir a con estas características es equivalente a elegir a_1, \dots, a_k tales que $a \equiv a_i \pmod{p_i^{\alpha_i}}$ cuyos órdenes r_i tengan 2^d como mayor potencia de 2 que lo divide.

El grupo multiplicativo $(\mathbb{Z}/p_i^{\alpha_i}\mathbb{Z})^*$ es cíclico de orden $\phi(p_i^{\alpha_i})$, de forma que si g es un generador, el orden de g^n es

$$(B.1) \quad \frac{\phi(p_i^{\alpha_i})}{\gcd(\phi(p_i^{\alpha_i}), n)}, \quad \text{donde } \gcd \text{ denota el máximo común divisor.}$$

En particular, como el orden de un elemento es divisor de $\phi(p_i^{\alpha_i})$, la mayor potencia de 2 que divide a un orden es la mayor potencia de 2 que divide a $\phi(p_i^{\alpha_i})$. De (B.1) deducimos que los órdenes que cumplen lo anterior se corresponden con n impar. Clasificando los elementos en función de la mayor potencia de 2 que divide su orden, notamos que el conjunto es $\{g^n : n \text{ impar}\}$ tiene $\phi(p_i^{\alpha_i})/2$ elementos, por lo que es el mayor conjunto con esta clasificación.

Así, fijado 2^d , la probabilidad p_i de que un elemento en $(\mathbb{Z}/p_i^{\alpha_i}\mathbb{Z})^*$ tenga un orden cuya mayor potencia de 2 sea 2^d satisfice, por la regla de Laplace, que

$$(B.2) \quad p_i \leq \frac{\phi(p_i^{\alpha_i})/2}{\phi(p_i^{\alpha_i})} = \frac{1}{2}.$$

Fijando la potencia 2^d con el primer elemento a_1 , la probabilidad de que a_1, \dots, a_k cumplan la propiedad buscada es la probabilidad de que los órdenes r_2, \dots, r_k tengan 2^d como mayor potencia de 2 que los divide. Por (B.2), esto es $p \leq \frac{1}{2^{k-1}}$. \square

Se concluye que la probabilidad de que el algoritmo funcione para un entero $2 \leq a < N$ es $p \geq 1 - \frac{1}{2^{k-1}}$. En particular, si N factoriza en dos primos distintos, la probabilidad es mayor que el 50%.

B.2. Algoritmo de la fracción continua y aproximación de números racionales

La fracción continua.

Sea $p/q \in \mathbb{Q}$ donde suponemos sin pérdida de generalidad que $q > 0$. Por el algoritmo de Euclides para calcular $\gcd(p, q)$, existen $c_0 \in \mathbb{Z}$, $c_1, \dots, c_k \in \mathbb{Z}_{\geq 0}$ y $r_0, \dots, r_{k+2} \in \mathbb{Z}_{\geq 0}$ tales que:

$$r_0 = p; \quad r_1 = q; \quad r_{k+1} = \gcd(p, q); \quad r_{k+2} = 0,$$

$$\text{para cada } 1 \leq n \leq k+1, \quad r_{n-1} = c_{n-1}r_n + r_{n+1}, \quad 0 \leq r_{n+1} < r_n.$$

Matricialmente, esto es:

$$(B.3) \quad \begin{pmatrix} r_{n-1} \\ r_n \end{pmatrix} = \begin{pmatrix} c_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_n \\ r_{n+1} \end{pmatrix}, \quad \text{para } 1 \leq n \leq k+1, \text{ de lo que se deduce que}$$

$$\begin{pmatrix} p \\ q \end{pmatrix} = A_k \begin{pmatrix} r_{k+1} \\ 0 \end{pmatrix}, \quad \text{donde } A_k = \begin{pmatrix} c_0 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} c_k & 1 \\ 1 & 0 \end{pmatrix}, \quad \det(A_k) = (-1)^{k+1}.$$

Si $0 \leq n < k$, $\frac{r_{n-1}}{r_n} = c_{n-1} + \frac{r_{n+1}}{r_n} = c_{n-1} + \frac{1}{r_n/r_{n+1}}$ implica, inductivamente, que

$$(B.4) \quad \frac{p}{q} = c_0 + \frac{1}{q/r_2} = c_0 + \frac{1}{c_1 + \frac{1}{r_2/r_3}} = \dots = c_0 + \frac{1}{c_1 + \frac{1}{\dots + \frac{1}{c_{k-1} + \frac{1}{c_k}}}}$$

Esta última expresión es la *fracción continua* de p/q y se denota $[c_0, \dots, c_{k-1}]$. Para todo número racional existe una fracción continua (finita).

Nótese que si $p/q < 0$, es decir, $p < 0$ entonces $c_0 \leq 0$ y como $q = r_1 > r_2 \geq 0$ se sigue que $c_1 \geq 1$. En general, $r_n > r_{n-1} \geq 0$ y por tanto $c_n \geq 1$ si $n \geq 1$.

Se puede generalizar la fracción continua a $[c_0, \dots, c_k, \lambda]$ para $\lambda \in \mathbb{R}$ definiendo

$$[c_0, \dots, c_k, \lambda] = c_0 + \frac{1}{c_1 + \frac{1}{\dots + \frac{1}{c_k + \frac{1}{\lambda}}}}$$

En particular, si $\lambda \in \mathbb{Q}$ tal que $\lambda = [a_0, \dots, a_j]$

$$(B.5) \quad \text{es su fracción continua, entonces } [c_0, \dots, c_k, \lambda] = [c_0, \dots, c_k, a_0, \dots, a_j].$$

El algoritmo de la fracción continua.

El algoritmo de la fracción continua permite calcular p y q dada su fracción continua $[c_0, \dots, c_k]$, cuando $\text{gcd}(p, q) = 1$. Es el equivalente a calcular $A_k \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

El algoritmo se basa en la regla *multiplicar por el de abajo y sumar el anterior*, es decir, en la columna j , el valor en fila 1 (resp.2) se calcula multiplicando los valores de las filas 0 y 1 (resp.0 y 2) de la columna $j - 1$ y sumando el valor de la fila 1 (resp.2) de la columna $j - 2$, donde las columnas 1 y 2 y la fila 0 son conocidas:

(B.6)	0	c_0	c_1	c_{k-1}	c_k	p/q	$p_0 = c_0; \quad q_0 = 1;$
	0	1	p_0	p_{k-2}	p_{k-1}	$p_k = p$	$q_1 = c_1;$
	1	0	q_0	q_{k-2}	q_{k-1}	$q_k = q$	$\frac{p_n}{q_n} = \frac{c_n p_{n-1} + p_{n-2}}{c_n q_{n-1} + q_{n-2}}, \text{ si } n \geq 1.$

Por recursividad, como $p_0, q_0, p_1, q_1 \in \mathbb{Z}$, se tiene que $p_n/q_n \in \mathbb{Q}$. Las fracciones p_n/q_n se llaman *convergentes*. Nótese que $q_n \geq 0$ por la recursión que las define y porque $c_n \geq 0$ si $n \geq 1$.

Además, para cada $0 \leq n \leq k$ tenemos $p_n/q_n = [c_0, \dots, c_n]$, así que por (B.3),

$$(B.7) \quad A_n = \begin{pmatrix} c_0 & 1 \\ 1 & 0 \end{pmatrix} \dots \begin{pmatrix} c_n & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} p_1 & p_0 \\ q_1 & q_0 \end{pmatrix} \dots \begin{pmatrix} c_n & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix}$$

y como $\det(A_n) = (-1)^{n+1}$, se deduce que $p_n q_{n-1} - q_n p_{n-1} = (-1)^{n+1}$.

Como ocurría en la fracción continua, el algoritmo puede generalizarse a expresiones $[c_0, \dots, c_k, \lambda]$ donde $\lambda \in \mathbb{R}$. En particular, si $\lambda \in \mathbb{Q}$ con $\lambda = [a_0, \dots, a_j]$

entonces por (B.5) $x = \frac{x_1}{x_2} = [c_0, \dots, c_k, \lambda]$ puede calcularse mediante el algoritmo y se simplifica a:

$$(B.8) \quad \begin{array}{c|c|c|c|c|c|c|c} 0 & c_0 & c_1 & & c_k & a_0 & & a_j & & x \\ 0 & 1 & p_0 & \cdots & p_{k-1} & p_k & \cdots & p_{k+j-1} & & x_1 \\ 1 & 0 & q_0 & & q_{k-1} & q_k & & q_{k+j-1} & & x_2 \end{array} \quad \text{donde} \quad x = \frac{\lambda p_k + p_{k-1}}{\lambda q_k + q_{k-1}}.$$

Por otro lado, $q_n > 2^{\frac{n-1}{2}}$. En efecto, para $n = 0$ se tiene que $q_0 = 1 > 2^{-1/2}$. Por inducción, si suponemos cierto para $0 \leq j < n$ entonces, como $c_n \geq 1$, obtenemos que $q_n = c_n q_{n-1} + q_{n-2} \geq q_{n-1} + q_{n-2} > 2^{\frac{n-2}{2}} + 2^{\frac{n-3}{2}} = (2^{\frac{1}{2}} + 1)2^{\frac{n-3}{2}} > 2 \cdot 2^{\frac{n-3}{2}}$.

Es decir, los denominadores de las convergentes tienen un crecimiento exponencial y por tanto hay una cantidad del orden de $\log N$ que satisfacen $q_n < N$.

Se concluye con esto la “equivalencia” entre los números racionales y las fracciones continuas. Por el algoritmo de Euclides, todo racional puede expresarse como una fracción continua finita, como se desarrolla en (B.4). Recíprocamente, por el algoritmo de la fracción continua descrito en (B.6), toda fracción continua se identifica con un número racional¹.

Buena aproximación

Dado $\alpha \in \mathbb{Q}$ se dice que $p/q \in \mathbb{Q}$ aproxima bien α si $|\alpha - \frac{p}{q}| < \frac{1}{2q^2}$.

Teorema. Sea $\alpha \in \mathbb{Q}$ un número racional y sea $p/q \in \mathbb{Q}$ una fracción irreducible que aproxima bien α , entonces p/q es una convergente de α .

Se deduce que las fracciones convergentes de un número son las únicas que aproximan bien dicho número.

Demostración:

Como $\alpha \in \mathbb{Q}$, α es en realidad la última convergente de su fracción continua. Así, si $\alpha = p/q$ ya hemos terminado.

Supongamos que $\alpha \neq p/q$. Desarrollamos $p/q = [c_0, \dots, c_k]$ como en (B.4). Como $\gcd(p, q) = 1$ por hipótesis, aplicando el algoritmo de la fracción continua (B.6) a $[c_0, \dots, c_k]$ obtenemos que $p_k/q_k = p/q$.

Como p/q aproxima bien α , existe δ necesariamente racional tal que:

$$0 < |\delta| < 1 \quad \text{y} \quad \alpha = \frac{p_k}{q_k} + \frac{\delta}{2q_k^2}. \quad \text{Definimos} \quad \lambda = 2 \frac{(-1)^k}{\delta} - \frac{q_{k-1}}{q_k} \in \mathbb{Q}.$$

$$\text{Aplicando (B.7),} \quad \lambda p_k + p_{k-1} = 2 \frac{(-1)^k p_k}{\delta} + \frac{-p_k q_{k-1}}{q_k} + p_{k-1} = 2 \frac{(-1)^k p_k}{\delta} + \frac{(-1)^k}{q_k},$$

$$\lambda q_k + q_{k-1} = 2 \frac{(-1)^k q_k}{\delta} + \frac{-q_k q_{k-1}}{q_k} + q_{k-1} = 2 \frac{(-1)^k q_k}{\delta}, \quad \text{luego} \quad \alpha = \frac{\lambda p_k + p_{k-1}}{\lambda q_k + q_{k-1}}.$$

Esta última expresión implica, por (B.8), que $\alpha = [c_0, \dots, c_k, \lambda]$, donde p_k/q_k es una convergente, y hemos concluido. \square

¹Un número racional puede considerarse sin pérdida de generalidad como una fracción irreducible p/q .

Aplicación a la estimación de fase en el algoritmo de Shor.

Dado un número racional α con el algoritmo de la fracción continua pueden calcularse las fracciones convergentes de dicho número —un ordenador convencional realiza la tarea de forma eficiente—. A partir de ello, determinar las convergentes cuyos denominadores son menores que un natural N fijado es fácil y hay del orden de $\log N$ que lo satisfacen. Además, si una fracción irreducible *aproxima bien* α entonces será una de las convergentes calculadas.

Bibliografía

- [1] A. GALINDO AND P. PASCUAL. *Mecánica cuántica*. Alhambra, Madrid, 1978.
- [2] E. HERNÁNDEZ RODRÍGUEZ, M. J. VÁZQUEZ GALLO, AND M. A. ZURRO MORO. *Álgebra Lineal y Geometría*. Pearson, Madrid, 2012, 3^a Ed.
- [3] F. CHAMIZO. *Un poco de física cuántica para chicos listos de primero (del grado de física o matemáticas)*.
<http://matematicas.uam.es/~fernando.chamizo/physics/files/qf.pdf>, 2015.
- [4] K. H. ROSEN. *Elementary number theory and its applications*. Addison-Wesley Publishing Company, Advanced Book Program, Reading, MA, third edition, 1993.
- [5] M. A. NIELSEN AND I. L. CHUANG. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2000.
- [6] M. NAKAHARA AND T. OHMI. *Quantum Computing. From linear algebra to physical realizations*. CRC Press, Boca Raton, FL, 2008.
- [7] S. J. MILLER AND R. TAKLOO-BIGHASH. *An invitation to modern number theory*. Princeton University Press, Princeton, NJ, 2006. With a foreword by P. Sarnak.

