Computación Cuántica

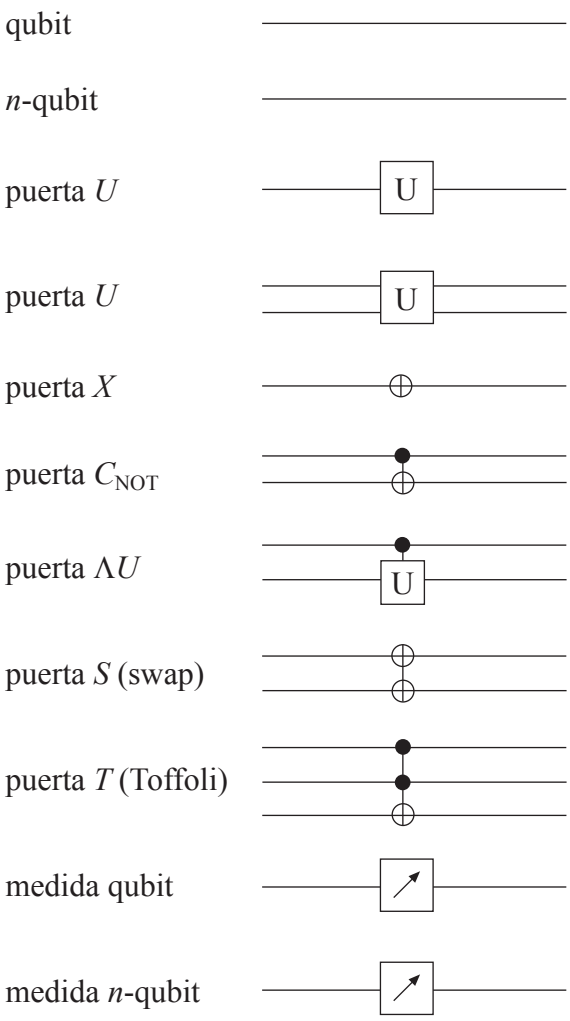
Jesús García López de Lacalle

Autores: Alfonsa García López
Jesús García López de Lacalle
Francisco García Mazarío

Máster en Ciencias y Tecnologías de la Computación
ETS de Ingeniería de Sistemas Informáticos
Universidad Politécnica de Madrid
2016

Índice de contenidos

1 Notación


qubit
n-qubit
puerta *U*
U
puerta *U*
U
puerta *X*
puerta C_{NOT}
puerta ΛU
U
puerta *S* (swap)
puerta *T* (Toffoli)
medida qubit
medida *n*-qubit

2 Introducción

La computación cuántica empezó a desarrollarse en la década de los ochenta a raíz de las propuestas de Paul Benioff, David Deutsch y Richard Feynman. En 1982, Benioff [2] y Feynman [16] sugirieron independientemente que, dado el elevado coste computacional del cálculo de la evolución de sistemas cuánticos, la evolución de estos sistemas se podría considerar como una herramienta de cálculo más que como un objeto a calcular. Poco después, en 1985, y también de forma independiente Deutsch [12] propone la búsqueda de un ordenador que sea capaz de simular eficientemente un sistema físico arbitrario. La conjunción de todas estas ideas ha conducido a la concepción actual de ordenador cuántico.

Cuestionar el sistema de computación clásico, que cuenta con una sólida base teórica y con el aval de infinidad de aplicaciones en todos los ámbitos de la vida cotidiana, sólo tiene sentido si el modelo que se propone como alternativo es potencialmente mejor que el actual. Efectivamente así lo hacen Benioff, Deutsch y Feynman, fundamentando sus propuestas sobre la posibilidad de que los sistemas cuánticos tengan mayor potencia de cálculo que los clásicos. El argumento que todos utilizan para apuntar esta posibilidad es el hecho de que la simulación de un ordenador cuántico (sistema cuántico) en un ordenador clásico requiere una gran cantidad de operaciones.

El principal método para aumentar la capacidad de cálculo de un ordenador clásico es el procesamiento en paralelo. Los ordenadores que soportan este esquema de programación disponen de varios cientos o miles de procesadores. Sabemos que la capacidad de almacenamiento de información y la capacidad de cálculo de un ordenador son proporcionales al número de celdas de memoria y al número de procesadores respectivamente, es decir, al tamaño del ordenador. Entonces la capacidad de un ordenador clásico (de almacenamiento y de cálculo) crece linealmente con respecto a su tamaño.

En un ordenador cuántico la situación cambia por completo, hasta el punto de que su capacidad crece exponencialmente con respecto a su tamaño. Este hecho, estrechamente relacionado con el principio de superposición de la mecánica cuántica, se denomina paralelismo cuántico. Llamamos qubits o bits cuánticos a los sistemas cuánticos elementales, es decir, a los sistemas cuánticos obtenidos a partir de dos estados. Los sistemas cuánticos de n qubits se describen mediante vectores de un espacio de Hilbert complejo de dimensión 2^n. Esto permite codificar una cantidad exponencial de información en el estado de un sistema cuántico de n qubits. Además, cualquier transformación del estado del sistema se traduce en la modificación simultánea de toda la información almacenada. Por tanto, la capacidad de un ordenador cuántico (tanto de almacenamiento como de cálculo) crece exponencialmente con respecto a su tamaño.

Sin embargo, la medición de estados cuánticos es un inconveniente importante para la computación cuántica. Hay que recordar que las medidas cuánticas no son deterministas. Esto quiere decir, por ejemplo, que si medimos dos estados iguales los resultados no tienen por qué ser iguales. El proceso de medida es, por tanto, un experimento aleatorio en el que la probabilidad de cada resultado está determinada por el estado del sistema.

Las dificultades para sacar provecho del paralelismo cuántico son tan notables que hubo que esperar más de una década para encontrar el primer gran resultado. En 1994 Peter

W. Shor sorprendió a todos presentando sendos algoritmos polinomiales para factorizar números enteros y para calcular logaritmos discretos [28]. Fueron los primeros problemas relevantes en los que se alcanzaba una aceleración exponencial con respecto a los mejores algoritmos clásicos conocidos. A raíz de este descubrimiento se generó una gran actividad, tanto en el desarrollo de la tecnología necesaria para la construcción de ordenadores cuánticos como en el estudio de algoritmos cuánticos.

El algoritmo de Shor rompió teóricamente el sistema criptográfico más difundido en la actualidad, el sistema RSA propuesto por Rivest, Shamir y Adleman en 1978 [26]. Este hecho contribuyó a su vez al desarrollo de los sistemas criptográficos cuánticos [5, 4, 6]. Las técnicas que se utilizan para garantizar la confidencialidad de los canales cuánticos se apoyan en una propiedad característica de la mecánica cuántica: los estados cuánticos no se pueden copiar (clonar) [33]. En el área de las comunicaciones, además del estudio de la confidencialidad, se están investigando otros problemas como, por ejemplo, la codificación de información clásica en canales cuánticos y el teletransporte de estados cuánticos.

Pero el estudio de este modelo de computación apenas si ha comenzado. Hasta el momento, sólo se han desarrollado ordenadores cuánticos basados en resonancia magnética nuclear [11, 18], en trampas de iones [10] y en cavidades cuánticas [13]. Con la primera de estas técnicas se han conseguido prototipos de hasta 10 qubits, sobre los que se ha probado el algoritmo de Shor. También se ha propuesto la construcción de ordenadores cuánticos aprovechando los conocimientos actuales sobre semiconductores [22, 32], aunque esta técnica está menos desarrollada.

Desde el punto de vista algorítmico, sólo se ha podido hacer efectiva una ganancia exponencial en el cálculo de transformadas de Fourier y, en estos momentos, ésta es la herramienta más importante de la computación cuántica. Otra técnica que permite mejorar la complejidad de algunos algoritmos clásicos, aunque con ganancia solamente cuadrática, es el método de Grover de búsqueda en conjuntos desordenados [20].

Los ordenadores cuánticos, a diferencia de los clásicos, son dispositivos analógicos. Este hecho plantea mayores dificultades para la construcción de ordenadores cuánticos que las que se tuvieron que afrontar para ordenadores clásicos. En primer lugar, los estados cuánticos se modifican por la influencia del entorno. Esto provoca el fenómeno denominado decoherencia que se convierte en una fuente de errores. Y, en segundo lugar, la imprecisión del propio ordenador cuántico al aplicar el algoritmo constituye una nueva fuente de errores. Estas son las razones fundamentales por las que la computación cuántica requiere un mecanismo para acotar la acumulación de errores durante el proceso de cálculo. Para ello hay que superar dificultades importantes: los errores cuánticos son continuos, no se pueden copiar estados y, hasta el final, no se puede leer (medir) la información codificada en un estado cuántico. Estos obstáculos fueron finalmente superados, una vez que Shor [29] y Steane [30] establecieran las ideas básicas para la construcción de códigos cuánticos y se formalizase de forma consistente la teoría cuántica de códigos [8, 19, 25].

3 Modelo de computación cuántica

En lo que sigue presentamos las nociones básicas del modelo de computación cuántica. Para estudiarlo con más detalle, se puede consultar [24] ó [21].

3.1 Representación de la información

En computación cuántica la unidad elemental de información es el *qubit* o bit cuántico, que se define a partir de dos estados básicos denotados por $|0\rangle$ y $|1\rangle$. Físicamente se representa por un sistema cuántico de dos estados, por ejemplo el spin de un electrón. En este sistema se puede representar el spin $-\frac{1}{2}$ por el estado $|0\rangle$ y el spin $\frac{1}{2}$ por el estado $|1\rangle$.

Pero, además, el estado cuántico puede ser una *superposición* de los dos estados básicos, que modelizaremos por una combinación lineal. Ésta es la principal diferencia con el modelo de computación clásico.

Actualmente en los canales cuánticos de comunicación se utiliza la polarización de un fotón como soporte físico de un qubit. Un estado de polarización puede ser modelizado por un vector unidad apuntando en la dirección adecuada y se puede representar como un vector de norma uno en un espacio de Hilbert complejo.

Así, en términos matemáticos, el estado cuántico de un qubit es una combinación lineal de los dos estados básicos $\phi = a|0\rangle + b|1\rangle$ tal que $|a|^2 + |b|^2 = 1$. Es decir ϕ es un vector unitario de un espacio de Hilbert complejo $\mathcal{H}$, en el que $B_1 = [|0\rangle, |1\rangle]$ es una base ortonormal.

Cualquier sistema de medida de estados cuánticos tiene asociada una base ortonormal, respecto de la cual se realiza la medición. Por ejemplo, usando la base B_1 se consideran las polarizaciones horizontal y vertical. Al medir un estado cuántico, éste se proyecta sobre uno de los vectores de la base ortonormal considerada. De este modo, al medir $\phi = a|0\rangle + b|1\rangle$, se obtendrá $\frac{a}{|a|}|0\rangle$ con probabilidad $|a|^2$ o $\frac{b}{|b|}|1\rangle$ con probabilidad $|b|^2$.

Cuando se mide un estado, éste cambia irreversiblemente, salvo los estados de la propia base respecto de la que se mide (o múltiplos de ellos).

En la tabla siguiente se resume el proceso de medida de un qubit, respecto a la base B_1. Decimos que el resultado de la medida es 0 (1) si el estado se proyecta sobre el subespacio generado por $|0\rangle$ ($|1\rangle$).

Estado	Medida	Estado final	Probabilidad
$a\|0\rangle + b\|1\rangle$	0	$\frac{a}{\|a\|}\|0\rangle$	$p_0 = \|a\|^2$
$a\|0\rangle + b\|1\rangle$	1	$\frac{b}{\|b\|}\|1\rangle$	$p_1 = \|b\|^2$

Aparatos o sistemas de medidas diferentes tienen asociadas bases distintas y las mediciones correspondientes dan lugar a resultados diferentes. Por ejemplo, en el espacio $\mathcal{H}$ podemos considerar la base ortogonal $B_\times = [|+\rangle, |-\rangle]$, donde

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

(que corresponde a las polarizaciones 45^o y -45^o).

Para obtener el resultado de la medición de un estado respecto de $B_\times$, basta conocer las coordenadas del vector respecto de esta base. Dado que se puede escribir

$$|0\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle) \quad \text{y} \quad |1\rangle = \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle),$$

si se mide el estado $|0\rangle$ respecto de la base $B_\times$ se obtendrá $|+\rangle$ con probabilidad $1/2$, o $|-\rangle$ con probabilidad $1/2$ y análogamente para el estado $|1\rangle$. Nótese que si hubiéramos medido uno de estos estados usando B_1 se hubiera obtenido el propio estado con probabilidad 1.

La información que contiene un qubit es evidentemente muy pequeña. Para poder representar más información es necesario recurrir a estados de n-qubits.

Un *estado cuántico* de n-qubits es un vector de norma 1 del espacio de Hilbert complejo $\mathcal{H}_n = \mathcal{H} \overset{(n)}{\otimes \cdots \otimes} \mathcal{H}$, de dimensión 2^n, cuya base es $B_n = [|x_0\rangle, \cdots, |x_{2^n-1}\rangle]$, donde cada $x_j \in \{0,1\}^n$ es la representación binaria (con n bits) del número $j \in \{0, \ldots, 2^n - 1\}$. Es decir, un estado cuántico de n-qubits se puede escribir de la forma

$$\phi = \sum_{j=0}^{2^n-1} a_j |x_j\rangle, \quad \text{con} \quad \sum_{j=0}^{2^n-1} |a_j|^2 = 1.$$

Los vectores de B_n son todos los productos tensoriales de n vectores de B_1. Su identificación con cadenas de n bits resulta muy conveniente para codificar la información.

Generalmente, cuando no haya lugar a confusión sobre la dimensión, escribiremos $|j\rangle$ en lugar de $|x_j\rangle$, de este modo se puede poner $B_n = [|0\rangle, \cdots, |2^n - 1\rangle]$ y

$$\phi = \sum_{j=0}^{2^n-1} a_j |j\rangle.$$

Conviene resaltar que la dimensión exponencial de $\mathcal{H}_n$ es la clave de la potencia de la computación cuántica que se basa en el llamado paralelismo cuántico, derivado del hecho de que manejar un estado cuántico, superposición de todos los estados de la base, es como operar simultáneamente con todas las 2^n cadenas de n bits. Esto permite un incremento exponencial de la velocidad de cálculo.

3.2 Estados producto y estados entrelazados

Para entender un poco cómo son los estados cuánticos, veamos un ejemplo sencillo con $n = 2$. Un 2-qubit es un vector unitario del espacio vectorial $\mathcal{H}_2$. En este caso, $B_2 = [|00\rangle, |01\rangle, |10\rangle, |11\rangle]$ es una base ortonormal de $\mathcal{H}_2$, cuyos elementos son todos los productos tensoriales de los elementos de la base $[|0\rangle, |1\rangle]$ de $\mathcal{H}$. (Podemos escribir $B_2 = [|0\rangle, |1\rangle, |2\rangle, |3\rangle]$.)

En $\mathcal{H}_2$ hay estados que son producto tensorial de dos estados de $\mathcal{H}$, como por ejemplo

$$\frac{1}{4}|00\rangle - \frac{\sqrt{3}}{4}|01\rangle + \frac{\sqrt{3}}{4}|10\rangle - \frac{3}{4}|11\rangle = \left(\frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle\right) \otimes \left(\frac{1}{2}|0\rangle - \frac{\sqrt{3}}{2}|1\rangle\right).$$

Pero también hay otros estados que no se pueden poner como producto tensorial de estados de $\mathcal{H}$, por ejemplo

$$\left(\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle\right).$$

Estos estados se denominan *entrelazados* (entangled) y tienen gran importancia en algorítmica y criptografía cuántica.

En un 2-qubit se puede medir el primer qubit o el segundo, usando una base ortonormal de $\mathcal{H}$. Por ejemplo, sea $\phi = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$, con $|a|^2 + |b|^2 + |c|^2 + |d|^2 = 1$. Si se mide el primer qubit usando la base B_1, el resultado de la medición puede ser 0, con probabilidad $|a|^2 + |b^2|$, y el estado resultante tras la medida será

$$\frac{a}{\sqrt{|a|^2 + |b|^2}}|00\rangle + \frac{b}{\sqrt{|a^2 + |b|^2}}|01\rangle,$$

ó bien 1, con probabilidad $|c|^2 + |d^2|$, y el estado se proyectará en

$$\frac{c}{\sqrt{|c|^2 + |d|^2}}|10\rangle + \frac{d}{\sqrt{|c^2 + |d|^2}}|11\rangle.$$

Veamos qué ocurre si se mide el primer qubit de un estado entrelazado, por ejemplo el estado

$$\left(\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle\right).$$

Si el resultado de la medición es 0, lo que ocurre con probabilidad 1/2, el estado resultante sería $|00\rangle$; si el resultado de la medición es 1, el estado resultante sería $|11\rangle$. En cualquier caso, si después se mide el segundo qubit, se obtendrá con probabilidad 1 el mismo resultado de la primera medición. Es decir el resultado de la medida del segundo qubit está condicionado por el de la primera medición. Esta es una característica de los estados entrelazados, que da lugar a la famosa paradoja de Einstein, Podolsky y Rosen (EPR). Los dos qubits de un estado cuántico entrelazado pueden representar sistemas físicos que estén en el espacio arbitrariamente lejos y si se mide el primero de ellos ya se conoce el resultado de la medida del segundo.

En un n-qubit podemos medir cualquiera de los qubits. Por ejemplo si en

$$\phi = \sum_{x=0}^{2^n-1} a_x|x\rangle$$

medimos el primer qubit y el resultado de la medida es 0, el estado se proyectará sobre el subespacio generado por los vectores de la base que tienen cero en el primer qubit y si el resultado es 1 se proyectará sobre el subespacio correspondiente. La probabilidad de cada opción viene dada por la suma de los cuadrados de las amplitudes correspondientes. Observemos que el primer bit de la expresión binaria (con n bits) de x es 0 para $x = 0, \ldots, 2^{n-1} - 1$ y es 1 para el resto. Así el resultado de la medida de ϕ puede ser:

- 0 con probabilidad $p_0 = \sum_{x=0}^{2^{n-1}-1} |a_x|^2$ y estado final $\frac{1}{\sqrt{p_0}} \sum_{x=0}^{2^{n-1}-1} a_x|x\rangle$.

- 1 con probabilidad $p_1 = \sum_{x=2^{n-1}}^{2^n-1} |a_x|^2$ y estado final $\frac{1}{\sqrt{p_1}} \sum_{x=2^{n-1}}^{2^n-1} a_x|x\rangle$.

La tabla siguiente muestra el esquema de medición del k-ésimo qubit.

Estado	Medida	Estado final	Probabilidad
$\sum_{0\le x<2^n} a_x\|x\rangle$	0	$\frac{1}{\sqrt{p_0}} \sum_{\substack{0\le x<2^n \\ x_k=0}} a_x\|x\rangle$	$p_0 = \sum_{\substack{0\le x<2^n \\ x_k=0}} \|a_x\|^2$
$\sum_{0\le x<2^n} a_x\|x\rangle$	1	$\frac{1}{\sqrt{p_1}} \sum_{\substack{0\le x<2^n \\ x_k=1}} a_x\|x\rangle$	$p_1 = \sum_{\substack{0\le x<2^n \\ x_k=1}} \|a_x\|^2$

3.3 Puertas cuánticas de un qubit

La evolución de un estado cuántico se describe mediante *transformaciones cuánticas*, que son operadores lineales unitarios definidos en el espacio de Hilbert. Si se considera una base ortonormal, cada operador lineal se identifica con una matriz cuadrada, cuyas columnas son las coordenadas de las imágenes de los elementos de esta base. Si M es la matriz asociada a un operador y M^* su traspuesta conjugada, el operador lineal es unitario cuando el producto de M por M^* es la matriz unidad. Un algoritmo cuántico se llevará a cabo mediante una de estas transformaciones. Aquí tenemos una de las características importantes de la computación cuántica, que la diferencia del modelo clásico. La computación cuántica es reversible, es decir para cualquier transformación es sencillo obtener la transformación inversa.

Como no cabe esperar que sea posible implementar en un ordenador cuántico cualquier transformación unitaria, será necesario descomponerla en transformaciones elementales denominadas *puertas cuánticas*.

Las puertas cuánticas más simples son las transformaciones unitarias de un qubit, definidas en el espacio de Hilbert $\mathcal{H}$, en el que consideramos la base $B_1 = [|0\rangle, |1\rangle]$. Para definir un operador basta conocer las imágenes de los elementos de la base, o lo que es equivalente disponer de su matriz asociada.

Las cuatro puertas simples más sencillas son las marices de Pauli y están asociadas a las matrices siguientes:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

I es la transformación identidad,

X es la negación definida por $X|0\rangle = |1\rangle, X|1\rangle = |0\rangle$,

Z es el cambio de fase: $Z|0\rangle = |0\rangle, Z|1\rangle = -|1\rangle$,

$Y = iXZ$: $Y|0\rangle = i|1\rangle, Y|1\rangle = -i|0\rangle$.

Obviamente todos son operadores unitarios y, además $X^* = X$, $Z^* = Z$ e $Y^* = Y$. Por tanto $X^2 = Z^2 = Y^2 = I$.

También se puede considerar un cambio de fase con ángulo distinto de π:

$$Z_\alpha = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi\alpha} \end{pmatrix}.$$

En este caso $Z_\alpha^* = Z_{-\alpha}$.

Para el algoritmo de transformada cuántica de Fourier usaremos este tipo de puertas, concretamente denominaremos R_k a la transformación $Z_{2^{1-k}}$, cuya matriz asociada es

$$R_k = \begin{pmatrix} 1 & 0 \\ 0 & \sigma_k \end{pmatrix},$$

con $\sigma_k = e^{2\pi i/2^k}$.

También usaremos otra transformación muy habitual, la de *Hadamard*, definida por $H|0\rangle = |+\rangle$ y $H|1\rangle = |-\rangle$ y cuya matriz asociada es

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Nótese que $H^* = H$ y que por tanto $H^2 = I$.

Toda transformación unitaria de un qubit se puede poner como producto de transformaciones H y Z_α (ver [15]).

En la figura 1 se ve la representación gráfica de una puerta simple U.

Figure 1: Una puerta de 1-qubit

El circuito evoluciona de izquierda a derecha. Es decir $|y\rangle = U|x\rangle$.

Para el caso de la transformación X en la representación gráfica de circuitos se sustituye la caja con el nombre correspondiente por el símbolo $\oplus$.

3.4 Puertas cuánticas de más de un qubit

Vamos a considerar ahora transformaciones unitarias en el espacio de dimensión cuatro de los 2-qubits, $\mathcal{H}_2$. En dicho espacio, se pueden definir transformaciones que sean producto tensorial de dos puertas simples (que actúan separadamente en cada uno de los qubits) y transformaciones específicas, que no se pueden poner como producto tensorial de dos puertas simples.

Dadas dos puertas simples U y V se define el *producto tensorial* $U \otimes V$ de la forma:

$$U \otimes V|x_1x_2\rangle = U|x_1\rangle \otimes V|x_2\rangle$$

En definitiva la transformación consiste en aplicar U en el primer qubit y V en el segundo. La matriz asociada es el producto tensorial de las matrices de U y V y la representación gráfica de esta transformación definida en el sistema de 2-qubits es la natural (figura 2).

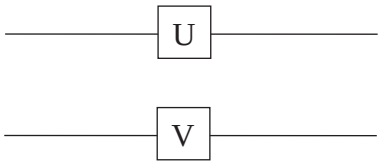


Figure 2: Dos puertas de 1-qubit

Si una de las transformaciones U o V es la identidad, obviamente no aparece representada.

El producto tensorial de una puerta simple U por la identidad es una transformación de 2-qubits consistente en modificar sólo uno de los qubits. Así, $U \otimes I$ es la puerta simple U actuando sobre el primer qubit, mientras que $I \otimes U$ representa la actuación de U sobre el segundo qubit.

Se verifica la siguiente propiedad $U \otimes V = (I \otimes V)(U \otimes I) = (U \otimes I)(I \otimes V)$. Lo que significa que puertas cuánticas que afectan a qubits distintos conmutan.

Ejemplo 1

La trasformación X en el primer qubit es $X_1 = X \otimes I$ y su matriz asociada es:

$$X_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Análogamente se define $X_2 = I \otimes X$, que modifica el segundo qubit dejando invariante el primero.

La transformación $X \otimes Z$ actúa del siguiente modo:

$$\begin{array}{ll} X \otimes Z|00\rangle = X|0\rangle \otimes Z|0\rangle = |10\rangle & X \otimes Z|01\rangle = X|0\rangle \otimes Z|1\rangle = -|11\rangle \\ X \otimes Z|10\rangle = X|1\rangle \otimes Z|0\rangle = |00\rangle & X \otimes Z|11\rangle = X|1\rangle \otimes Z|1\rangle = -|01\rangle \end{array}$$

Y la matriz asociada es $X \otimes Z = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$.

La transformación de *Walsh-Hadamard* es un ejemplo importante de transformación producto tensorial. Es la transformación definida en $\mathcal{H}_2$ por $W_2 = H \otimes H$, cuya matriz asociada es:

$$W_2 = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

Nótese que la imagen del $|0\rangle$, $W_2|00\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$, es la suma de los cuatro estados de la base. La imagen de los otros tres estados básicos también es una suma de los cuatro vectores de la base pero con la mitad de los signos positivos y la mitad negativos. En general

$$\begin{aligned} W|x_0x_1\rangle = H|x_0\rangle \otimes H|x_1\rangle &= \frac{1}{2}(|0\rangle + (-1)^{x_0}|1\rangle) \otimes (|0\rangle + (-1)^{x_1}|1\rangle) \\ &= \frac{1}{2}(|00\rangle + (-1)^{x_1}|01\rangle + (-1)^{x_0}|10\rangle + (-1)^{x_0+x_1}|11\rangle) \end{aligned}$$

Respecto a las transformaciones específicas de 2-qubits, la más sencilla de implementar es la C_{NOT} que actúa sobre $|x_1x_2\rangle$ cambiando el valor de uno de los qubits usando el otro como qubit de control. Denotaremos por C_{ij} a la transformación que cambia x_j cuando $x_i = 1$. Se pueden definir de la forma:

$$C_{12}|x_1,\, x_2\rangle = |x_1,\, x_1 \oplus x_2\rangle, \qquad C_{21}|x_1,\, x_2\rangle = |x_1 \oplus x_2,\, x_2\rangle,$$

donde $\oplus$ es la suma módulo 2. Las matrices asociadas son:

$$C_{12} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{y} \quad C_{21} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

La representación gráfica de estas puertas se puede ver en la figura 3.

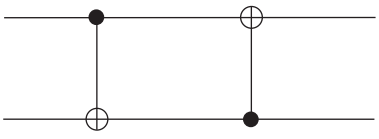


Figure 3: Puertas C_{12} y C_{21}

La puerta C_{NOT} junto con las puertas de 1-qubit constituyen un sistema universal (ver [1]), lo que significa que cualquier transformación unitaria se puede poner como composición de puertas C_{NOT} y puertas de 1-qubit.

Para cualquier puerta simple U, se puede definir la puerta de 2-qubits *control-U*, que denotaremos ΛU, de la siguiente manera:

$$\Lambda U|0x\rangle = |0x\rangle \quad \text{y} \quad \Lambda U|1x\rangle = |1\rangle \otimes U|x\rangle.$$

Es decir la puerta U actúa sobre el segundo qubit cuando el primero es 1. Con esta notación $C_{12} = \Lambda X$, porque aplica X al segundo qubit usando el primero como qubit de control.

Representaremos gráficamente la transformación ΛU como aparece en la figura 4.

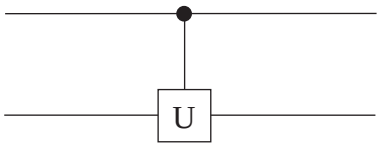


Figure 4: Puerta ΛU

Otra puerta específica de 2-qubits es la transformación *Swap*, definida por $S|x_1x_2\rangle = |x_2x_1\rangle$, cuya matriz es

$$S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Se verifica que $S = C_{12}C_{21}C_{12}$.

La puerta clásica de *Toffoli* T es una transformación de 3-qubits que cambia el tercer qubit cuando los dos primeros son 1. Evidentemente es una transformación unitaria, definida en $\mathcal{H}_3$ y su representación gráfica es la de la figura 5.

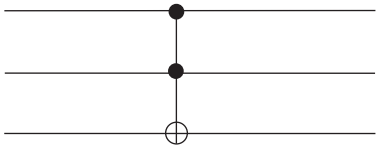


Figure 5: Puerta de Toffoli

La puerta de Toffoli, además de su utilidad en el diseño de algoritmos cuánticos, permite realizar cualquier computación clásica, ya que es universal para la computación booleana. Para demostrarlo, basta comprobar que permite construir las puertas lógicas NOT y AND:

$$T|11x\rangle = |11\overline{x}\rangle \quad \text{y} \quad T|xy0\rangle = |xy\, x \wedge y\rangle$$

La transformación de *Walsh-Hadamard* en $\mathcal{H}_n$, $W_n = H^{\otimes n}$, consiste en aplicar H a cada uno de los qubits. Su matriz asociada es $W_n = H \otimes \cdots \otimes H$, que se construye recursivamente de la forma: $W_1 = H$, y $W_n = H \otimes W_{n-1}$.

W_n transforma el $|0\rangle$ en la suma de todos los estados de la base:

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

Para todo $x = 0, \ldots, 2^n - 1$ se verifica que

$$W_n|x\rangle = \frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} (-1)^{x \cdot y} |y\rangle.$$

donde $x \cdot y = x_1 y_1 \oplus \ldots \oplus x_n y_n$, siendo $x_1 \ldots x_n$ e $y_1 \ldots y_n$ las expresiones binarias de x e y respectivamente y $\oplus$ la suma módulo 2.

Cualquier función booleana $f : \{0,1\}^n \to \{0,1\}^m$ se puede implementar mediante una transformación unitaria que trabaje con un $(n+m)$-qubit de modo que $|x\rangle \otimes |0\rangle \to |x\rangle \otimes |f(x)\rangle$. Para ello se define el operador $U_f : \mathcal{H}_{n+m} \to \mathcal{H}_{n+m}$, como

$$U_f|x, b\rangle = |x, b \oplus f(x)\rangle,$$

donde $\oplus$ es la suma módulo 2.

Ejemplo 2

Sea f una función booleana $f : \{0,1\} \to \{0,1\}$. Consideremos $U_f : \mathcal{H}_2 \to \mathcal{H}_2$, dada por $U_f|x, b\rangle = |x, b \oplus f(x)\rangle$. Si tomamos $b = 0$, entonces

$$U_f \cdot H_1|0, b\rangle = \frac{1}{\sqrt{2}}(U_f|0,0\rangle + U_f|1,0\rangle) = \frac{1}{\sqrt{2}}(|0\rangle \otimes |f(0)\rangle + |1\rangle \otimes |f(1)\rangle).$$

Tenemos los dos valores de $f(x)$ en un solo qubit. Esto es la base del paralelismo cuántico.

3.5 Teorema de no-clonning

Supongamos que tenemos un n-qubit

$$\Psi = \sum_{x=0}^{2^n-1} a_x |x\rangle$$

y que queremos conocer su estado, es decir, los coeficientes a_x para $0 \leq x < 2^n$. Si medimos cada uno de los n qubits el estado resultante será $|y\rangle$ para un $y \in \{0, 1, \ldots, 2^n - 1\}$ y los resultados de las medidas de los n qubits serán los dígitos binarios de y. Sin embargo este proceso nos da muy poca información sobre el estado Ψ: el coeficiente $|a_y|^2$ es no nulo y es probable que sea uno de los de mayor módulo. Además, con las medidas hemos destruido el estado Ψ y, por tanto, ya no podemos obtener ninguna información adicional sobre dicho estado.

Existe otro planteamiento que, en principio, daría mejores resultados. Este método consiste en hacer una copia del estado Ψ antes de medir el n-qubit. Esto nos permitiría repetir la medida de Ψ tantas veces como quisiéramos. De este modo podríamos, por ejemplo, hacer una estimación estadística de los módulos de los coeficientes de Ψ:

$$|a_y| \approx \sqrt{\frac{\text{frecuencia de } y}{\text{número total de medidas}}} \qquad \text{para todo } 0 \leq y < 2^n$$

Este estudio estadístico se podría refinar para poder estimar, además del módulo, la parte real e imaginaria de cada coeficiente. Sin embargo existen dos problemas que no permiten que este método sea eficaz para conocer el estado Ψ. La primera dificultad es que para hacer una estimación razonable de todos los coeficientes se necesitarían más de 2^n medidas y esto, para valores de n grandes, no es factible. La segunda dificultad es mucho más fundamental. Se deriva del hecho de que en el modelo de computación cuántica no se pueden copiar estados.

Teorema 1 *No existe ninguna transformación unitaria $U : \mathcal{H}_{2n} \to \mathcal{H}_{2n}$ tal que $U(\Psi \otimes |0\rangle) = \Psi \otimes \Psi$ para todo n-qubit Ψ.*

Demostración. Sea U una transformación unitaria $U : \mathcal{H}_{2n} \to \mathcal{H}_{2n}$ tal que $U(|a\rangle \otimes |0\rangle) = |a\rangle \otimes |a\rangle$ y $U(|b\rangle \otimes |0\rangle) = |b\rangle \otimes |b\rangle$ con $a \neq b$ y $0 \leq a, b < 2^n$. Consideremos el n-qubit

$$\Psi = \frac{1}{\sqrt{2}}(|a\rangle + |b\rangle).$$

Entonces

$$U(\Psi \otimes |0\rangle) = \frac{1}{\sqrt{2}}\left(U(|a\rangle \otimes |0\rangle) + U(|b\rangle \otimes |0\rangle)\right) = \frac{1}{\sqrt{2}}(|a\rangle \otimes |a\rangle + |b\rangle \otimes |b\rangle) \neq \Psi \otimes \Psi$$

En definitiva, ninguna transformación puede copiar los estados $|a\rangle$, $|b\rangle$ y $\frac{1}{\sqrt{2}}(|a\rangle + |b\rangle)$. □

3.6 Estados no distinguibles

Toda medida cuántica está asociada a una suma directa ortogonal del espacio de Hilbert que describe el sistema y viceversa, toda suma directa ortogonal está asociada a una medida cuántica. Nosotros nos vamos a restringir al espacio $\mathcal{H}_n$ de los n-qubits:

$$\mathcal{H}_n = W_1 \perp W_2 \perp \cdots \perp W_k$$

Entonces, para calcular el resultado de la medida de un n-qubit Ψ debemos hallar la representación (única) de Ψ como suma de vectores de los subespacios ortogonales W_j $(1 \leq j \leq k)$:

$$\Psi = w_1 + w_2 + \cdots + w_k \quad \text{donde} \quad w_j \in W_j \text{ para todo } 1 \leq j \leq k$$

Esta representación permite describir completamente el proceso de medida. El estado resultante de la medida será

$$\tilde{\Psi} = \frac{1}{\|w_j\|} w_j \quad \text{con probabilidad} \quad \|w_j\|^2$$

que corresponde a una proyección aleatoria sobre uno de los subespacios W_j $(1 \leq j \leq k)$ y una renormalización del estado, con probabilidad igual a la norma al cuadrado del estado proyectado.

Como consecuencia de este principio cuántico no es posible distinguir con certeza dos estados cuánticos no ortogonales. El problema de distinguir dos estados consiste en determinar el estado Ψ sabiendo que

$$\Psi = \Psi_1 \quad \text{o} \quad \Psi = \Psi_2$$

Supongamos que los estados a distinguir Ψ_1 y Ψ_2 no son ortogonales. Entonces, ninguna medida puede distinguirlos con certeza puesto que es imposible que dichos estados pertenezcan a sendos subespacios ortogonales. Si antes de medir intentamos transformarlos la situación no cambia, puesto que cualquier transformación debe ser unitaria y, por tanto, mantendrá la no ortogonalidad de los estados. Finalmente, debido al teorema de no-cloning, tampoco podemos hacer hacer copias de Ψ para, mediante un estudio estadístico de las copias, determinar cuál es el estado Ψ.

3.7 Teletransporte de estados cuánticos

Los estados entrelazados son fundamentales en computación cuántica. De hecho, si prescindiéramos de ellos los algoritmos cuánticos se podrían simular en ordenadores clásicos con un incremento polinomial de la complejidad, tanto en espacio como en tiempo.

De entre todos los estados entrelazados vamos a considerar los más significativos. Se denominan estados de Bell o estados EPR y son los siguientes:

$$\beta_{00} = \frac{|00\rangle + |11\rangle}{\sqrt{2}} \qquad \beta_{01} = \frac{|01\rangle + |10\rangle}{\sqrt{2}}$$
$$\beta_{10} = \frac{|00\rangle - |11\rangle}{\sqrt{2}} \qquad \beta_{11} = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

Se pueden construir aplicando las puertas H y C_{NOT} tal como se indica en el siguiente circuito (figura 6).

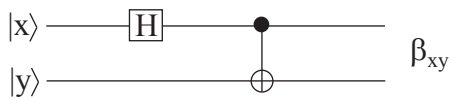


Figure 6: Circuito para generar los estados de Bell

En las comunicaciones cuánticas entre dos interlocutores (Alice y Bob) es muy útil compartir estados de Bell. Entre otras aplicaciones, si Alice y Bob comparten un estado β_{00} (Alice tiene el primer qubit y Bob el segundo) entonces Alice puede transferir el estado de un qubit arbitrario Ψ al qubit del par EPR de Bob enviando para ello únicamente dos bits de información. Este proceso se denomina teletransporte de estados cuánticos y se implementa del siguiente modo (figura 7):

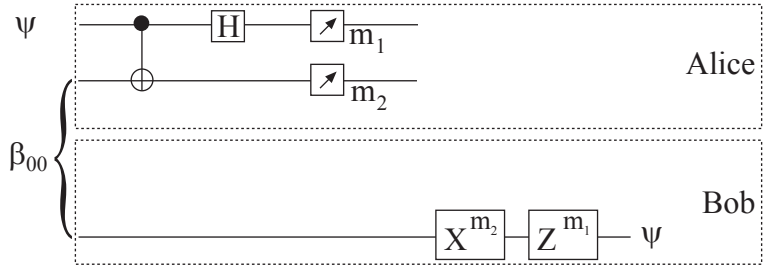


Figure 7: Circuito para teletransportar qubits

Para comprobar que, en efecto, se produce el teletransporte del qubit $\Psi = a|0\rangle + b|1\rangle$ al qubit del par EPR de Bob vamos a seguir la evolución de los tres qubits. Obsérvese que Alice comunica a Bob los resultados de las medidas del qubit Ψ y de su qubit del par EPR (m_1 y m_2) y que Bob manipula su qubit en función de estos valores para recuperar el estado Ψ.

$$\begin{aligned}
\Psi \otimes \beta_{00} &= \frac{1}{\sqrt{2}}\big[a|0\rangle \otimes \big(|00\rangle + |11\rangle\big) + b|1\rangle \otimes \big(|00\rangle + |11\rangle\big)\big] \\
&\longrightarrow \frac{1}{\sqrt{2}}\big[a|0\rangle \otimes \big(|00\rangle + |11\rangle\big) + b|1\rangle \otimes \big(|10\rangle + |01\rangle\big)\big] \\
&\longrightarrow \frac{1}{2}\big[a\big(|0\rangle + |1\rangle\big) \otimes \big(|00\rangle + |11\rangle\big) + b\big(|0\rangle - |1\rangle\big) \otimes \big(|10\rangle + |01\rangle\big)\big] \\
&= \frac{1}{2}\big[|00\rangle \otimes \big(a|0\rangle + b|1\rangle\big) + |01\rangle \otimes \big(a|1\rangle + b|0\rangle\big) \\
&\quad + |10\rangle \otimes \big(a|0\rangle - b|1\rangle\big) + |11\rangle \otimes \big(a|1\rangle - b|0\rangle\big)\big]
\end{aligned}$$

Dependiendo de las medidas obtenidas por Alice la evolución del sistema es diferente:

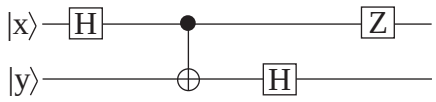
$$\begin{aligned}
m_1m_2 = 00: &\longrightarrow |00\rangle \otimes \big(a|0\rangle + b|1\rangle\big) \longrightarrow |00\rangle \otimes \big(a|0\rangle + b|1\rangle\big) \\
m_1m_2 = 01: &\longrightarrow |01\rangle \otimes \big(a|1\rangle + b|0\rangle\big) \longrightarrow |01\rangle \otimes \big(a|0\rangle + b|1\rangle\big) \\
m_1m_2 = 10: &\longrightarrow |10\rangle \otimes \big(a|0\rangle - b|1\rangle\big) \longrightarrow |10\rangle \otimes \big(a|0\rangle + b|1\rangle\big) \\
m_1m_2 = 11: &\longrightarrow |11\rangle \otimes \big(a|1\rangle - b|0\rangle\big) \longrightarrow |11\rangle \otimes \big(a|0\rangle + b|1\rangle\big)
\end{aligned}$$

Nótese que Alice pierde el estado del qubit Ψ en el proceso. En caso contrario dispondríamos de un dispositivo para copiar estados, en contra del teorema de no-cloning.

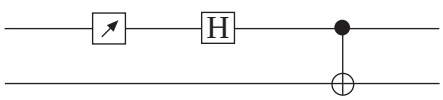
Compartir estados EPR les permite a Alice y Bob teletransportar estados cuánticos usando exclusivamente comunicaciones clásicas. Pero esta no es la única aplicación. Si disponen de suficientes pares EPR compartidos pueden generar una clave privada aleatoria segura. Basta con que ambos midan sus qubits ya que, por las propiedades especiales del estado de Bell β_{00}, los resultados de sus medidas siempre coinciden.

3.8 Ejercicios

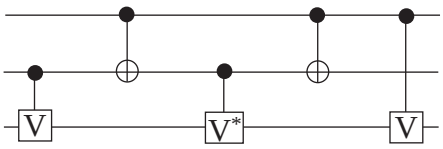
1. Calcular la medida del estado $\Psi = a|0\rangle + b|1\rangle$, con a y b reales, en la base $B_\times$.
2. Demostrar que el estado $\beta_{01} = \dfrac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ es entrelazado.
3. Probar que $XYZ = iI$, $HXH = Z$ y $HXH = Z$.
4. Hallar las matrices de las transformaciones de dos qubits Z_1 y Z_2.
5. Calcular el efecto del siguiente circuito sobre los estados de la base B_2.



6. Expresar la puerta S (swap) como producto de puertas C_{NOT} y dibujar el circuito correspondiente.
7. Demostrar que C_{NOT} no es producto tensorial de dos puertas de un qubit.
8. Construir un circuito que transforme el estado $(a|0\rangle + b|1\rangle) \otimes |00\rangle$ en $a|000\rangle + b|111\rangle$.
9. Si sólo podemos medir en la base de computación B_1, ¿qué deberemos hacer para medir el estado $\Psi = a|0\rangle + b|1\rangle$ en la base $B_\times$?
10. Hallar las matrices de las transformaciones de tres qubits T_{123} y T_{132}.
11. Probar que $H_2\, \Lambda X\, H_2 = \Lambda Z$ y $H_2\, \Lambda Z\, H_2 = \Lambda X$.
12. Obtener un circuito equivalente al de la figura realizando la medida al final, añadiendo para ello un qubit auxiliar.



13. (a) Calcular una transformación de un qubit V tal que $V^2 = X$.
    (b) Comprobar que el siguiente circuito es equivalente a la puerta de Toffoli.



14. Encontrar un circuito que transforme

$$|0\rangle \otimes |000\rangle \longrightarrow \tfrac{1}{2}\big(|00\rangle + |11\rangle\big) \otimes \big(|00\rangle + |11\rangle\big)$$
$$|1\rangle \otimes |000\rangle \longrightarrow \tfrac{1}{2}\big(|00\rangle - |11\rangle\big) \otimes \big(|00\rangle - |11\rangle\big)$$

4 Algoritmos cuánticos básicos

4.1 Problema de Deutsch

Problema 1 *Dada una función booleana $f : \{0,1\} \to \{0,1\}$ y su transformación unitaria asociada U_f, encontrar un algoritmo que determine si f es constante o no aplicando U_f el menor número posible de veces.*

Es obvio que clásicamente hay que evaluar $f(0)$ y $f(1)$ para resolver el problema planteado. Sin embargo, cuánticamente sólo es preciso evaluar U_f una vez. Esto se puede conseguir gracias al paralelismo cuántico que permite evaluar simultáneamente $f(0)$ y $f(1)$:

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

Aunque generalmente los algoritmos cuánticos son probabilísticos, en este caso es posible obtener un algoritmo determinista para resolver el problema. Generalmente es fácil utilizar el paralelismo cuántico, lo más difícil es conseguir que la probabilidad de obtener el resultado buscado sea grande. Vamos a ver cómo se consigue en el problema de Deutsch que dicha probabilidad sea igual a 1.

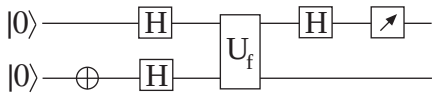


Figure 8: Circuito para el problema de Deutsch

A continuación se muestra el estado del 2-qubit después de aplicar cada una de las transformaciones unitarias del circuito y el resultado de la medida final (figura 8). Denotamos por $\bar{f}$ al complemento de la función booleana f, es decir $\bar{f}(x) = 1 \oplus f(x)$ y llamamos α a una variable booleana que nos indica si f es una función constante tomando el valor 0 o si no lo es tomando el valor 1.

$$\begin{aligned}
|0\rangle \otimes |0\rangle &\longrightarrow |0\rangle \otimes |1\rangle \\
&\longrightarrow \tfrac{1}{2}\left(|0\rangle + |1\rangle\right) \otimes \left(|0\rangle - |1\rangle\right) \\
&\longrightarrow \tfrac{1}{2}\left(|0\rangle \otimes \left(|f(0)\rangle - |\bar{f}(0)\rangle\right) + |1\rangle \otimes \left(|f(1)\rangle - |\bar{f}(1)\rangle\right)\right) \\
&= \tfrac{1}{2}\left(|0\rangle + (-1)^{\alpha}|1\rangle\right) \otimes \left(|f(0)\rangle - |\bar{f}(0)\rangle\right) \\
&\longrightarrow \tfrac{1}{\sqrt{2}}|\alpha\rangle \otimes \left(|f(0)\rangle - |\bar{f}(0)\rangle\right) \\
&\longrightarrow \alpha \quad \text{(probabilidad = 1)}
\end{aligned}$$

Por tratarse de un problema de complejidad constante, es decir, que no depende de un parámetro $n \in \mathbb{N}$, no podemos sacar ninguna conclusión al compararlo con los algoritmos clásicos. Para poder hacerlo vamos a introducir una generalización del problema. Se dice que una función booleana $f : \{0,1\}^n \to \{0,1\}$ es *balanceada* si toma el valor 0 el mismo número de veces que el valor 1.

4.2 Problema de Deutsch-Jozsa

Problema 2 *Dada una función booleana* $f : \{0,1\}^n \to \{0,1\}$ *constante o balanceada y su transformación unitaria asociada* U_f*, encontrar un algoritmo que determine si* f *es constante o es balanceada aplicando* U_f *el menor número posible de veces.*

Es sencillo probar que clásicamente hay que evaluar f sobre la mitad más uno de los elementos del dominio, es decir sobre $2^{n-1}+1$ elementos, para resolver el problema planteado de forma determinista. Sin embargo, cuánticamente basta evaluar U_f una sola vez para obtener la solución determinista. Por tanto en la resolución de este problema el modelo cuántico de computación permite una ganancia exponencial respecto al modelo clásico. Éste es el primer problema en el que se ha conseguido un aprovechamiento óptimo del paralelismo cuántico.

En la siguiente figura 8 se muestra el circuito correspondiente al algoritmo cuántico de Deutsch-Jozsa, en la primera linea el input es un n-qubit y en la segunda un qubit.

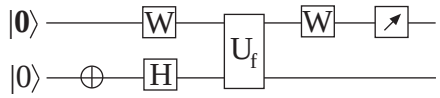


Figure 9: Circuito para el problema de Deutsch-Jozsa

A continuación se muestra el estado del $(n+1)$-qubit después de aplicar cada una de las transformaciones unitarias del circuito. El resultado final se puede separar en dos sumandos, atendiendo al estado del n-qubit: en el primero el n-qubit está en estado $|0\rangle$ y en el segundo es ortogonal al estado $|0\rangle$.

$$
\begin{aligned}
|0\rangle \otimes |0\rangle &\longrightarrow |0\rangle \otimes |1\rangle \\
&\longrightarrow \frac{1}{\sqrt{2^{n+1}}} \sum_{0\le x<2^n} (|x\rangle \otimes (|0\rangle - |1\rangle)) \\
&\longrightarrow \frac{1}{\sqrt{2^{n+1}}} \sum_{0\le x<2^n} \left(|x\rangle \otimes \left(|f(x)\rangle - |\bar{f}(x)\rangle\right)\right) \\
&\longrightarrow \frac{1}{2^n\sqrt{2}} \sum_{0\le x<2^n} \sum_{0\le y<2^n} \left((-1)^{x\cdot y}|y\rangle \otimes \left(|f(x)\rangle - |\bar{f}(x)\rangle\right)\right) \\
&= \frac{1}{2^n\sqrt{2}}|0\rangle \otimes \sum_{0\le x<2^n} \left(|f(x)\rangle - |\bar{f}(x)\rangle\right) + \\
&\quad \frac{1}{2^n\sqrt{2}} \sum_{0<y<2^n} \left(|y\rangle \otimes \sum_{0\le x<2^n} \left((-1)^{x\cdot y}\left(|f(x)\rangle - |\bar{f}(x)\rangle\right)\right)\right)
\end{aligned}
$$

Analicemos el resultado anterior suponiendo que f es constante y que es balanceada. Si es constante, el estado final del n-qubit es $|0\rangle$ y si es balanceada, es ortogonal a $|0\rangle$. Por lo tanto la medida del n-qubit permite determinar de forma determinista si f es constante o balanceada.

$$\text{constante:}\quad \frac{1}{\sqrt{2}}|0\rangle \otimes \left(|f(0)\rangle - |\bar{f}(0)\rangle\right) \longrightarrow 0$$

$$\text{balanceada:}\quad \frac{1}{2^n\sqrt{2}} \sum_{0<y<2^n} \left(|y\rangle \otimes \sum_{0\leq x<2^n} \left((-1)^{x\cdot y}\left(|f(x)\rangle - |\bar{f}(x)\rangle\right)\right)\right) \longrightarrow \neq 0$$

Si sólo se busca una solución probabilística entonces existen algoritmos clásicos que evalúan la función $O(1)$ veces, donde la constante que esconde la O depende de la cota de error ϵ. Basta elegir aleatoriamente $O(1)$ elementos distintos del dominio de f y evaluar la función en los mismos. Si se obtienen valores distintos f es con seguridad balanceada mientras que si todos los valores son iguales f es probablemente constante, con un margen de error ϵ.

4.3 Problema de Simon

Una de las primeras cuestiones que se resolvieron en computación cuántica, con ganancia exponencial respecto a los algoritmos clásicos, fue el problema de Simon. En el espacio vectorial $(\mathbb{Z}_2^n, \oplus, \odot_{\mathbb{Z}_2})$ una función booleana $f : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ es *periódica de periodo* $T \in \mathbb{Z}_2^n$ si para todo $k \in \mathbb{Z}_2^n$ se cumple $f(k \oplus T) = f(k)$ y es *2 a 1* si para todo $k \in \mathbb{Z}_2^n$ se cumple que $f^{-1}(k)$ tiene cardinal 0 ó 2. Y se plantea el siguiente problema.

Problema 3 *El problema consiste en determinar el periodo de una función periódica y 2 a 1, f, con el menor número posible de evaluaciones de la función. En el modelo cuántico de computación la función f se implementa con la transformación unitaria de $2n$ qubits*

$$U_f\left(|k\rangle \otimes |j\rangle\right) = |k\rangle \otimes |j \oplus f(k)\rangle.$$

Desde el punto de vista clásico hay que evaluar f sobre la mitad más uno de los elementos del dominio, es decir sobre $2^{n-1}+1$ elementos, para estar seguros de encontrar dos elementos k y j tales que $f(k) = f(j)$. A partir de estos valores k y j se puede calcular el periodo de la función: $T = k \oplus j$. Si sólo buscamos una solución probabilística, con cota de error ϵ, habría que evaluar la función al menos $2^{(n-1)/2}\sqrt{1-\epsilon}$ veces.

En efecto, sea h el número de veces que se evalúa la función y P la probabilidad de obtener la solución del problema de Simon, es decir la probabilidad de evaluar f en dos puntos del dominio k y j tales que $f(k) = f(j)$. Supondremos que $h \leq 2^{n-1}$, en caso contrario $P = 1$. Entonces se cumple que

$$1 - P = \frac{2^h \binom{2^{n-1}}{h}}{\binom{2^n}{h}} = 2^h \, \frac{2^{n-1}}{2^n} \cdots \frac{2^{n-1}-h+1}{2^n-h+1} = \frac{2^n}{2^n} \cdots \frac{2^n-2h+2}{2^n-h+1}$$

Acotando cada uno de los factores por el más pequeño y usando la desigualdad $(1-x)^h \geq 1-hx$, para $0 \leq x \leq 1$, queda

$$1-P \geq \left(\frac{2^n-2(h-1)}{2^n-(h-1)}\right)^h = \left(1-\frac{h-1}{2^n-(h-1)}\right)^h \geq 1-\frac{h(h-1)}{2^{n-1}+(2^{n-1}-h+1)} \geq 1-\frac{h^2}{2^{n-1}}$$

Para obtener la solución con una cota de error ϵ se debe cumplir que $P \geq 1-\epsilon$, es decir, $1-P \leq \epsilon$. Y teniendo en cuenta la desigualdad anterior se debe cumplir

$$1-\frac{h^2}{2^{n-1}} \leq \epsilon \quad \Longrightarrow \quad h \geq 2^{(n-1)/2}\sqrt{1-\epsilon}$$

Sin embargo, como veremos enseguida, cuánticamente son suficientes $O((n-1)\log(\epsilon^{-1}))$ evaluaciones. El circuito del algoritmo propuesto puede verse en la figura 10.

Algoritmo 1 (Simon)

1. Inicializar el $2n$-qubit $\Psi = |0\rangle \otimes |0\rangle$.
2. Aplicar la transformada de Walsh-Hadamard W_n a los n primeros qubits.
3. Aplicar U_f.
4. Medir los n últimos qubits: $j_1, \ldots, j_n$ (resultado $j = j_1 \ldots j_n$).
5. Aplicar de nuevo W_n a los n primeros qubits.
6. Medir los n primeros qubits: $k_1, \ldots, k_n$. Devolver $k = k_1 \ldots k_n$.

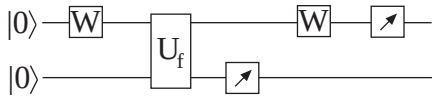


Figure 10: Circuito para el problema de Simon

El resultado final es un número entero k tal que $T \cdot k = 0$. Para comprobarlo basta desarrollar el algoritmo paso a paso. La primera medida da como resultado cualquier valor de la imagen de f con probabilidad 2^{-n+1}, por ser una función 2 a 1, y la segunda cualquier valor k que cumple $T \cdot k = 0$ también con probabilidad 2^{-n+1}.

$$\begin{aligned}
|0\rangle \otimes |0\rangle \quad &\xrightarrow{2} \quad \frac{1}{\sqrt{2^n}} \sum_{0\leq k<2^n} (|k\rangle \otimes |0\rangle) \quad \xrightarrow{3} \quad \frac{1}{\sqrt{2^n}} \sum_{0\leq k<2^n} (|k\rangle \otimes |f(k)\rangle) \\
&\xrightarrow{4} \quad \frac{1}{\sqrt{2}} (|l\rangle + |l \oplus T\rangle) \otimes |j\rangle \quad \big(\text{medida } j,\ \Rightarrow\ \{\, l, l \oplus T \,\} = f^{-1}(j)\big) \\
&\xrightarrow{5} \quad \frac{1}{\sqrt{2^{n+1}}} \sum_{0\leq k<2^n} \left((-1)^{l\cdot k} + (-1)^{(l\oplus T)\cdot k}\right) |k\rangle \otimes |j\rangle \\
&= \quad \frac{1}{\sqrt{2^{n+1}}} \sum_{0\leq k<2^n} (-1)^{l\cdot k} \left(1 + (-1)^{T\cdot k}\right) |k\rangle \otimes |j\rangle \\
&\xrightarrow{6} \quad k \quad \text{tal que} \quad T \cdot k = 0
\end{aligned}$$

Al aplicar el algoritmo evaluamos una vez la transformación unitaria U_f y obtenemos una ecuación lineal homogénea, $T \cdot k = 0$, con n incógnitas y coeficientes en el cuerpo $\mathbb{Z}_2$. Debemos repetir el algoritmo hasta obtener un sistema lineal homogéneo de rango $n-1$. La solución no nula de este sistema será el periodo de la función. Aplicando el algoritmo $n-1$ veces la probabilidad de obtener un sistema de rango $n-1$ es mayor que $1/5$, por tanto son suficientes $O((n-1)\log(\epsilon^{-1}))$ repeticiones.

Vamos a comprobar este resultado: Sea P la probabilidad de que con $n-1$ evaluaciones del algoritmo 1 se obtenga un sistema lineal homogéneo de rango máximo. La ecuación $T \cdot k = 0$ tiene exactamente 2^{n-1} soluciones en $\mathbb{Z}_2^n$ (número de vectores en un subespacio de dimensión $n-1$), todas con la misma probabilidad de ser el resultado del algoritmo. Un argumento análogo permite probar que si tenemos j ecuaciones independientes, $0 \leq j < n-1$, la $(j+1)$-ésima ecuación será independiente para $2^{n-1} - 2^j$ valores de k que, igual que antes, son equiprobables. Por tanto se cumple

$$P = \frac{2^{n-1} - 2^0}{2^{n-1}} \frac{2^{n-1} - 2^1}{2^{n-1}} \cdots \frac{2^{n-1} - 2^{n-2}}{2^{n-1}} = \prod_{j=1}^{n-1} \left(1 - \frac{1}{2^j}\right)$$

El producto anterior decrece cuando n tiende a ∞. Si llamamos l el límite de dicho producto, tomamos logaritmos y los desarrollamos en series de potencias queda

$$-\log(l) = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \frac{1}{k} \left(\frac{1}{2}\right)^{jk} = \sum_{m=1}^{\infty} b_m \left(\frac{1}{2}\right)^m$$

donde b_m es la suma de los inversos de los divisores de m. Puesto que $b_m \leq 1 + \dfrac{m-1}{2}$ para todo $m \geq 1$ se cumple

$$-\log(l) \leq \sum_{m=1}^{\infty} \left(1 + \frac{m-1}{2}\right) \left(\frac{1}{2}\right)^m = \frac{3}{2} \qquad \Longrightarrow \qquad P \geq l \geq e^{-3/2} > \frac{1}{5}$$

4.4 Ejercicios

1. Obtener un algoritmo clásico para resolver el problema de Deutsch-Jozsa con una cota de error ϵ, evaluando la función un número de veces N que dependa de ϵ pero no del número de qubits n.

2. Resolver clásicamente el problema de Simon suponiendo que la función f es una aplicación afín en $\mathbb{Z}_2^n$. ¿Cuál es el menor número de evaluaciones de f que siempre permiten encontrar su periodo?

3. Demostrar que una función $f : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ periódica y 2 a 1 no tiene por qué ser una aplicación afín.

5 Algoritmo de Grover

Muchos problemas, que se pueden denominar problemas de búsqueda, se pueden plantear de la siguiente forma: "Hallar x en un conjunto de posibles soluciones tal que la sentencia $f(x)$ sea cierta".

Por ejemplo, el problema de buscar un factor para un número entero q se puede tratar con un algoritmo de búsqueda en un lista que contenga a todos los enteros entre 2 y $\sqrt{q}$, buscando uno que verifique la propiedad $q \bmod x = 0$.

Un problema de búsqueda no estructurada es aquel para cuya resolución no se puede usar ninguna hipótesis relativa a la estructura del espacio de soluciones.

Por ejemplo, una búsqueda en una lista alfabetizada, como la búsqueda del número de un usuario en la guía telefónica, es un problema de búsqueda estructurada. Mientras que la búsqueda en la misma guía del titular de un número concreto sería una búsqueda no estructurada. En una búsqueda estructurada, se puede usar la estructura ordenada de la lista para construir algoritmos eficientes, pero en una búsqueda no estructurada lo habitual es comprobar aleatoriamente la veracidad de la sentencia. En el modelo de computación clásico, para un espacio de búsqueda de tamaño N, sería necesario evaluar f un promedio de $N/2$ veces y N veces en el peor de los casos. Los algoritmos clásicos de búsqueda no estructurada requieren por tanto $O(N)$ evaluaciones de f.

Grover [20] probó en 1996 que, con computación cuántica, un problema de búsqueda no estructurada, con solución única, se puede resolver con $O(\sqrt{N})$ evaluaciones de f. Posteriormente, Boyer, Brassard, Hoyer y Tapp [7] generalizaron el algoritmo para problemas con solución múltiple (obviamente, el caso interesante es cuando el número de soluciones es pequeño en relación al tamaño de la lista). También estudiaron la posibilidad de que no se conozca a priori el número de soluciones.

Por otra parte, en el mismo artículo, los autores obtuvieron una cota inferior para la eficiencia de cualquier algoritmo cuántico de búsqueda, probando que el de Grover es asintóticamente óptimo. Un artículo más reciente de Zalka [34] mejora este resultado y prueba que para un problema con solución única, con cualquier número de iteraciones por debajo de $\pi\sqrt{N}/4$, el algoritmo de Grover da la máxima probabilidad de encontrar el elemento deseado.

5.1 Descripción del algoritmo

Partimos de una lista de tamaño N. Supondremos, incrementando la lista si es preciso, que $N = 2^n$ para algún n. Trabajaremos con los índices de los elementos de la lista, es decir con $x = 0, \ldots, 2^n - 1$, y queremos localizar aquellos x tales que $f(x) = 1$, para una cierta función booleana f.

La computación cuántica permite evaluar f simultáneamente sobre todos los posibles inputs, sin más que construir el estado

$$\psi = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle,$$

que se obtiene, a partir de $|0\rangle$, con la transformación de Walsh-Hadamard. El problema es que no se puede leer el resultado obtenido sin destruir el estado.

La idea que aplicaremos es la siguiente: partiremos del estado ψ y lo iremos modificando de modo que se vaya incrementando la amplitud de los x tales que $f(x) = 1$ y disminuyendo la de aquellos que no verifican la propiedad. Así conseguiremos, al medir el registro resultante, tener un acierto con probabilidad alta. Para conseguir este efecto, en un estado cuántico

$$\sum_{x=0}^{N-1} a_x|x\rangle,$$

primero cambiaremos a_x por $-a_x$ para los x tales que $f(x) = 1$ y después llevaremos a cabo la operación denominada inversión sobre el promedio, que transforma

$$\sum_{x=0}^{N-1} a_x|x\rangle \quad \text{en} \quad \sum_{x=0}^{N-1}(2A - a_x)|x\rangle,$$

donde A es el promedio de los a_x.

Nótese que si los a_x son números reales, después de realizar estas operaciones se tendrá un estado cuyas amplitudes serán también números reales.

Veamos cómo llevar a cabo cada una de estas operaciones:

1. Cambio de signo de la amplitud. Se trata de implementar la transformación

$$U|x\rangle = (-1)^{f(x)}|x\rangle,$$

que no modifica los $|x\rangle$ tales que $f(x) = 0$ y pone un coeficiente -1 a los que verifican $f(x) = 1$.

Sabemos que la puerta cuántica $U_f : |x, b\rangle \to |x, b \oplus f(x)\rangle$ implementa la evaluación de la función booleana f.

Si tomamos $|b\rangle = \dfrac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ tendremos:

$$\begin{aligned} f(x) = 0 \quad &\Rightarrow \quad |b \oplus f(x)\rangle = |b\rangle \\ f(x) = 1 \quad &\Rightarrow \quad |b \oplus f(x)\rangle = \frac{1}{\sqrt{2}}(|1\rangle - |0\rangle) = -|b\rangle \end{aligned}$$

Luego $U_f|x, b\rangle = (-1)^{f(x)}|x, b\rangle$.

Con esta transformación se realiza simultáneamente la evaluación de f y la inversión de las amplitudes de los $|x\rangle$ que satisfacen la propiedad.

Vamos a detallar la actuación de U_f sobre un estado cualquiera $\phi = \sum_{x=0}^{N-1} a_x|x_x\rangle$.

Sean $X_0 = \{x : f(x) = 0\}$ y $X_1 = \{x : f(x) = 1\}$. Entonces:

$$\begin{aligned} U_f(|\phi\rangle \otimes |b\rangle) &= U_f\left(\sum_{x\in X_0} a_x|x,b\rangle + \sum_{x\in X_1} a_x|x,b\rangle\right) \\ &= \left(\sum_{x\in X_0} a_x|x,b\rangle - \sum_{x\in X_1} a_x|x,b\rangle\right) \\ &= \left(\sum_{x\in X_0} a_x|x\rangle - \sum_{x\in X_1} a_x|x\rangle\right) \otimes |b\rangle \end{aligned}$$

Por tanto el efecto de U_f es el cambio de signo de las amplitudes, tal como queríamos.

Nótese que en la salida de U_f no se modifica el estado b, por tanto podemos omitirlo y en lo que sigue nos referiremos a esta transformación, que en la literatura especializada se suele denominar el "oráculo", como

$$U : |x\rangle \to (-1)^{f(x)}|x\rangle.$$

2. Inversión sobre el promedio. Dado $\psi = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle$, es fácil ver que la transformación $G = 2|\psi\rangle\langle\psi| - I$ cuya matriz asociada es

$$G = \begin{pmatrix} \frac{2}{N} - 1 & \frac{2}{N} & \cdots & \frac{2}{N} \\ \frac{2}{N} & \frac{2}{N} - 1 & \cdots & \frac{2}{N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{2}{N} & \frac{2}{N} & \cdots & \frac{2}{N} - 1 \end{pmatrix}$$

transforma $\sum_{x=0}^{N-1} a_x|x\rangle$ en $\sum_{x=0}^{N-1} (2A - a_x)|x\rangle$, donde A es el promedio de los a_x.

Grover propuso una implementación eficiente de esta transformación con $O(\log(N))$ puertas elementales. Para ello basta probar que $G = W_n R W_n$, donde W_n es la transformación de Walsh-Hadamard y $R = (r_{ij})$ es una matriz diagonal tal que $r_{11} = 1$ y $r_{ii} = -1$ para $2 \leq i < N$.

La matriz R se puede escribir como $R = R' - I$ donde $R' = 2|0\rangle\langle 0|$ es una matriz con todos sus elementos iguales a 0 menos el de la primera fila y primera columna que es $r'_{11} = 2$. Entonces

$$W_n R W_n = W_n(2|0\rangle\langle 0| - I)W_n = W_n 2|0\rangle\langle 0|W_n - I$$

y como $W_n|0\rangle = |\psi\rangle$, se concluye que

$$W_n R W_n = 2|\psi\rangle\langle\psi| - I = G.$$

La transformación R se puede implementar con $n = \log(N)$ puertas de Toffoli y el coste de W_n también es $\log(N)$.

El algoritmo de Grover consiste en aplicar reiteradamente las transformaciones U y G. Pero debemos determinar cuántas veces hacerlo para maximizar la probabilidad de acierto.

Ejemplo 3

Supongamos que se tiene una lista de 64 elementos de los que sólo uno, que denominamos y, verifica la propiedad. Entonces se verifica que $|X_0| = 63$, $|X_1| = 1$ y $X_1 = \{\, y \,\}$.

En primer lugar se construye el estado $\frac{1}{8}\sum_{x=0}^{63} |x\rangle$. Inicialmente todos los coeficientes son iguales a $\frac{1}{8}$.

Cambiamos el signo de la amplitud de x_0 y hacemos el promedio, que será

$$A = \left(63\frac{1}{8} - \frac{1}{8}\right)\frac{1}{64} \approx 0.12109$$

Hacemos la operación de inversión en el promedio y la nueva amplitud para $|y\rangle$ es 0.367817, mientras que para el resto es 0.117187.

Si repetimos el proceso, después de 6 iteraciones, el coeficiente de $|y\rangle$ es 0.998291, mientras que el resto de los coeficientes son iguales a -0.00736174.

Si en este momento medimos el estado, tendremos una probabilidad de acierto de 0.998291^2, pero si seguimos iterando las cosas empiezan a cambiar. Por ejemplo después de 10 iteraciones la amplitud de $|y\rangle$ es aproximadamente 0.487922. Por esto es importante determinar el número adecuado de veces que hay que aplicar el algoritmo.

5.2 Visión geométrica del algoritmo

Partimos de una lista de $N = 2^n$ elementos de los cuales s verifican la propiedad en cuestión.

Construimos el estado $\psi = \frac{1}{\sqrt{N}}\sum_{x=0}^{N-1} |x\rangle$ y consideramos los estados

$$|\alpha\rangle = \frac{1}{\sqrt{N-s}}\sum_{x\in X_0} |x\rangle \quad \text{y} \quad |\beta\rangle = \frac{1}{\sqrt{s}}\sum_{x\in X_1} |x\rangle$$

con $X_0 = \{x : f(x) = 0\}$ y $X_1 = \{x : f(x) = 1\}$.

Evidentemente $|\alpha\rangle$ y $|\beta\rangle$ son ortonormales y podemos considerar el subespacio de Hilbert real $L_{I\!R}(|\alpha\rangle, |\beta\rangle) = \{a|\alpha\rangle + b|\beta\rangle, a, b \in I\!R\}$, que tiene como base $[|\alpha\rangle, |\beta\rangle]$.

En este espacio podemos escribir

$$\psi = \sqrt{\frac{N-s}{N}}|\alpha\rangle + \sqrt{\frac{s}{N}}|\beta\rangle$$

y las transformaciones U y G verifican que

$$U(L_{I\!R}(|\alpha\rangle,|\beta\rangle)) = L_{I\!R}(|\alpha\rangle,|\beta\rangle) \quad \text{y} \quad G(L_{I\!R}(|\alpha\rangle,|\beta\rangle)) = L_{I\!R}(|\alpha\rangle,|\beta\rangle)$$

Trabajando con este plano euclídeo, la transformación U es una simetría respecto a $|\alpha\rangle$ ya que transforma el vector $a|\alpha\rangle + b|\beta\rangle$ en $a|\alpha\rangle - b|\beta\rangle$, mientras que $G = 2|\psi\rangle\langle\psi| - I$ es una simetría respecto de $|\psi\rangle$, pues su matriz asociada es

$$G = 2\begin{pmatrix} \sqrt{\frac{N-s}{N}} \\ \sqrt{\frac{s}{N}} \end{pmatrix}\left(\sqrt{\frac{N-s}{N}} \quad \sqrt{\frac{s}{N}}\right) - I = \begin{pmatrix} 1-\frac{2s}{N} & 2\frac{\sqrt{(N-s)s}}{N} \\ 2\frac{\sqrt{(N-s)s}}{N} & \frac{2s}{N}-1 \end{pmatrix}.$$

La composición de dos simetrías es un giro de ángulo doble del que forman los dos vectores. Por lo tanto, trabajando en el plano $L_{I\!R}(|\alpha\rangle,|\beta\rangle)$ el algoritmo de Grover consiste en hacer un giro de ángulo θ donde $\theta \in [0, \frac{\pi}{2}]$ y $\cos(\frac{\theta}{2})$ es el producto escalar de los vectores $|\alpha\rangle$ y $|\psi\rangle$:

$$\cos\left(\frac{\theta}{2}\right) = \langle\alpha|\psi\rangle = \sqrt{\frac{N-s}{N}}$$

Obviamente $\text{sen}\left(\frac{\theta}{2}\right) = \sqrt{\frac{s}{N}}$.

Sea $\gamma = \theta/2$. En la base $[|\alpha\rangle,|\beta\rangle]$, el estado inicial es $\psi = \cos(\gamma)|\alpha\rangle + \text{sen}(\gamma)|\beta\rangle$.

La aplicación del algoritmo supone un giro de ángulo $\theta = 2\gamma$ y el estado se transforma en

$$\cos(3\gamma)|\alpha\rangle + \text{sen}(3\gamma)|\beta\rangle.$$

Después de k iteraciones tendremos el estado $\cos((2k+1)\gamma)|\alpha\rangle + \text{sen}((2k+1)\gamma)|\beta\rangle$ que, sustituyendo $|\alpha\rangle$ y $\beta\rangle$, es:

$$\cos((2k+1)\gamma)\frac{1}{\sqrt{N-s}}\sum_{x\in X_0}|x\rangle + \text{sen}((2k+1)\gamma)\frac{1}{\sqrt{s}}\sum_{x\in X_1}|x\rangle$$

El número óptimo de iteraciones k será el número de veces que hace falta hacer un giro de ángulo $\theta = 2\gamma$ para transformar $|\psi\rangle$ en un estado lo más cerca posible de $|\beta\rangle$:

$$(2k+1)\gamma \approx \frac{\pi}{2} \quad \Longrightarrow \quad k \approx \frac{\pi}{4\gamma} - \frac{1}{2}$$

5.3 Cálculo del número óptimo de iteraciones

Partimos de una lista de $N = 2^n$ elementos de los cuales s verifican la propiedad. Construimos el estado

$$\psi = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle$$

y aplicamos reiteradamente las transformaciones U y G. Después de k iteraciones, los elementos que verifican la propiedad tendrán todos la misma amplitud, que denominamos b_k, y los que no la verifican tendrán todos otra amplitud, que denominamos m_k.

Entonces, podemos escribir el estado en la forma:

$$m_k \sum_{x \in X_0} |x\rangle + b_k \sum_{x \in X_1} |x\rangle$$

En cada iteración del algoritmo, las amplitudes b_k cambian de signo (efecto de la aplicación U) y se invierten respecto del promedio (efecto de la aplicación G). Por tanto se verifican la siguientes ecuaciones recursivas:

$$\begin{cases} m_0 = \frac{1}{\sqrt{N}} & b_0 = \frac{1}{\sqrt{N}} \\ m_{k+1} = 2A_k - m_k & b_{k+1} = 2A_k + b_k \end{cases}$$

donde

$$A_k = \frac{(N-s)m_k - s\, b_k}{N}$$

Es decir, para $k \geq 0$ se verifica:

$$\begin{pmatrix} m_{k+1} \\ b_{k+1} \end{pmatrix} = \begin{pmatrix} \frac{N-2s}{N} & -\frac{2s}{N} \\ \frac{2N-2s}{N} & \frac{N-2s}{N} \end{pmatrix} \begin{pmatrix} m_k \\ b_k \end{pmatrix}$$

Se puede demostrar por inducción que la solución de este sistema de ecuaciones en diferencias es precisamente el resultado obtenido geométricamente. Es decir:

$$m_k = \frac{1}{\sqrt{N-s}} \cos((2k+1)\gamma) \quad \text{y} \quad b_k = \frac{1}{\sqrt{s}} \operatorname{sen}((2k+1)\gamma)$$

donde

$$\cos(\gamma) = \sqrt{\frac{N-s}{N}} \quad \text{y} \quad \operatorname{sen}(\gamma) = \sqrt{\frac{s}{N}}.$$

Para conseguir la máxima probabilidad de acierto, habría que minimizar $|m_k|$. Se verifica que $m_k = 0$ si $(2k+1)\gamma = \pi/2$. Es decir si

$$k = \frac{\pi}{4\gamma} - \frac{1}{2},$$

como hemos visto antes. Pero esto no es posible porque el número de iteraciones debe ser entero.

Consideramos el valor óptimo de k, como número real, y tomamos $\tilde{k} = \left\lfloor \frac{\pi}{4\gamma} \right\rfloor$. Entonces se verifica que $\left|k - \tilde{k}\right| \leq \frac{1}{2}$, con lo que se obtiene

$$\left|\frac{\pi}{2} - (2\tilde{k}+1)\gamma\right| = \left|(2k+1)\gamma - (2\tilde{k}+1)\gamma\right| = \left|2\gamma(k-\tilde{k})\right| \leq \gamma$$

Después de $\tilde{k}$ iteraciones la probabilidad de fallo es:

$$(N-s)(m_{\tilde{k}})^2 = \cos^2((2\tilde{k}+1)\gamma) = \text{sen}^2(\frac{\pi}{2} - (2\tilde{k}+1)\gamma) \leq \text{sen}^2(\gamma) = \frac{s}{N}$$

En consecuencia, la probabilidad de fallo después de $\tilde{k} = \lfloor \pi/(4\gamma) \rfloor$ iteraciones es menor que s/N y el número de iteraciones es $O(\sqrt{N})$ ya que:

$$\tilde{k} = \left\lfloor \frac{\pi}{4\gamma} \right\rfloor \leq \frac{\pi}{4\,\text{sen}(\gamma)} = \frac{\pi}{4\sqrt{s/N}} = \frac{\pi}{4}\sqrt{\frac{N}{s}}$$

Veamos algunos casos particulares:

1. Si $s = 1$, $\text{sen}(\gamma) = \sqrt{1/N}$. Si N es grande, $\text{sen}(\gamma)$ es pequeño y $\gamma \approx \text{sen}(\gamma) = \sqrt{1/N}$ y entonces

   $$\left\lfloor \frac{\pi}{4\gamma} \right\rfloor \approx \frac{\pi}{4}\sqrt{N}.$$

   Con $\lfloor \frac{\pi}{4}\sqrt{N} \rfloor$ iteraciones, la probabilidad de fallo está acotada por $1/N$.

2. Si $s = N/4$, $\text{sen}(\gamma) = \sqrt{s/N} = 1/2$, luego $\gamma = \pi/6$ y el número óptimo de iteraciones es

   $$\left\lfloor \frac{\pi}{4\gamma} \right\rfloor = \left\lfloor \frac{3}{2} \right\rfloor = 1.$$

   Con una sola iteración del algoritmo se consigue una solución, con probabilidad 0.75.

3. Si $s = N/2$, $\text{sen}(\gamma) = 1/\sqrt{2}$, luego $\gamma = \pi/4$. El número óptimo de iteraciones también es 1. Pero en este caso no se mejora la probabilidad de acierto que sigue siendo 1/2.

5.4 Algoritmo de Grover conociendo el número de soluciones

Se dispone de una lista de $N = 2^n$ elementos de los cuales se sabe que hay s que verifican la propiedad. Se numeran los elementos de la lista y se trabaja con los índices correspondientes.

El siguiente algoritmo proporciona una solución con probabilidad de fallo menor que s/N. El circuito del algoritmo propuesto puede verse en la figura 11:

Algoritmo 2 (Grover)

1. Tomar como estado inicial del n-qubit $|0\rangle$.
2. Aplicar la transformada de Walsh-Hadamard, W_n.
3. Aplicar $U : |x\rangle \to (-1)^{f(x)}|x\rangle$.
4. Aplicar $G = 2|\psi\rangle\langle\psi| - I$.
5. Repetir $\lfloor \pi/(4\gamma) \rfloor$ veces los pasos 3 y 4.
6. Medir el resultado.

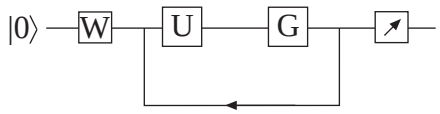


Figure 11: Circuito para el problema de Grover

Teniendo en cuenta que el número de iteraciones es $O(\sqrt{N})$ y que el coste de la transformación G es $O(\log(N))$, la complejidad total del algoritmo es $O(\sqrt{N}\log(N))$.

Ejemplo 4

Supongamos que partimos de una lista de 8 elementos $[0, 1, \ldots, 7]$ y buscamos un x que verifique la propiedad de ser un múltiplo no nulo de 5. En este caso hay una única solución $(x = 5)$ y la función f podría definirse por la propiedad $f(x) = x \bmod (\text{mcd}(x, 5))$.

La transformación unitaria U correspondiente a esta propiedad tiene como matriz:

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

Pero esta matriz no es conocida, si no se conoce la solución del problema. La transformación U actúa como caja negra (oráculo).

La matriz de la aplicación G sólo depende del tamaño de la lista y en este ejemplo es:

$$G = \frac{1}{4}\begin{pmatrix} -3 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -3 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -3 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -3 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & -3 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & -3 \end{pmatrix}$$

A continuación aplicamos el algoritmo:

Paso 1: Partimos del n-qubit $|0\rangle$.

Paso 2: $W_n(|0\rangle) = \frac{1}{\sqrt{8}} \sum_{x=0}^{7} |x\rangle$.

Pasos 3 a 5: El número de iteraciones es $\lfloor \pi\sqrt{N}/4 \rfloor = 2$.

Tras la primera aplicación de U y G las amplitudes resultantes son:

$$\frac{1}{4\sqrt{2}}[1, 1, 1, 1, 1, 5, 1, 1]$$

(Nótese que este resultado es desconocido, ya que si medimos el estado se destruye.)

Después de la segunda iteración resulta:

$$\frac{1}{8\sqrt{2}}[-1, -1, -1, -1, -1, 11, -1, -1]$$

Paso 6: Al medir tenemos la probabilidad real de éxito es $121/128 \approx 0.9453125$.

El método garantiza una probabilidad de éxito mayor o igual que $7/8 \approx 0.875$.

5.5 Ejercicios

1. Comprobar que la matriz asociada de $R = 2|0\rangle\langle 0| - I$ es

$$R = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 0 & \dots & 0 \\ 0 & 0 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -1 \end{pmatrix}$$

2. Dado $\psi = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle$, comprobar que la matriz asociada de $G = 2|\psi\rangle\langle\psi| - I$ es

$$G = \begin{pmatrix} \frac{2}{N} - 1 & \frac{2}{N} & \dots & \frac{2}{N} \\ \frac{2}{N} & \frac{2}{N} - 1 & \dots & \frac{2}{N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{2}{N} & \frac{2}{N} & \dots & \frac{2}{N} - 1 \end{pmatrix}$$

y que G transforma $\sum_{x=0}^{N-1} a_x|x\rangle$ en $\sum_{x=0}^{N-1}(2A - a_x)|x\rangle$, donde A es el promedio de los a_x.

3. Poner la transformación R como producto de puertas cuánticas. Utilizar la puerta generalizada de Toffoli

$$\Lambda T_{i_1 i_2 \dots i_k, j}$$

que aplica X al qubit j-ésimo si los qubits i_1, i_2, $\dots$, i_k están en estado $|1\rangle$, siendo k un entero entre 3 y n.

6 Transformada cuántica de Fourier

La factorización de números enteros es un problema para el que no se ha encontrado un método eficiente de resolución dentro de la computación clásica. Dicho problema se puede reducir (tal como se verá en la sección siguiente) al de encontrar el periodo de una función entera, y para este propósito se puede usar la transformada discreta de Fourier. A continuación presentamos un método cuántico de cálculo eficiente de esta transformada. La transformada cuántica de Fourier de un n-qubit es el operador lineal $F_n : \mathcal{H}_n \to \mathcal{H}_n$ que se define sobre el vector $|j\rangle$ de la base de computación, $0 \leq j < 2^n$, del siguiente modo:

$$F_n|j\rangle = \frac{1}{\sqrt{Q}} \sum_{k=0}^{Q-1} \sigma_n^{jk} |k\rangle$$

donde σ_n es la raíz 2^n-ésima de la unidad $e^{\frac{2\pi i}{2^n}}$ y $Q = 2^n$. Realmente se trata de la transformada discreta de Fourier cuya expresión habitual $(\hat{f}_0, \cdots, \hat{f}_{Q-1}) = F_n(f_0, \cdots, f_{Q-1})$, en coordenadas de la base de computación, es la siguiente:

$$\hat{f}_k = \frac{1}{\sqrt{Q}} \sum_{j=0}^{Q-1} \sigma_n^{kj} f_j$$

La transformada cuántica de Fourier es una transformación unitaria. Para obtener su inversa basta sustituir en la expresión de F_n el parámetro $\sigma_n = e^{\frac{2\pi i}{2^n}}$ por $\sigma_n^* = e^{-\frac{2\pi i}{2^n}}$.

Observemos además que F_n actúa sobre el vector $|0\rangle$ del mismo modo que la transformada de Walsh-Hadamard:

$$F_n|0\rangle = \frac{1}{\sqrt{Q}} \sum_{k=0}^{Q-1} |k\rangle$$

Ejemplo 5
Veamos cómo actúa la transformada cuántica de Fourier para $n = 1$ y para $n = 2$:

(a) Si $n = 1$, $F_1|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ y $F_1|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Por lo tanto la matriz asociada a la transformación F_1 es

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

y en este caso F_1 coincide con la transformación de Hadamard.

(b) Tomemos ahora $n = 2$.

Para $k = 0$ se tiene que $\sigma_n^{jk} = 1$ para todo j y por lo tanto

$$F_2|0\rangle = \frac{1}{2}(|0\rangle + |1\rangle + |2\rangle + |3\rangle).$$

Para $k=1$, $\sigma_n^{jk}=1$ si $j=0$, $\sigma_n^{jk}=e^{\frac{2\pi i}{4}}=i$ si $j=1$, $\sigma_n^{jk}=e^{\frac{4\pi i}{4}}=-1$ si $j=2$ y $\sigma_n^{jk}=e^{\frac{6\pi i}{4}}=-i$ si $j=3$. Por lo tanto

$$F_2|1\rangle=\frac{1}{2}(|0\rangle+i|1\rangle-|2\rangle-i|3\rangle).$$

Análogamente para $k=2$ se obtiene que $(\sigma_n^{jk})_{j=0}^3=(1,-1,1,-1)$ y entonces

$$F_2|2\rangle=\frac{1}{2}(|0\rangle-|1\rangle+|2\rangle-|3\rangle).$$

Y finalmente, para $k=3$, $(\sigma_n^{jk})_{j=0}^3=(1,-i,-1,i)$ y entonces

$$F_2|3\rangle=\frac{1}{2}(|0\rangle-i|1\rangle-|2\rangle+i|3\rangle).$$

De esta forma, la matriz asociada a la transformación F_2 es

$$\frac{1}{2}\begin{pmatrix}1&1&1&1\\1&i&-1&-i\\1&-1&1&-1\\1&-i&-1&i\end{pmatrix}$$

(c) Para un n arbitrario la matriz asociada a F_n es

$$\frac{1}{\sqrt{2^n}}\begin{pmatrix}1&1&1&1&\cdots&1\\1&\sigma_n&\sigma_n^2&\sigma_n^3&\cdots&\sigma_n^{2^n-1}\\1&\sigma_n^2&\sigma_n^4&\sigma_n^6&\cdots&\sigma_n^{(2^n-1)2}\\\vdots&\vdots&\vdots&\vdots&\ddots&\vdots\\1&\sigma_n^{2^n-1}&\sigma_n^{2(2^n-1)}&\sigma_n^{3(2^n-1)}&\cdots&\sigma_n^{(2^n-1)(2^n-1)}\end{pmatrix}$$

6.1 Propiedades de la transformada de Fourier

Si denominamos periodo de la función discreta f al menor número T, $1\leq T\leq Q$, que verifica $f_{j+T}=f_j$ para todo $0\leq j<Q$, considerando los índices módulo Q, entonces T es un divisor de Q. Si no lo fuera $T'=\mathrm{mcd}(T,Q)<T$ cumpliría $f_{j+T'}=f_j$ para todo $0\leq j<Q$, contradiciendo la minimalidad de T. En efecto, según el teorema de Bézout, existirían $m_1,m_2\in\mathbb{Z}$ tales que $T'=m_1T+m_2Q$ y, por tanto, se satisfaría $f_{j+T'}=f_{j+m_1T+m_2Q}=f_j$.

Como consecuencia de esto se puede obtener el siguiente resultado:

Proposición 1 *Dada una función $f:\mathbb{Z}\to\mathbb{C}$ de periodo T (divisor de Q), su transformada cuántica de Fourier $\hat{f}$ se anula en todos los elementos del dominio salvo en los múltiplos de la frecuencia w de la función ($wT=Q$), es decir*

$$F_n\left(\sum_{j=0}^{Q-1}f_j|j\rangle\right)=\sum_{k=0}^{T-1}\hat{f}_{wk}|wk\rangle \tag{1}$$

Demostración. Pongamos, como es habitual, $F_n(f_0, f_1, \ldots, f_{Q-1}) = (\hat{f}_0, \hat{f}_1, \cdots, \hat{f}_{Q-1})$, donde $\hat{f}_k = \frac{1}{\sqrt{Q}} \sum_{j=0}^{Q-1} \sigma_n^{jk} f_j$.

Si k no es múltiplo de w, teniendo en cuenta que $f_j = f_{j+T}$ para todo j, esta suma la podemos desarrollar del siguiente modo:

$$\begin{aligned}
\hat{f}_k &= \frac{1}{\sqrt{Q}} \left(\sigma_n^0 f_0 + \sigma_n^k f_1 + \cdots + \sigma_n^{Tk} f_T + \sigma_n^{(T+1)k} f_{T+1} + \cdots + \sigma_n^{2Tk} f_{2T} + \cdots \right) \\
&= \frac{1}{\sqrt{Q}} \left(\left(\sigma_n^0 + \sigma_n^{Tk} + \cdots + \sigma_n^{(w-1)Tk} \right) f_0 + \left(\sigma_n^k + \sigma_n^{(T+1)k} + \cdots \right) f_1 + \cdots \right) \\
&= \frac{1}{\sqrt{Q}} \left(\frac{1 - \sigma_n^{wTk}}{1 - \sigma_n^{Tk}} f_0 + \sigma_n \frac{1 - \sigma_n^{wTk}}{1 - \sigma_n^{Tk}} f_1 + \cdots \right) \\
&= 0
\end{aligned}$$

Observemos que el desarrollo anterior no es válido cuando k es múltiplo de w. Finalmente, aplicando la linealidad de F_n y el hecho de que $\hat{f}_k = 0$ si k no es múltiplo de w, se obtiene de manera inmediata la expresión (1). □

Esta propiedad permite obtener fácilmente la frecuencia w de la función f y, en consecuencia, también el periodo T. Para ello se aplica la transformada cuántica de Fourier a f y, a continuación, se miden todos los qubits. De este modo obtenemos un valor wk tal que $0 \leq k < T$ y devolvemos como resultado $w' = \text{mcd}(wk, Q)$.

- Si se cumple que el $\text{mcd}(k, T) = 1$ entonces $w' = w$, puesto que $Q = wT$.
- En caso contrario w' es un múltiplo de w.

En este último caso, si todos los valores de k entre 0 y $T - 1$ son equiprobables entonces se verifica que $P(w' = w) = \frac{\phi(T)}{T}$ (donde ϕ es la función de Euler). Teniendo en cuenta el siguiente resultado clásico de teoría de números [9]:

$$\liminf_{T \to \infty} \frac{\phi(T) \, \text{loglog}(T)}{T} = e^{-\gamma},$$

siendo $\gamma \approx 0.577215$ la constante de Euler, se llega a que

$$\liminf_{T \to \infty} P(w' = w) \, \text{loglog}(T) = e^{-\gamma} = 0.561459...$$

Así pues, para conseguir una probabilidad positiva, independiente de T y de Q, habría que repetir el proceso $O(\text{loglog}(Q))$ veces. Por tanto, tendríamos un algoritmo polinomial en el número de dígitos de Q.

6.2 Algoritmo de la transformada cuántica de Fourier

La transformada cuántica de Fourier se puede definir para valores de Q arbitrarios, sin embargo, la elección más simple para implementarla consiste en tomar $Q = 2^n$. En la

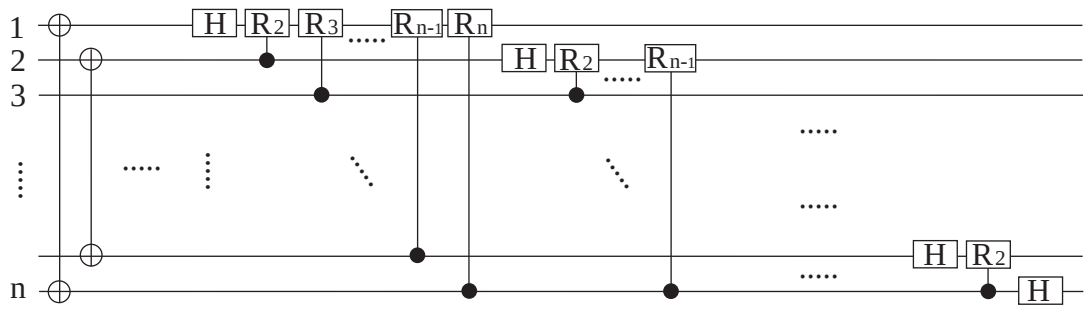


Figure 12: Algoritmo para la transformada cuántica de Fourier

figura 12 se muestra el primer algoritmo que se propuso para calcular la transformada cuántica de Fourier en este caso.

Recordemos que las líneas horizontales representan qubits que evolucionan temporalmente de izquierda a derecha y se numeran desde arriba. El símbolo H sobre una línea especifica la aplicación de la puerta cuántica H (transformación de Hadamard) sobre el qubit correspondiente a la línea. La aplicación de la puerta ΛR_k (control-R_k) se indica uniendo con un segmento vertical los símbolos $\bullet$ y R_k que se colocan sobre el qubit de control y el qubit afectado respectivamente, siendo

$$R_k = \begin{pmatrix} 1 & 0 \\ 0 & \sigma_k \end{pmatrix}$$

La puerta S (Swap) se representa uniendo con un segmento vertical dos símbolos $\oplus$ colocados sobre los dos qubits afectados.

Calculamos en primer lugar el número de puertas cuánticas que usa el algoritmo: F_n utiliza $\lfloor \frac{n}{2} \rfloor$ puertas Swap, n puertas H y

$$(n-1) + (n-2) + \cdots + 1 = \frac{n(n-1)}{2}$$

puertas ΛR_k. Por lo tanto, el algoritmo anterior calcula la transformada de Fourier de un vector de longitud $Q = 2^n$ aplicando $O(n^2)$ puertas cuánticas mientras que, clásicamente, el algoritmo de la transformada rápida de Fourier realiza $O(Q \log(Q)) = O(n2^n)$ operaciones. Shor utiliza la ganancia exponencial del algoritmo cuántico para obtener un factor propio del número natural N en tiempo $O(\log^4(N)\text{loglog}(N))$.

Sólo queda comprobar el siguiente resultado:

Teorema 2 *El algoritmo de la transformada cuántica de Fourier de la figura 12 es correcto.*

Demostración. Utilizaremos la notación $H(k)$ para indicar la transformación consistente en aplicar la puerta H al qubit k-ésimo, $S(k, j)$ para representar la puerta Swap actuando sobre los qubits k y j y $\Lambda R_k(j, m)$ para aplicar la puerta R_k sobre el qubit m-ésimo si el j-ésimo está en estado $|1\rangle$.

Dado un número natural de n dígitos binarios $k = k_n \dots k_1$ llamamos k' al número de $n-1$ dígitos que se obtiene al suprimir el dígito más significativo de k, es decir $k' = k_{n-1} \dots k_1$.

Utilizando esta notación vamos a demostrar que la transformada cuántica de Fourier se puede expresar de forma recursiva:

$$F_n|k\rangle = F_{n-1}|k'\rangle \otimes \frac{1}{\sqrt{2}}\left(|0\rangle + \sigma_n^k|1\rangle\right)$$

Para probarlo observemos que se verifican las siguientes relaciones para $0 \leq j < 2^{n-1}$:

$$|2j\rangle = |j\rangle \otimes |0\rangle \quad \text{y} \quad |2j+1\rangle = |j\rangle \otimes |1\rangle$$

$$\sigma_n^{k(2j)} = \sigma_n^{(2^{n-1}k_n + k')(2j)} = \sigma_n^{2^n k_n j}\sigma_n^{k'(2j)} = \sigma_n^{k'(2j)} = \sigma_{n-1}^{k'j}$$

$$\sigma_n^{k(2j+1)} = \sigma_n^{k(2j)}\sigma_n^k = \sigma_{n-1}^{k'j}\sigma_n^k$$

A partir de estas ecuaciones, descomponiendo la suma de la definición de la transformada cuántica de Fourier en potencias pares e impares, de modo análogo a como se hace en el algoritmo de la transformada rápida de Fourier clásica, se obtiene de forma sencilla:

$$\begin{aligned}
F_n|k\rangle = \frac{1}{\sqrt{2^n}}\sum_{j=0}^{2^n-1}\sigma_n^{kj}|j\rangle &= \frac{1}{\sqrt{2^n}}\sum_{j=0}^{2^{n-1}-1}\sigma_n^{k(2j)}|2j\rangle + \frac{1}{\sqrt{2^n}}\sum_{j=0}^{2^{n-1}-1}\sigma_n^{k(2j+1)}|2j+1\rangle \\
&= \frac{1}{\sqrt{2^n}}\sum_{j=0}^{2^{n-1}-1}\sigma_{n-1}^{k'j}|j\rangle \otimes |0\rangle + \frac{1}{\sqrt{2^n}}\sum_{j=0}^{2^{n-1}-1}\sigma_{n-1}^{k'j}\sigma_n^k|j\rangle \otimes |1\rangle \\
&= \frac{1}{\sqrt{2^{n-1}}}\sum_{j=0}^{2^{n-1}-1}\sigma_{n-1}^{k'j}|j\rangle \otimes \frac{1}{\sqrt{2}}\left(|0\rangle + \sigma_n^k|1\rangle\right) \\
&= F_{n-1}|k'\rangle \otimes \frac{1}{\sqrt{2}}\left(|0\rangle + \sigma_n^k|1\rangle\right)
\end{aligned}$$

con lo que queda demostrada la expresión recursiva de F_n.

Dado un número natural de n dígitos $k = k_n \ldots k_1$ llamamos $\overline{k}$ al número de n dígitos que se obtiene al invertir el orden de los dígitos de k, es decir $\overline{k} = k_1 \ldots k_n$. Claramente se verifica que

$$\cdots S(2, n-1)\, S(1, n)\, |k\rangle = |\overline{k}\rangle$$

Por otra parte, llamemos W_n a la transformación cuántica definida por el circuito de la figura 12, pero prescindiendo de las puertas Swap. Observemos las dos siguientes propiedades importantes:

(a) Aplicando reiteradamente la siguiente propiedad: dos puertas cuánticas que actúan sobre qubits diferentes conmutan, y reordenando las puertas cuánticas

$$\Lambda R_2(n, n-1),\ \ldots,\ \Lambda R_{n-1}(n, 2),\ \Lambda R_n(n, 1)$$

respecto a las posiciones en las que aparecen en la figura 12, y colocándolas al final del algoritmo, se puede comprobar que la transformación W_n se puede expresar recursivamente de la forma

$$W_n = H(n)\,\Lambda R_2(n, n-1)\,\cdots\,\Lambda R_{n-1}(n,2)\,\Lambda R_n(n,1)\,W_{n-1}$$

(b) La transformada cuántica de Fourier F_n se puede escribir como

$$F_n = W_n\,\cdots\,S(2,n-1)\,S(1,n)$$

Por lo tanto, por la propiedad vista anteriormente, para demostrar la corrección del algoritmo sólo es necesario probar que $W_n|\overline{k}\rangle = F_n|k\rangle$.

Dicha demostración se puede hacer por inducción sobre el número n de qubits:

1. Paso base: $n = 1$. En este caso $|\overline{k}\rangle = |k\rangle$ y, además, se cumple

$$W_1 = F_1 = H$$

2. Paso de inducción: supongamos por H.I. que $W_n|\overline{k}\rangle = F_n|k\rangle$ para un $n \geq 1$. Entonces

$$\begin{aligned} F_{n+1}|k\rangle &= F_n|k'\rangle \otimes \frac{1}{\sqrt{2}}\left(|0\rangle + \sigma_{n+1}^{k}|1\rangle\right) = W_n|\overline{k'}\rangle \otimes \frac{1}{\sqrt{2}}\left(|0\rangle + \sigma_{n+1}^{k}|1\rangle\right) \\ &= W_n|\overline{k'}\rangle \otimes \frac{1}{\sqrt{2}}\left(|0\rangle + \sigma_{n+1}^{2^n k_{n+1}+\cdots+2^0 k_1}|1\rangle\right) \\ &= W_n|\overline{k'}\rangle \otimes \frac{1}{\sqrt{2}}\left(|0\rangle + \sigma_{n+1}^{k_1}\cdots\sigma_2^{k_n}\sigma_1^{k_{n+1}}|1\rangle\right) \end{aligned}$$

Ahora observemos que los factores que afectan al último qubit $\sigma_{n+2-j}^{k_j}$ valen 1 si $k_j = 0$ y σ_{n+2-j} si $k_j = 1$, para $j = 1 \ldots n$. Por lo tanto, cada uno de ellos puede considerarse como el efecto de una transformación ΛR_{n+2-j} sobre el qubit $(n+1)$-ésimo, controlada por el j-ésimo.

Y por otra parte observemos que $\sigma_1^{k_{n+1}} = \pm 1$, según que el dígito k_{n+1} valga 0 ó 1, de donde se deduce que $\frac{1}{\sqrt{2}}\left(|0\rangle + \sigma_1^{k_{n+1}}|1\rangle\right) = H(n+1)|k_{n+1}\rangle$. Entonces se verifica que

$$\begin{aligned} F_{n+1}|k\rangle &= \Lambda R_2(n+1,n)\,\cdots\,\Lambda R_{n+1}(n+1,1)\left(W_n|\overline{k'}\rangle \otimes \frac{1}{\sqrt{2}}\left(|0\rangle + \sigma_1^{k_{n+1}}|1\rangle\right)\right) \\ &= H(n+1)\,\Lambda R_2(n+1,n)\,\cdots\,\Lambda R_{n+1}(n+1,1)\left(W_n|\overline{k'}\rangle \otimes |k_{n+1}\rangle\right) \\ &= W_{n+1}\left(|\overline{k'}\rangle \otimes |k_{n+1}\rangle\right) = W_{n+1}|\overline{k}\rangle \end{aligned}$$

sin más que tener en cuenta la expresión recursiva de W_n. □

6.3 Ejercicios

1. Hallar la transformada cuántica de Fourier del estado $\frac{1}{\sqrt{2}}(|010\rangle + |101\rangle)$.

2. Demostrar que F_n es una transformación unitaria.

3. Dadas dos funciones f, $g : \mathbb{Z}_{2^n} \to \mathbb{Z}_{2^n}$ se define el producto de convolución

$$(f * g)(k) = \sum_{j=0}^{2^n-1} f(k-j)g(j)$$

donde los índices se toman módulo 2^n. Calcular el producto de convolución de las funciones f y g definidas sobre $\mathbb{Z}_{16}$ del siguiente modo:

$$f(j) = \begin{cases} 1 & \text{si } j \bmod 4 = 0 \\ 0 & \text{en otro caso} \end{cases} \qquad g(j) = \begin{cases} 1 & \text{si } j \neq 5 \\ -1 & \text{si } j = 5 \end{cases}$$

4. Demostrar que $\widehat{(f * g)} = \hat{f}\hat{g}$.

7 Algoritmo de Shor

La factorización de números enteros es un problema muy actual. No porque hayamos encontrado un método eficiente para resolverlo sino más bien por todo lo contrario. La dificultad de este problema, infructuosamente atacado hasta el momento, nos permite considerar la multiplicación de números enteros como una función de dirección única. La aplicación más importante de este hecho es el sistema criptográfico de clave pública RSA [26], introducido por Rivest, Shamir y Adleman en 1978. La seguridad de este sistema radica en la imposibilidad práctica de factorizar números grandes, con centenares de dígitos.

El algoritmo clásico de factorización con mejor complejidad demostrada es el obtenido por Pollard y Strassen en 1976 que tiene complejidad $O(2^{n/4}n^2)$ [17], siendo n el número de dígitos. Sin embargo existen otros algoritmos más eficientes, aunque no se haya conseguido probar rigurosamente su complejidad. Por ejemplo, se conjetura que el algoritmo de Lenstra, Lenstra, Manasse y Pollard [23] tiene complejidad $e^{O\left(\sqrt[3]{n\log^2(n)}\right)}$.

7.1 Reducción del problema de factorización

Para encontrar un factor propio de un número N, impar y con al menos dos factores primos distintos, vamos a trabajar con el grupo multiplicativo de unidades de $\mathbb{Z}_N$, que denotaremos U_N. Se trata de un grupo conmutativo de cardinal $\phi(N)$, donde ϕ es la función de Euler. Un elemento $a \in \mathbb{Z}_N$ es una unidad, es decir pertenece a U_N, si y solo si $\mathrm{mcd}(a, N) = 1$. Con este planteamiento podemos reducir el problema de obtener un factor propio de N al de encontrar el orden de un elemento $a \in U_N$, es decir, al de calcular el menor número natural t tal que $a^t \equiv 1 \bmod N$.

Algoritmo 3 (Cálculo del orden de un elemento)

1. Elegir aleatoriamente a entre 1 y $N-1$, ambos incluidos.
2. Si $\mathrm{mcd}(a, N) \neq 1$ **devolver** $\mathrm{mcd}(a, N)$.
3. Calcular el orden t de a en U_N.
4. Si t es impar **devolver** fallo.
5. Si $\mathrm{mcd}(a^{t/2}+1, N) \neq N$ **devolver** $\mathrm{mcd}(a^{t/2}+1, N)$.
6. **Devolver** fallo.

Si el algoritmo no falla entonces devuelve un factor propio de N. Si termina en el paso 2 el resultado es obviamente un factor propio de N. Si termina en el paso 5 entonces se cumple que a es una unidad de orden par y $\mathrm{mcd}(a^{t/2}+1, N) \neq N$. En este caso obtenemos dos divisores de cero en $\mathbb{Z}_N$: $m_1 = a^{t/2} - 1$ y $m_2 = a^{t/2} + 1$. En efecto $m_1 \not\equiv 0 \bmod N$ pues en caso contrario $a^{t/2} \equiv 1 \bmod N$ y el orden de a no sería t, $m_2 \not\equiv 0 \bmod N$ ya que estamos suponiendo que $\mathrm{mcd}(m_2, N) \neq N$ y $m_1 m_2 = a^t - 1 \equiv 0 \bmod N$ puesto que t es el orden de a. Finalmente cada uno de los divisores de cero, y en particular m_2, contiene un factor propio de N.

El algoritmo propuesto para reducir el problema es un algoritmo probabilístico. Pero, para que ambos problemas sean polinomialmente equivalentes, debe funcionar con probabilidad mayor que una constante positiva independiente del dato de entrada N. Efectivamente, el algoritmo calcula un factor propio de N con probabilidad mayor que

$$1 - \left(\frac{1}{2}\right)^{h-1} \tag{2}$$

donde h es el número de factores primos distintos que tiene N. La demostración de este resultado se incluye en el apéndice A.

Antes de abordar cuánticamente el cálculo del orden de una unidad $a \in U_N$, vamos a modificar de nuevo el problema. Consideremos la función $f : \mathbb{Z} \to \mathbb{Z}_N$ definida por $f(k) = a^k \bmod N$. Se trata de una función periódica cuyo periodo T coincide con el orden t de a. En efecto:

1. $f(k+t) = a^{k+t} \bmod N = a^k a^t \bmod N = a^k \bmod N = f(k)$, por lo tanto $T \leq t$.
2. $a^T \bmod N = f(T) = f(0) = a^0 \bmod N = 1 \bmod N$, por lo tanto $t \leq T$.

De este modo el problema de encontrar un factor propio de N se ha reducido al de encontrar el periodo de la función $f(k) = a^k \bmod N$ para una unidad $a \in U_N$. Con este propósito Shor introduce en su algoritmo la transformada cuántica de Fourier (QFT), que le permite calcular la transformada discreta de Fourier (DFT) de forma mucho más eficiente que con los algoritmos clásicos conocidos.

7.2 Algoritmo de factorización

El algoritmo de Shor consiste en realizar cuánticamente la parte más costosa del algoritmo clásico descrito anteriormente (algoritmo 4). Realiza cuánticamente el cálculo del orden t de a en U_N (paso 3) o, equivalentemente, el cálculo del periodo T de la función $f(k) = a^k \bmod N$ definida sobre $\mathbb{Z}_{\phi(N)}$.

Ya hemos comentado que si $Q = \phi(N)$ no podemos calcular la transformada cuántica de Fourier. El motivo es evidente: no conocemos el valor de $\phi(N)$ y su cálculo es tan costoso como la factorización de N. Esto significa que no podemos tomar $\mathbb{Z}_{\phi(N)}$ como dominio de la función f. En su lugar consideramos el conjunto $\mathbb{Z}_Q$ donde $Q = 2^n$ es la única potencia de dos que verifica $N^2 < Q < 2N^2$. Generalmente f no será una función periódica en $\mathbb{Z}_Q$, puesto que T no tiene por qué ser un divisor de Q. Por tanto, vamos a trabajar con una extensión no periódica de la función f.

Para representar los valores que toma la función f necesitamos exactamente m qubits, siendo m el único número entero tal que $N < 2^m < 2N$, puesto que f toma valores en $\mathbb{Z}_N$. Por tanto, la parte cuántica del algoritmo de Shor trabajará con un $(n+m)$-qubit y con el siguiente operador unitario, asociado a la función f:

$$U_f\left(|k\rangle \otimes |j\rangle\right) = |k\rangle \otimes |j \oplus f(k)\rangle = |k\rangle \otimes |j \oplus (a^k \bmod N)\rangle, \quad 0 \leq k < Q \ \text{ y } \ 0 \leq j < 2^m$$

En el $(n+m)$-qubit se distinguen dos conjuntos de qubits o registros. El primero está formado por los n primeros y permite representar todos los elementos del dominio $\mathbb{Z}_Q$ de

la función. El segundo está constituido por los m últimos y en él se codifican los valores que toma la función en los distintos elementos del dominio. El papel de estos dos registros y del operador U_f se apreciará mejor viendo el funcionamiento del algoritmo.

Algoritmo 4 (Shor)

1. Elegir aleatoriamente a entre 1 y $N-1$, ambos incluidos.
2. Si $\text{mcd}(a,N) \neq 1$ **devolver** $\text{mcd}(a,N)$.
3. Calcular el periodo T de la función $f(k) = a^k \bmod N$ definida sobre $\mathbb{Z}_{\phi(N)}$:
   (a) Inicializar el (n,m)-qubit $\Psi = |0\rangle \otimes |0\rangle$.
   (b) Aplicar F_n al primer registro, es decir aplicar $F_n \otimes I$.
   (c) Aplicar el operador U_f asociado a la función f.
   (d) Aplicar nuevamente F_n al primer registro, es decir aplicar $F_n \otimes I$.
   (e) Obtener la medida $k \in \{\, 0, 1, \ldots, Q-1 \,\}$ del primer registro.
   (f) Si es posible, calcular T a partir de k y si no **devolver** fallo.
4. Si T es impar **devolver** fallo.
5. Si $\text{mcd}(a^{T/2}+1, N) \neq N$ **devolver** $\text{mcd}(a^{T/2}+1, N)$.
6. **Devolver** fallo.

Para ver el funcionamiento del algoritmo vamos a mostrar, a lo largo del paso 3, la evolución del $(n+m)$-qubit. Este seguimiento nos permitirá determinar posteriormente la probabilidad de que el algoritmo encuentre un factor propio de N.

$$\begin{aligned}
&\xrightarrow{(b)} \quad \frac{1}{\sqrt{Q}} \sum_{j=0}^{Q-1} |j\rangle \otimes |0\rangle \\
&\xrightarrow{(c)} \quad \frac{1}{\sqrt{Q}} \sum_{j=0}^{Q-1} |j\rangle \otimes |f(j)\rangle \\
&\xrightarrow{(d)} \quad \frac{1}{Q} \sum_{j=0}^{Q-1} \sum_{k=0}^{Q-1} \sigma_n^{jk} |k\rangle \otimes |f(j)\rangle = \frac{1}{Q} \sum_{k=0}^{Q-1} |k\rangle \otimes |A(k)\rangle \quad \left(A(k) = \sum_{j=0}^{Q-1} \sigma_n^{jk} |f(j)\rangle \right) \\
&\xrightarrow{(e)} \quad k \in \{\, 0, 1, \ldots, Q-1 \,\} \text{ con probabilidad } P(k) = \frac{\|A(k)\|^2}{Q^2}
\end{aligned}$$

La primera transformada de Fourier sirve exclusivamente para inicializar el $(n+m)$-qubit, poniéndolo como combinación lineal (superposición) de todos los elementos del dominio de f. Si tomásemos como dominio $\mathbb{Z}_{\phi(N)}$, sustituyendo Q por $\phi(N)$ y σ_n por $e^{\frac{2\pi i}{\phi(N)}}$, la probabilidad de obtener k al medir el primer registro verificaría

$$P(k) = \frac{\|A(k)\|^2}{\phi^2(N)} = \begin{cases} \frac{1}{T} & \text{si } k \text{ es múltiplo de } w \\ 0 & \text{si } k \text{ no es múltiplo de } w \end{cases}$$

siendo w la frecuencia definida por la expresión $wT = \phi(N)$. Si $Q = 2^n$ la función f no tiene por qué ser periódica en el dominio $\mathbb{Z}_Q$ pues, generalmente, T no será divisor de Q.

Sin embargo, la transformada de Fourier tiene picos muy pronunciados en los elementos del dominio próximos a los valores jw', $0 \leq j < T$, siendo $w' = Q/T$ el valor formal de la frecuencia en la extensión no periódica de f. En la siguiente figura se observa el efecto característico de dicha extensión: reducción de la altura y esanchamiento de los picos de la transformada de Fourier. A pesar de este efecto, los valores próximos a los múltiplos de la frecuencia formal acumulan una parte importante de la probabilidad.

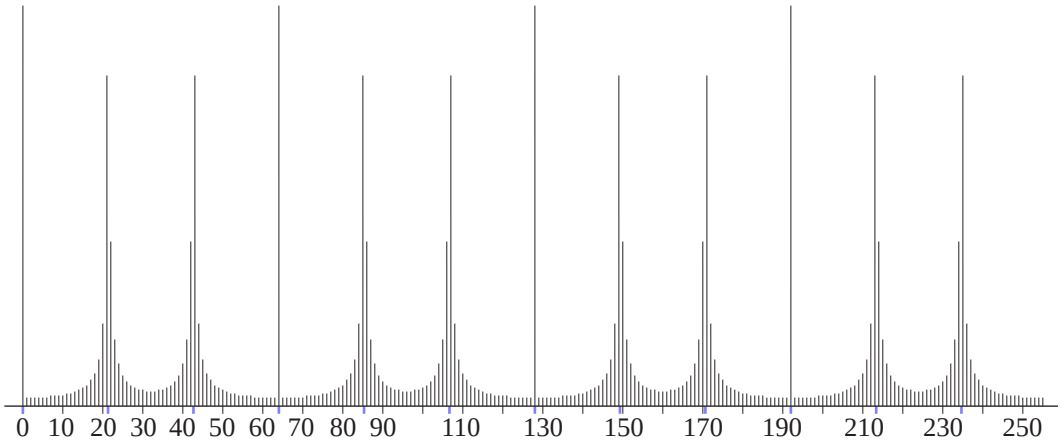


Figure 13: Gráfica de $\sqrt{P(k)}$ para $N = 13$, $a = 2$ ($T = 12$), $n = 8$ ($Q = 256$) y $w' = 21.33$

7.3 Análisis de la probabilidad de éxito

En el apéndice B se demuestra que la probabilidad de que el resultado k de la medida sea el entero más cercano a un múltiplo dado de la frecuencia formal jw' ($0 \leq j < T$), es decir que para dicho valor de j se verifique $|k - jw'| \leq 1/2$, satisface

$$P(k)\,T = \frac{\|A(k)\|^2}{Q^2}\,T \geq \frac{4}{\pi^2}\left(1 - \frac{1}{N}\right)^2 \frac{\cos^2\left(\frac{\pi}{2N}\right)}{\left(1 + \frac{1}{N}\right)^2} \quad \xrightarrow[N\to\infty]{} \quad \frac{4}{\pi^2} \tag{3}$$

Entonces, la probabilidad de obtener el entero más cercano a uno cualquiera de los múltiplos de la frecuencia formal tiene una cota inferior que se aproxima, cuando N tiende a infinito, a $4/\pi^2$; una constante independiente de N. Esto significa que el efecto de la extensión no periódica de f es irrelevante. El motivo por el que estamos especialmente interesados en esos valores de k es que permiten obtener el periodo T de la función f con probabilidad alta. En efecto, teniendo en cuenta que $T < N$ y $N^2 < Q$, se concluye que

$$|k - jw'| \leq 1/2 \implies |Tk - jQ| \leq \frac{T}{2} \implies \left|\frac{k}{Q} - \frac{j}{T}\right| \leq \frac{1}{2Q} < \frac{1}{2T^2}$$

Finalmente, aplicando un teorema clásico de teoría de fracciones continuas (véase el apéndice C), se concluye que j/T es una de las convergentes de la fracción continua del número k/Q, del que conocemos tanto el numerador como el denominador. Si, además, se cumpliese que $\mathrm{mcd}(j, T) = 1$ obtendríamos el periodo T de la función f sin más que probar con los denominadores de todas las convergentes. Puesto que existe una biyección

entre los valores de j y los valores de k y hay $\phi(T)$ valores de j que son primos relativos con T, entonces exiten $\phi(T)$ valores de k que permiten calcular el periodo de la función. Teniendo en cuenta este hecho y los resultados expresados por las fórmulas 2 y 3, la probabilidad de éxito del algoritmo de Shor, P, verifica:

$$\begin{aligned} P\,\mathrm{loglog}(T) &\geq \left(1-\left(\frac{1}{2}\right)^{h-1}\right)\frac{4}{\pi^2}\frac{\phi(T)}{T}\,\mathrm{loglog}(T)\left(1-\frac{1}{N}\right)^2\frac{\cos^2\left(\frac{\pi}{2N}\right)}{\left(1+\frac{1}{N}\right)^2} \\ &\geq \frac{2}{\pi^2}\frac{\phi(T)}{T}\,\mathrm{loglog}(T)\left(1-\frac{1}{N}\right)^2\frac{\cos^2\left(\frac{\pi}{2N}\right)}{\left(1+\frac{1}{N}\right)^2} \end{aligned}$$

En la segunda desigualdad hemos utilizado la hipótesis de que el número h de divisores primos de N distintos verifica $h \geq 2$. Tomando límites en la expresión anterior, tanto en N como en T, obtenemos el resultado clave del algoritmo de Shor: la probabilidad de éxito P verifica

$$P\,\mathrm{loglog}(T) \quad \xrightarrow[N\to\infty]{} \quad \frac{2}{\pi^2}\frac{\phi(T)}{T}\,\mathrm{loglog}(T) \quad \xrightarrow[\liminf]{} \quad \frac{2}{\pi^2}e^{-\gamma}$$

Para conseguir una probabilidad de éxito independiente de N y de T es suficiente repetir el algoritmo $O(\mathrm{loglog}(N))$ veces. Además, veremos un poco más adelante que el número de operaciones que realiza el algoritmo es polinomial respecto del número de dígitos de N. Por tanto, el algoritmo de Shor realiza uno de los sueños más antiguos de las matemáticas: factorizar números enteros de forma eficiente.

A modo de ejemplo, la siguiente figura muestra la probabilidad de éxito del algoritmo de Shor para todos los números compuestos e impares entre 9 y 255. Para calcular estas probabilidades se ha simulado la ejecución del algoritmo de Shor en un ordenador clásico utilizando números complejos en coma flotante con doble precisión.

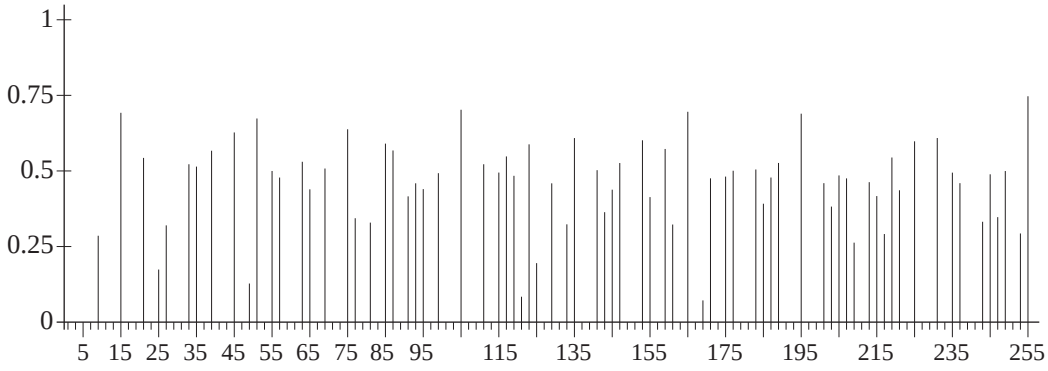


Figure 14: Probabilidad de éxito del algoritmo de Shor

7.4 Análisis de la complejidad

Una vez demostrado que la probabilidad de éxito del algoritmo de Shor es razonablemente alta, sólo queda por determinar el número de operaciones que requiere su ejecución. Las complejidades de las operaciones que necesitamos, tanto clásicas como cuánticas, verifican:

1. Aritmética básica para dos números naturales A y B tales que $0 < A, B < 2N^2$:
   (a) Expresión booleana $A = B$ o $A \neq B$: $O(\log(N))$.
   (b) Suma $A + B$ o diferencia $A - B$: $O(\log(N))$.
   (c) Producto $A \cdot B$, cociente A div B o módulo A mod B: $O(\log^2(N))$.
   (d) Máximo común divisor $\mathrm{mcd}(A, B)$: $O(\log^3(N))$.
   (e) Exponenciación modular $A^B \bmod N$: $O(\log^3(N))$.
   (f) Expresión booleana $A^B = 1 \bmod N$: $O(\log^3(N))$.
2. Elección aleatoria de un número entre 1 y $N - 1$: $O(\log(N))$.
3. Determinar si el denominador de alguna convergente de k/Q verifica $a^x = 1 \bmod N$: $O(\log^4(N))$.
4. Transformada cuántica de Fourier F_n: $O(\log^2(N))$
5. Exponenciación modular cuántica U_f: $O(\log^3(N))$.
6. Medida cuántica del primer registro: $O(\log(N))$.

Entre estos resultados hay dos que requieren una explicación más detallada: puntos 3 y 5. Respecto al primero de ellos, para determinar la complejidad basta observar que el tamaño de la fracción continua de k/Q es $O(\log(N))$ (véase el apéndice C). Además, no es preciso determinar si el periodo T de la función coincide con el denominador de alguna de las convergentes puesto que, en tal caso, deberíamos resolver un problema tan complicado como el problema original de factorizar N. En realidad, basta con verificar si algún denominador cumple la ecuación $a^x = 1 \bmod N$. En cuanto al segundo punto, la exponenciación modular cuántica se puede implementar a partir de algoritmos cuánticos de aritmética entera y modular [31] que, esencialmente, imitan los algoritmos clásicos.

Finalmente, para obtener una probabilidad de éxito próxima a $2e^{-\gamma}/\pi^2 = 0.1137...$, o mayor, habría que repetir el algoritmo $\mathrm{loglog}(N)$ veces. Por tanto, extrictamente hablando, la complejidad del algoritmo de Shor es $O(\log^4(N)\mathrm{loglog}(N))$. Se trata de un algoritmo probabilístico polinomial para factorizar números enteros que, además, se puede adaptar fácilmente para resolver el problema del logaritmo discreto. Estos resultados tienen una gran trascendencia para las matemáticas, especialmente en teoría de números y en criptografía de clave pública.

7.5 Ejercicios

1. Sean k y n números naturales. Obtener un algoritmo (clásico) polinomial que determine si la ecuación $n = x^k$ tiene solución entera.
2. Usando el resultado anterior, obtener un algoritmo polinomial que determine si un número natural n es una potencia no trivial de otro.
3. Probar que $|e^{ix} - 1|^2 = 4\sin^2\left(\frac{x}{2}\right)$.

4. Demostrar que

$$\left| \sum_{q=0}^{s} e^{\frac{2\pi i p q r}{m}} \right|^2 = \frac{\sin^2\left(\frac{\pi p r (s+1)}{m}\right)}{\sin^2\left(\frac{\pi p r}{m}\right)}$$

5. Demostrar que $\sin^2\left(\frac{\pi}{2}(1+x)\right) \geq 1 - \left(\frac{\pi}{2}x\right)^2$.

8 Criptografía cuántica

En esta sección veremos una de las aplicaciones más importantes de la teoría de la información cuántica. Como hemos visto el algoritmo de Shor abre la posibilidad de que los ordenadores cuánticos puedan romper los criptosistemas de seguridad de clave pública. Afortunadamente la criptografía cuántica proporciona un sistema seguro de distribución de claves privadas.

Uno de los problemas de mayor dificultad práctica a la hora de llevar a cabo una comunicación segura mediante un sistema de clave privada es la distribución y almacenamiento de las claves. Un resultado de Shannon de 1949 establece que si la clave es aleatoria, de la misma longitud que el mensaje a cifrar y se usa una única vez, la codificación es segura. De hecho es el único sistema cuya seguridad ha sido probada. Sin embargo, la necesidad de distribuir y almacenar de manera segura las claves en general largas y de un solo uso, limita claramente las posibilidades de este sistema.

Una de las razones del éxito obtenido por los sistemas de clave pública es que permiten prescindir de acordar y distribuir la clave secreta. Sin embargo la seguridad de estos sistemas nunca ha sido probada matemáticamente. No se sabe si factorizar un número puede hacerse en tiempo polinómico, simplemente no se ha encontrado un algoritmo que lo haga, de modo que la seguridad práctica del sistema RSA viene proporcionada por el coste computacional de la factorización. La construcción de un ordenador cuántico en el que se implemente el algoritmo de Shor, que permite calcular logaritmos discretos y factorizar en tiempo polinómico, supondría claramente la fractura definitiva de los sistemas de clave pública más importantes.

Las leyes de la mecánica cuántica, sin embargo, proporcionan herramientas para abordar el problema de la distribución segura de claves privadas. Esencialmente consiste en que los comunicantes se transmiten la clave privada a través de un canal cuántico (por ejemplo un cable de fibra óptica, si la clave se codifica mediante estados de polarización de fotones) y la aportación cuántica a la seguridad del proceso es que un espía no puede extraer información sin revelar su presencia a los comunicantes, ya que al medir estados cuánticos éstos se modifican.

Existen diferentes protocolos cuánticos, ideados con el fin de obtener claves privadas de un solo uso, que se pueden usar en sistemas simétricos de seguridad. De hecho, la criptografía cuántica es la primera aplicación comercial de la mecánica cuántica. Ya se comercializa actualmente un dispositivo que permite la distribución cuántica de claves.

8.1 Protocolos de distribución de claves

En un proceso de distribución cuántica de claves intervienen un emisor y un receptor, comúnmente denominados Alicia y Benito, un espía, Eva, y dos canales de comunicación, uno cuántico para enviar fotones y otro clásico para reconciliar y depurar información. Eva puede acceder al canal clásico y también puede acceder al canal cuántico y usar todos los medios que desee con la única restricción de que sean compatibles con las leyes de la mecánica cuántica. Los dos comunicantes legítimos usan un trozo de su clave para detectar la presencia de espías.

Los estados de polarización de un fotón se pueden usar para diseñar un sistema criptográfico cuántico de un solo uso.

En el espacio de Hilbert $\mathcal{H}_1$ consideraremos las bases ortogonales: $B_1 = [|0\rangle, |1\rangle]$, que se identifica con la polarización horizontal y vertical, y $B_\times = [|+\rangle, |-\rangle]$ identificada con la polarización 45^o y -45^o.

Recordemos que $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ y que si se mide el estado $|+\rangle$ respecto de B_1 se obtendrá el estado $|0\rangle$ con probabilidad 1/2 o el estado $|1\rangle$ con la misma probabilidad. Lo mismo ocurre al medir $|-\rangle$ respecto de B_1. Por otra parte si se mide el estado $|0\rangle$, o el estado $|1\rangle$, respecto de $B_\times$ se obtendrá el estado $|+\rangle$ con probabilidad 1/2 o el estado $|-\rangle$ con la misma probabilidad.

En los protocolos cuánticos de generación de claves la idea será enviar una cadena de qubits, usando para codificarla fotones polarizados.

8.1.1 Protocolo BB84

Uno de los esquemas más sencillos para la generación y distribución de una clave aleatoria es el protocolo BB84, propuesto en 1984 por Bennett y Brassard [5], que se puede describir del siguiente modo:

Paso 1: Alicia utiliza una moneda para generar una cadena aleatoria de ceros y unos.

Paso 2: Para cada bit de la cadena, Alicia elige aleatoriamente una de las dos bases B_1 o $B_\times$ y envía a Benito, por un canal cuántico, el qubit correspondiente de acuerdo con el siguiente alfabeto:

Si elige B_1:	$0 \to \|0\rangle$	(pol. horizontal),	$1 \to \|1\rangle$	(pol. vertical).
Si elige $B_\times$:	$0 \to \|+\rangle$	(pol. 45^o),	$1 \to \|-\rangle$	(pol. -45^o).

Paso 3: Cuando Benito recibe cada fotón, no tiene modo de saber con qué alfabeto ha sido polarizado, así que él mide eligiendo, también aleatoriamente, para cada uno de ellos la base B_1 o $B_\times$.

Aproximadamente la mitad de las veces Benito elegirá el mismo alfabeto que Alicia y la otra mitad elegirá la base contraria a la utilizada por ella. Por tanto después de todas las mediciones Benito tendrá su secuencia binaria, que coincidirá con la de Alicia en un 75%.

Paso 4: Para localizar y eliminar los bits en que las mediciones se han realizado en la base inadecuada, se realiza el proceso de contraste de información, denominado *siftting* o *reconciliación de bases*. Benito comunica a Alicia, por un canal público, qué alfabeto ha usado en cada medición. Como respuesta, por el mismo canal, Alicia le comunica las posiciones en las que la medición se ha realizado con el alfabeto correcto.

Paso 5: Alicia y Benito borran de sus cadenas los bits en los que se han usado alfabetos diferentes y tienen así su palabra clave, que denominaremos clave bruta.

Si no ha habido ruido ni interferencia de espías, Alicia y Benito tienen una clave común, cuya longitud será aproximadamente la mitad de la de la cadena inicial.

Si ha habido espías, veremos que, si la fuente utilizada para emitir fotones, los emite de modo individual, cualquier estrategia de espionaje va a introducir discrepancias en la clave de Benito. Por ello, para detectar la presencia de espías, Alicia y Benito comparan sobre el canal público determinadas posiciones, aleatoriamente elegidas de sus claves brutas. Si quieren una clave de longitud n, llevan a cabo el protocolo para obtener una cadena de longitud $2n$, de los que usan n para chequear los errores. Es conocido (ver [24]) que, para todo $\delta > 0$, la probabilidad de que en los n bits testados haya menos de δn errores y en los restantes más de $(\delta + \varepsilon)n$ es menor que $2^{-O(\varepsilon^2 n)}$. Por lo tanto, si tras el chequeo de n bits no hallan discrepancias, pueden suponer que no ha habido espías y se quedan con la clave bruta como clave final. Si las no coincidencias superan una cota prefijada, abortan el protocolo y empiezan de nuevo.

8.1.2 Protocolo B92

Una generalización muy simple del protocolo BB84 fue propuesta por Bennet en 1992 [3] y se conoce como B92. La idea es que Alicia no utilice cuatro sino dos estados para codificar. El protocolo se puede describir como sigue:

Paso 1: Alicia genera una cadena aleatoria de ceros y unos.

Paso 2: Alicia codifica cada bit a de la cadena, usando el siguiente alfabeto:

$$a = 0 \to |0\rangle, \quad a = 1 \to |+\rangle$$

y envía a Benito el qubit resultante.

Paso 3: Benito genera una cadena aleatoria a' de ceros y unos.

Paso 4: Benito mide cada qubit recibido, usando B_1 si en la posición correspondiente $a' = 0$ y usando $B_\times$ cuando $a' = 1$.

Paso 5: De esta medición Benito obtiene una cadena b de ceros y unos con el siguiente criterio:

Si ha elegido B_1, pone $b = 0$ si obtiene $|0\rangle$ y $b = 1$ si obtiene $|1\rangle$.

Si ha elegido $B_\times$, pone $b = 0$ si obtiene $|+\rangle$ y $b = 1$ si obtiene $|-\rangle$.

Paso 6: Benito publica las posiciones en que $b = 1$ y, considerando solo esas posiciones, las claves son a para Alicia y $1 \oplus a'$ para Benito.

Proposición 2 *Con la notación anterior, si no hay espías ni ruido, se obtienen claves coincidentes con probabilidad 1. El valor esperado de la longitud de la clave es la cuarta parte de la de la cadena inicial.*

Demostración. En primer lugar notemos que $P(b = 1) = 1/4$. En efecto, si $a = 0$ (con probabilidad $1/2$), Alicia envía $|0\rangle$ y b sólo puede ser 1, con probabilidad $1/2$, en el caso de que $a' = 1$, lo que a su vez ocurre con probabilidad $1/2$.

Por otra parte si $a = 1$ (de nuevo con probabilidad $1/2$), Alicia envía $|+\rangle$ y b sólo puede ser 1, con probabilidad $1/2$, en el caso de que $a' = 0$, lo que a su vez ocurre con probabilidad $1/2$.

Observemos que en cualquiera de los dos casos, cuando se ha obtenido $b = 1$, es $a = 1 - a'$. Luego $P(a = 1 \oplus a'/b = 1) = 1$. □

8.1.3 Protocolo basado en pares EPR

Esta generalización del BB84 tiene especial interés conceptual, histórico y práctico. Fue propuesta por Ekert en el año 1991 [14] y basa su seguridad en el uso de pares EPR.

Un par EPR es un estado cuántico de 2 qubits entrelazados, por ejemplo de la forma

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

La idea consiste en reemplazar el canal cuántico usado para enviar qubits de Alicia a Benito por un emisor que envía dos qubits entrelazados, uno para Alicia y otro para Benito.

Partimos pues de la hipótesis de que Alicia y Benito comparten un conjunto de $2n$ pares EPR y seleccionan aleatoriamente n de ellos, para comprobar su grado de correlación que teóricamente debe ser perfecta. Si es así, tienen garantizado que los restantes son estados cuánticos entrelazados suficientemente puros y realizan con ellos el protocolo.

Alicia prepara una cadena aleatoria b de ceros y unos y de acuerdo con ella, mide su qubit de cada par con las bases B_1 o $B_\times$ obteniendo una cadena a. Benito hace lo mismo y obtiene una cadena a' con una cadena b' de elección de bases. Después comunican sus cadenas b y b' por un canal público y se quedan sólo con aquellos en los que las bases utilizadas coinciden. Si los pares EPR eran estados entrelazados puros las claves deben ser coincidentes. En este caso, las leyes de la mecánica cuántica garantizan que Eva no puede tener información sobre esta clave.

8.1.4 Protocolo SARG04

El protocolo BB84 ha sido el más probado experimentalmente y se ha visto que, usando pulsos laser atenuados hay una probabilidad significativa de emitir pulsos que contengan más de un fotón. Para estos pulsos, Eva puede hacer lo que se llama un ataque PNS (photon number splitting) que consiste básicamente en quedarse con uno de los fotones y acceder a la información sin modificar el fotón que llega a Benito. Scarani, Acín, Ribordy y Gisin (2004 [27]), han visto que, usando los mismos cuatro estados, pero con una nueva forma de codificar y contrastar la información, se puede definir un protocolo más robusto frente a este tipo de ataques. Este protocolo, el SARG04, se puede llevar a cabo con la misma tecnología desarrollada para el BB84 y se describe del siguiente modo:

Paso 1: Alicia genera una cadena aleatoria de ceros y unos.

Paso 2: Alicia codifica cada bit de la cadena, usando aleatoriamente uno de los dos estados de la base B_1 para el 0 y uno de los de la base $B_\times$ para el 1. Es decir, si tiene un 0 envía $|0\rangle$ ó $|1\rangle$ y si tiene un 1 envía $|+\rangle$ o $|-\rangle$.

Paso 3: Cuando Benito recibe cada fotón, lo mide eligiendo aleatoriamente para cada uno de ellos la base B_1 o $B_\times$.

Paso 4: En la fase de contraste de información, Alicia no dice qué base ha usado (ya que sería como decir cuál es el bit), sino que le facilitará a Benito uno de los siguientes conjuntos en el que se encuentre el qubit que ella ha enviado junto con uno de los que se usan para codificar el otro bit:

$$S_{++} = \{|0\rangle, |+\rangle\}, SS_{+-} = \{|0\rangle, |-\rangle\}, S_{-+} = \{|1\rangle, |+\rangle\}, S_{--} = \{|1\rangle, |-\rangle\}$$

Paso 5:

Caso a) Alicia dice S_{++}.

Si Benito ha medido con $B_\times$ y ha obtenido $|+\rangle$ descarta el bit, si ha obtenido $|-\rangle$ pone un 0 (ya que sabe que Alicia no puede haber enviado $|+\rangle$ y tiene que haber enviado $|0\rangle$).

Si Benito ha medido con B_1 y ha obtenido $|0\rangle$ descarta el bit, si ha obtenido $|1\rangle$ pone un 1 (ya que sabe que Alicia no ha enviado $|0\rangle$).

Caso b) Alicia dice S_{+-}.

Si Benito ha medido con $B_\times$ y ha obtenido $|+\rangle$ pone un 0 , si ha obtenido $|-\rangle$ descarta el bit.

Si Benito ha medido con B_1 y ha obtenido $|0\rangle$ descarta el bit, si ha obtenido $|1\rangle$ pone un 1.

Así sucesivamente, en la siguiente tabla se muestra la actuación de Benito indicando lo que pone en su cadena en función de lo que le dice Alicia y del resultado de su medición (D significa que descarta el bit).

	$\|0\rangle$	$\|1\rangle$	$\|+\rangle$	$\|-\rangle$
S_{++}	D	1	D	0
S_{+-}	D	1	0	D
S_{-+}	1	D	D	0
S_{--}	1	D	0	D

A continuación, comunica a Alicia las posiciones descartadas.

Para determinar la longitud de la clave bruta resultante, hay que tener en cuenta que hay casos de probabilidad nula. Por ejemplo si Alicia ha enviado $|0\rangle$ y Benito mide con B_1 no puede obtener $|1\rangle$ en la medición. Por tanto, en la fila correspondiente a S_{++}, en este caso, la probabilidad de poner 1 es 0 y la de descartar el bit es 3/4. Razonando de este modo se obtiene que, si no ha habido espías, tendrán una clave común, cuya longitud será la cuarta parte de la de la cadena inicial.

8.2 Seguridad de los protocolos criptográficos cuánticos

Como se ha comentado antes, la seguridad de los protocolos cuánticos de distribución de claves se basa en que los comunicantes legítimos pueden detectar la presencia de un espía porque, al medir los estados cuánticos, éste introduce discrepancias en la clave.

Así, pueden chequear un trozo de la clave generada y rechazarla, en caso de encontrar discrepancias. Pero, en un sistema práctico de distribución cuántica de claves, pueden aparecer errores, debidos a imperfecciones técnicas del emisor o del receptor y no a la presencia de un espía. Por ello es importante determinar una cota de errores admisibles por encima de la cual se debe abortar el proceso y para ello es necesario analizar los efectos de las posibles estrategias de espionaje sobre el protocolo a utilizar.

Cuando la tasa de error esté por debajo de la cota, se admitirá la clave generada y se llevará a cabo un proceso de depuración en el que los errores se pueden detectar y corregir con códigos clásicos sobre un canal clásico de comunicación.

8.2.1 Estrategias de ataque individual al protocolo BB84

Como primera opción, vamos a suponer que Eva ha espiado la comunicación y que actúa del siguiente modo: Intercepta el fotón enviado por Alicia, elige aleatoriamente una de las dos bases, mide el qubit recibido, guarda el resultado de la medición y envía a Benito el qubit resultante.

Esta estrategia de espionaje (Interceptar-Reenviar) es de las denominadas de *ataque individual* ya que Eva ataca independientemente cada qubit. Más adelante comentaremos otras posibles estrategias de este tipo.

Alicia y Benito pueden detectar la presencia de Eva, ya que ésta habrá modificado el estado de algunos de los qubits que envía a Benito.

Para hacer el estudio de la seguridad del protocolo BB84 analizaremos las probabilidades sobre cada bit de la clave.

Denominamos *discrepancia* y denotamos por d, a la probabilidad de que un bit de la clave de Benito no coincida con el correspondiente de Alicia.

Alicia y Benito van a utilizar una parte de sus claves para estimar la discrepancia y han de decidir previamente una cota admisible por debajo de la cual consideran que el protocolo es seguro y que los errores son debidos a ruidos o interferencias y no a la presencia de espías.

Proposición 3 *Con la estrategia Interceptar-Reenviar, la cadena de Eva tendrá el 75% de aciertos y la discrepancia introducida en la clave de Benito será del 25%.*

Demostración. Sea a el bit de Alicia, a'' el de Benito y B la base elegida por ambos. Denominamos B' la base elegida por Eva y a' el bit resultante tras su medición. Entonces las probabilidad de coincidencia entre los bits de Alicia y Eva es:

$$P(a = a') = \frac{1}{2}P(a = a'/B = B') + \frac{1}{2}P(a = a'/B \neq B') = \frac{3}{4}$$

Por otra parte, la probabilidad de coincidencia de los bits de Alicia y Benito tras la reconciliación de bases $P(a = a'')$ es:

$$P(a = a''/B = B')P(B = B') + P(a = a''/B \neq B')P(B \neq B') = 1 \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4}$$

□

Eva puede usar otras estrategias de espionaje, con el objetivo de aumentar su probabilidad de acierto y/o disminuir la discrepancia introducida en la clave de Benito. Por ejemplo, en lugar de elegir aleatoriamente la base, puede medir siempre con una base que le maximice la probabilidad de acierto. Puede considerar una base ortonormal $B = [|u\rangle, |v\rangle]$, de la forma

$$|u\rangle = \cos(\alpha)|0\rangle + \sin(\alpha)|1\rangle, \quad |v\rangle = -\sin(\alpha)|0\rangle + \cos(\alpha)|1\rangle$$

y usarla para medir y descodificar su clave, poniendo 0 si obtiene $|u\rangle$, y 1 si obtiene $|v\rangle$.

En este caso, la expresión respecto de B de los vectores de las bases B_1 y $B_\times$ es

$$\begin{pmatrix} |0\rangle \\ |1\rangle \end{pmatrix} = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} |u\rangle \\ |v\rangle \end{pmatrix}$$

$$\begin{pmatrix} |+\rangle \\ |-\rangle \end{pmatrix} = \begin{pmatrix} \cos(\frac{\pi}{4} - \alpha) & -\sin(\frac{\pi}{4} - \alpha) \\ \sin(\frac{\pi}{4} - \alpha) & \cos(\frac{\pi}{4} - \alpha) \end{pmatrix} \begin{pmatrix} |u\rangle \\ |v\rangle \end{pmatrix}$$

Con esta estrategia, la probabilidad de acierto de Eva es

$$p(\alpha) = \frac{1}{2}\left(\cos^2(\alpha) + \cos^2(\frac{\pi}{4} - \alpha)\right),$$

que se hace máxima máxima cuando $\alpha = \dfrac{\pi}{4} - \alpha$, en cuyo caso $\alpha = \pi/8$.

La discrepancia es constante e igual a $1/4$.

Luego, de acuerdo con el plantemiento de espionaje propuesto, la base que debe elegir Eva es $B_i = [|u\rangle, |v\rangle]$, con

$$|u\rangle = \cos(\pi/8)|0\rangle + \text{sen}(\pi/8)|1\rangle \quad \text{y} \quad |v\rangle = -\,\text{sen}(\pi/8)|0\rangle + \cos(\pi/8)|1\rangle,$$

que es la denominada base intermedia.

En este caso, la probabilidad de acierto de Eva es

$$p = \cos^2(\pi/8) = \frac{1}{2} + \frac{\sqrt{2}}{4} \approx 0.854.$$

Ante diferentes estrategias de espionaje, parece razonable considerar mejor aquélla que permita a Eva obtener mayor probabilidad de acierto y menor discrepancia. Diremos que una estrategia de espionaje, E_1, es mejor que otra, E_2, si para la primera es mayor el valor de $p + f$, donde p es la probabilidad de acierto de Eva y f es la fidelidad de la clave de Benito (que será $f = 1 - d$).

Con esta idea, la estrategia de la base intermedia es mejor que la de Interceptar-Reenviar. Para la primera es $p + f = (5 + \sqrt{2})/4 \approx 1'603$, mientras que para la segunda este valor es $1'5$.

Pero, además de las consideradas, Eva puede llevar a cabo otras estrategias, ya que puede hacer con los qubits interceptados cualquier manipulación que no viole las leyes de la mecánica cuántica.

Con la hipótesis de ataque individual, podemos establecer, para el protocolo BB84, una estrategia general de espionaje en los siguientes términos:

1. Eva intercepta el qubit $|x\rangle$ emitido por Alicia.
2. Añade un n-qubit de prueba en estado $|0\rangle \in \mathcal{H}_n$ y le aplica una transformación unitaria arbitraria T, con lo que obtiene el estado $T(|x\rangle \otimes |0\rangle)$.
3. Envía a Benito el primer qubit de dicho estado.
4. Espera a la reconciliación de bases y mide uno de sus qubits para obtener su clave.

Se puede demostrar que, en este caso, la máxima información de Eva verifica

$$p(d) = \frac{1}{2} + \sqrt{d(1-d)}$$

Así, se establece una cota admisible de discrepancia (ver figura 15), por debajo de la cual conocemos la posible información de Eva y cuánto mayor que ésta es la probabilidad $(1-d)$ de coincidencia entre las claves de Alicia y Benito.

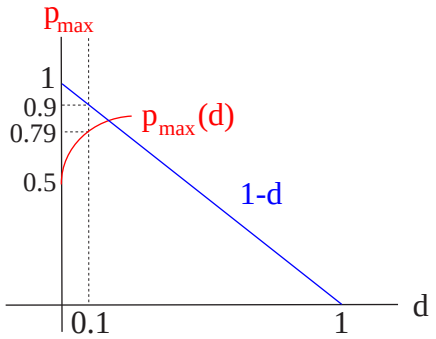


Figure 15: Gráficas de $p(d)$ y de $1-d$

De este modo, o bien se aborta el protocolo o bien éste proporciona una clave en la que la información de Eva está por debajo de una cota previamente establecida.

8.3 Corrección de errores y amplificación de la privacidad

Una vez que se tiene una clave y se sabe que contiene un número de errores menor que una cota establecida de antemano, hay varias técnicas clásicas que permiten corregir dichos errores y reducir la información del espía.

Por ejemplo, Alicia y Benito pueden localizar los errores usando técnicas de control de paridad en bloques o bien usar un código corrector adecuado para depurar la clave. Supongamos que el número máximo de errores por bloque es t, entonces es posible usar un

código (clásico o cuántico) con distancia $2t+1$ para obtener una nueva clave sin errores. La idea consiste en que Alicia elige aleatoriamente una palabra x del código y le envía a Benito la suma de su cadena con x. Cuando Benito la recibe, primero le suma su cadena y después descodifica el resultado obteniendo la palabra x.

A la clave resultante la llaman clave reconciliada. Pero esta clave sólo es parcialmente secreta porque no se reduce la información del espía. Incluso, dado que la comunicación se hace por un canal público, puede ocurrir que la estrategia de corrección de errores utilizada permita aumentar la información del espía. Por ello es necesario pasar a la fase conocida como *amplificación de la privacidad*.

La amplificación de privacidad se podría definir como el arte de conseguir una clave altamente secreta a partir de otra clave, de mayor longitud que es sólo parcialmente secreta. Para llevar a cabo esta fase los comunicantes usarán el canal público, y por tanto el espía tiene acceso (pasivo) al mismo.

Un método muy sencillo de amplificación de la privacidad es el siguiente:

Alicia elige aleatoriamente pares de bits y anuncia su elección (diciendo solamente las posiciones, por ejemplo bits 23 y 47). Alicia y Benito hacen la suma módulo 2 para cada una de las parejas y sustituyen los dos bits por el resultado obtenido.

De esta forma acortan su clave pero disminuyen la información de Eva, ya que si ésta conoce sólo uno de los bits de la pareja, no puede obtener el valor de la suma. Por otra parte si la probabilidad de que Eva conozca cada bit es $p > 1/2$, la probabilidad de acertar la paridad es $p^2 + (1-p)^2$ que es menor que p. Por ejemplo, si $p = 0.6$, se tiene $0.6^2 + 0.4^2 = 0.52$. Con bloques de mayor longitud disminuye más la probabilidad de acierto de Eva, pero la clave se acorta más.

Una técnica bastante generalizada en amplificación de privacidad es el uso de funciones de hash. Una clase $\mathcal{G}$ de funciones de hash $g : \{0,1\}^n \to \{0,1\}^r$ es *universal* si dada g elegida uniformemente en $\mathcal{G}$, para todo $x \neq y$ la probabilidad de que $g(x) = g(y)$ está acotada por el inverso del cardinal del conjunto imagen, esto es $p(g(x) = g(y)) \leq \dfrac{1}{2^r}$.

Por ejemplo, la clase de funciones g_a, definidas para cada $a = (a_1 \dots a_n) \in \{0,1\}^n$, por $g(x_1 \dots x_n) = (x_1 \oplus a_1, \dots, x_r \oplus a_r)$ es una clase universal de funciones de hash.

8.4 Ejercicios

1. Estudiar el efecto del espionaje de Eva en el protocolo B92 si utiliza la estrategia Interceptar-Reenviar.

2. Calcular la probabilidad de acierto de Eva $p(\alpha)$ y la discrepancia d que introduce en el protocolo BB84 si mide en la base $B = [|u\rangle, |v\rangle]$.

9 Apéndices

9.1 Apéndice A

Sean un número N, compuesto e impar, y un elemento $a \in \mathbb{Z}_N \setminus \{0\}$ elegido al azar. Vamos a probar que la probabilidad de que a pertenezca a la unión de los conjuntos disjuntos

$$S_1 = \{\, a \mid \text{mcd}(a, N) \neq 1 \,\} \quad \text{y} \quad S_2 = \{\, a \mid a \text{ tiene orden } t \text{ par y } a^{t/2} \neq -1 \bmod N \,\}$$

es mayor que $1 - (1/2)^{h-1}$, siendo h el número de factores primos distintos de N. Para el primer conjunto se cumple

$$P\,(S_1) = \frac{N - 1 - \phi(N)}{N - 1}$$

Para calcular la probabilidad del segundo conjunto vamos a utilizar la estructura del grupo de unidades U_N [15]. Sea $N = N_1 \ldots N_h$ con $N_1 = p_1^{\alpha_1}$, $N_2 = p_2^{\alpha_2}$, $\ldots$, $N_h = p_h^{\alpha_h}$ potencias de primos impares distintos. La aplicación

$$\begin{array}{rccccc} \psi: & U_N & \longrightarrow & U_{N_1} \times \cdots \times U_{N_h} & & \\ & k & \longrightarrow & k_1 \times \cdots \times k_h & \text{con} & k_i = k \bmod N_i \end{array}$$

es un isomorfismo, siendo U_{N_i} un grupo cíclico de orden par $n_i = \phi(N_i) = p_i^{\alpha_i - 1}(p_i - 1)$ para todo $1 \leq i \leq h$. Entonces, elegir al azar un elemento $a \in U_N$ es equivalente a elegir al azar $a_1 \in U_{N_1}$, $a_2 \in U_{N_2}$, $\ldots$ y $a_h \in U_{N_h}$ y tomar $a = \psi^{-1}(a_1 \times \cdots \times a_h)$. El orden de a en U_N, t, y el orden de a_i en U_{N_i}, t_i, para todo $1 \leq i \leq h$ verifican

$$t = \text{mcm}(t_1 \ldots, t_h)$$

Obsérvese que t_1, t_2, $\ldots$ y t_h dividen a t y, si llamamos $w(t)$, $w(t_1)$, $w(t_2)$, $\ldots$ y $w(t_h)$ a las mayores potencias de dos que dividen a t, t_1, t_2, $\ldots$ y t_h respectivamente, también se cumple que $w(t_i) \leq w(t)$ para todo $1 \leq i \leq h$. En primer lugar vamos a probar que el complementario de S_2 verifica

$$P\left(\bar{S}_2\right) = P\,(\, w(t) = w(t_i) \text{ para todo } 1 \leq i \leq h \,)$$

En efecto si t es impar, es decir $w(t) = 1$, entonces t_1, t_2, $\ldots$ y t_h también son impares, es decir se cumple $w(t_1) = w(t_2) = \cdots = w(t_h) = 1$. En otro caso, si t es par y $a^{t/2} = -1 \bmod N$ vamos a probar por reducción al absurdo que $w(t_i) = w(t)$ para todo $1 \leq i \leq h$. Por hipótesis asumimos que existe un i, $1 \leq i \leq h$, tal que $w(t_i) < w(t)$. De este hecho se deduce que t_i divide a $t/2$ y, por tanto, $a_i^{t/2} = 1 \bmod N_i$. Por otro lado $a_i^{t/2} = -1 \bmod N_i$, ya que $a_i^{t/2} \bmod N_i$ es igual a la proyección sobre U_{N_i} de $\psi(a^{t/2})$ y $a^{t/2} = -1 \bmod N$. Por tanto, la hipótesis asumida conduce a la contradicción $1 = -1 \bmod N_i$.

En segundo lugar vamos a demostrar que, dado un valor de $w(t)$, la probabilidad de que $w(t_i) = w(t)$ es menor o igual que $1/2$, $1 \leq i \leq h$. Sea g un generador de U_{N_i}. Entonces $a_i = g^s$ para algún s tal que $1 \leq s \leq n_i$ y el orden de a_i verifica $st_i = n_i k$ para un entero k tal que $\text{mcd}(t_i, k) = 1$. Si denotamos por $w(s)$, $w(n_i)$ y $w(k)$ a las mayores

potencias de dos que dividen a s, n_i y k respectivamente, entonces se cumple $w(t_i) = 1$ ó $w(k) = 1$ y por tanto $w(s)$ verifica

$$\begin{cases} w(s) \geq w(n_i) \geq 2 & \text{si} \quad w(t_i) = w(t) = 1 \\ w(s) = w(n_i)/w(t_i) & \text{si} \quad w(t_i) = w(t) > 1 \end{cases}$$

Finalmente, en cualquiera de los casos, $a_i = g^s \in \bar{S}_2$ para, a lo sumo, la mitad de los valores de s. Entonces la probabilidad de que a pertenezca al complementario de S_2 cumple

$$P\left(\bar{S}_2\right) = \sum_{w(t)} P\left(\, w(t_1) = w(t)\,\right) \cdots P\left(\, w(t_h) = w(t)\,\right) \leq \sum_{w(t)} P\left(\, w(t_1) = w(t)\,\right) \left(\frac{1}{2}\right)^{h-1} = \left(\frac{1}{2}\right)^{h-1}$$

A partir de los resultados anteriores ya podemos calcular la probabilidad de la unión de los conjuntos S_1 y S_2, es decir, la probabilidad de que el algoritmo 2 obtenga un factor propio de N.

$$P(S_1 \cup S_2) \geq \frac{N-1-\phi(N)}{N-1} + \frac{\phi(N)}{N-1}\left(1 - \left(\frac{1}{2}\right)^{h-1}\right) = 1 - \frac{\phi(N)}{N-1}\left(\frac{1}{2}\right)^{h-1} \geq 1 - \left(\frac{1}{2}\right)^{h-1}$$

Si $h = 1$ la cota inferior que se obtiene es $1/(p_1 + 1)$ y se alcanza para $\alpha_1 = 2$.

9.2 Apéndice B

Vamos a calcular el valor de $\|A(k)|^2$, $0 \leq k < Q$, que aparece en la segunda transformada cuántica de Fourier del algoritmo de Shor y determina la probabilidad de obtener k al medir el primer registro. En primer lugar vamos a obtener las coordenadas de $A(k)$ en una base ortonormal: la base de computación. Para ello se introducen el cociente y el resto de la división de Q entre T, w y r respectivamente. Entonces se cumple $Q = Tw + r$ y $0 \leq r < T$. Estos elementos nos permiten representar $A(k)$ del siguiente modo:

$$A(k) = \sum_{j=0}^{Q-1} \sigma_n^{jk} |a^j \bmod N\rangle = \sum_{j=0}^{T-1} \left(\sum_{i=0}^{w-1} \sigma_n^{(Ti+j)k} \right) |a^j \bmod N\rangle + \sum_{j=0}^{r-1} \sigma_n^{(Tw+j)k} |a^j \bmod N\rangle$$

Teniendo en cuenta que los estados $|a^j \bmod N\rangle$, $0 \leq j < T$, son ortogonales y sumando las sucesiones geométricas que aparecen en la expresión anterior, se obtiene que $\|A(k)\|^2$ es igual a

$$\sum_{j=0}^{r-1} \left| \sum_{i=0}^{w} \sigma_n^{(Ti+j)k} \right|^2 + \sum_{j=r}^{T-1} \left| \sum_{i=0}^{w-1} \sigma_n^{(Ti+j)k} \right|^2 = \begin{cases} r(w+1)^2 + (T-r)w^2 & \text{si} \quad \sigma_n^{Tk} = 1 \\ \dfrac{r\left|\sigma_n^{T(w+1)k} - 1\right|^2 + (T-r)\left|\sigma_n^{Twk} - 1\right|^2}{\left|\sigma_n^{Tk} - 1\right|^2} & \text{e.o.c.} \end{cases}$$

Obsérvese que si T divide a Q, es decir si $r = 0$, entonces se verifica

$$P(k) = \frac{\|A(k)\|^2}{Q^2} = \begin{cases} \dfrac{1}{T} & \text{si } k \text{ es múltiplo de } w \\ \\ 0 & \text{si } k \text{ no es múltiplo de } w \end{cases}$$

Estamos interesados en acotar $\|A(k)\|^2$ para los valores de k que verifican $|k - jw'| \leq 1/2$ o, equivalentemente, $|Tk - jQ| \leq T/2$. Para ello vamos a considerar $z = Tk \bmod Q$, con $-Q/2 < z \leq Q/2$. Entonces el rango de valores de z que nos interesan es $|z| \leq T/2$. La acotación es sencilla si $z = 0$ pues, en este caso, se cumple que $\sigma_n^{Tk} = 1$ y, teniendo en cuenta que $r < T < N$ y $N^2 < Q$, se obtiene:

$$\|A(k)\|^2 = r(w+1)^2 + (T-r)w^2 \geq Tw^2 = \frac{Q^2}{T}\left(1 - \frac{r}{Q}\right)^2 \geq \frac{Q^2}{T}\left(1 - \frac{1}{N}\right)^2$$

Para $z \neq 0$ la acotación se obtiene a partir de las siguientes desigualdades:

$$\frac{4}{\pi^2}\theta^2 \leq \left|e^{i\theta} - 1\right|^2 \leq \theta^2 \qquad \text{si} \quad |\theta| \leq \pi$$

$$\frac{4}{\pi^2}\theta^2 \frac{\cos^2\left(\frac{\pi}{2N}\right)}{\left(1 + \frac{1}{N}\right)^2} \leq \left|e^{i\theta} - 1\right|^2 \quad \text{si} \quad |\theta| \leq \pi\left(1 + \frac{1}{N}\right)$$

A partir de estos resultados, sabiendo que $\sigma_n = e^{\frac{2\pi}{Q}i}$, $|z| \leq T/2$, $T < N$ y $N^2 < Q$, se obtiene:

$$\left|\frac{2\pi z w}{Q}\right| \leq \frac{2\pi}{Q}\frac{T}{2}\frac{Q}{T} = \pi \qquad \Longrightarrow \quad |\sigma_n^{zw} - 1|^2 \geq 4^2\frac{w^2}{Q^2}z^2$$

$$\left|\frac{2\pi z}{Q}\right| \leq \frac{2\pi}{Q}\frac{T}{2} \leq \pi \qquad \Longrightarrow \quad |\sigma_n^{z} - 1|^2 \leq 4\pi^2\frac{z^2}{Q^2}$$

$$\left|\frac{2\pi z(w+1)}{Q}\right| \leq \frac{2\pi}{Q}\frac{T}{2}\frac{Q+T}{T} \leq \pi\left(1 + \frac{1}{N}\right) \quad \Longrightarrow \quad \left|\sigma_n^{z(w+1)} - 1\right|^2 \geq 4^2\frac{(w+1)^2}{Q^2}z^2\frac{\cos^2\left(\frac{\pi}{2N}\right)}{\left(1 + \frac{1}{N}\right)^2}$$

Introduciendo las acotaciones anteriores en la definición de $A(k)$ obtenemos finalmente

$$\begin{aligned} \|A(k)\|^2 &\geq \frac{r4^2\frac{(w+1)^2}{Q^2}z^2\frac{\cos^2\left(\frac{\pi}{2N}\right)}{\left(1+\frac{1}{N}\right)^2} + (T-r)4^2\frac{w^2}{Q^2}z^2}{4\pi^2\frac{z^2}{Q^2}} \\ &\geq \frac{4}{\pi^2}Tw^2\frac{\cos^2\left(\frac{\pi}{2N}\right)}{\left(1+\frac{1}{N}\right)^2} \\ &\geq \frac{4}{\pi^2}\frac{Q^2}{T}\left(1 - \frac{1}{N}\right)^2\frac{\cos^2\left(\frac{\pi}{2N}\right)}{\left(1+\frac{1}{N}\right)^2} \end{aligned}$$

En resumen, la probabilidad de que el resultado de la medida del primer registro sea un entero k tal que $|k - jw'| \leq 1/2$ o, equivalentemente, $|Tk - jQ| \leq T/2$ para un valor de j entre 0 y $T-1$ verifica

$$P(k) = \frac{\|A(k)\|^2}{Q^2} \geq \frac{4}{\pi^2}\frac{1}{T}\left(1 - \frac{1}{N}\right)^2 \frac{\cos^2\left(\frac{\pi}{2N}\right)}{\left(1+\frac{1}{N}\right)^2} \quad \xrightarrow[N\to\infty]{} \quad \frac{4}{\pi^2}\frac{1}{T}$$

Esto significa que la probabilidad de obtener para k el entero más próximo a un múltiplo de la frecuencia formal w' tiende, cuando N crece, a $4/\pi^2$. Por tanto, el ensanchamiento de los picos de la transformada cuántica de Fourier debido a la extensión no periódica de la función está perfectamente controlado.

9.3 Apéndice C

A continuación describimos algunos resultados básicos de teoría de fracciones continuas. Si estamos interesados en un estudio un poco más detallado podemos recurrir a otras fuentes bibliográficas como, por ejemplo, [9]. El algoritmo de Euclides proporciona un método para calcular la fracción continua de un número racional. Supongamos, por ejemplo, que hemos calculado el mcd$(12, 7)$:

$$\begin{array}{rcl} 12 &=& 7\cdot 1 + 5 \\ 7 &=& 5\cdot 1 + 2 \\ 5 &=& 2\cdot 2 + 1 \\ 2 &=& 1\cdot 2 + 0 \end{array} \quad \Longrightarrow \quad \frac{12}{7} = 1+\frac{5}{7} = 1+\frac{1}{\frac{7}{5}} = 1+\frac{1}{1+\frac{2}{5}} = 1+\frac{1}{1+\frac{1}{\frac{5}{2}}} = 1+\frac{1}{1+\frac{1}{2+\frac{1}{2}}}$$

Los números 1, 1, 2, 2 se denominan cocientes parciales y la fracción continua se representa linealmente del siguiente modo: $[1, 1, 2, 2]$. Dada la fracción continua de un número racional $[a_0, a_1, \ldots, a_n]$ se denominan convergentes a las fracciones irreducibles

$$\frac{p_k}{q_k} = [a_0, a_1, \ldots, a_k] \qquad 0 \leq k \leq n$$

que se pueden calcular a partir de las siguientes relaciones de recurrencia:

$$p_k = \begin{cases} a_0 & \text{si } k = 0 \\ a_0 a_1 + 1 & \text{si } k = 1 \\ a_k p_{k-1} + p_{k-2} & \text{si } 2 \leq k \leq n \end{cases} \qquad q_k = \begin{cases} 1 & \text{si } k = 0 \\ a_1 & \text{si } k = 1 \\ a_k q_{k-1} + q_{k-2} & \text{si } 2 \leq k \leq n \end{cases}$$

Finalizamos este escueto informe enunciando las propiedades más relevantes de las convergentes. Consideremos la fracción continua del número racional $x = [a_0, a_1, \ldots, a_n]$, entonces:

1. La secuencia de denominadores es monótona: $1 = q_0 \leq q_1 < q_2 < \cdots < q_n$.
2. La secuencia de convergentes de índice par es creciente.
3. La secuencia de convergentes de índice impar es decreciente.

4. Las convergentes aproximan el número racional x:

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \frac{p_4}{q_4} < \cdots x \cdots < \frac{p_3}{q_3} < \frac{p_1}{q_1}$$

5. Si $\frac{p}{q}$ es una fracción irreducible entonces:

$$\left| x - \frac{p}{q} \right| < \frac{1}{2q^2} \qquad \Longleftrightarrow \qquad \frac{p}{q} \text{ es una convergente}$$

6. Si $x = \frac{p}{q}$ verifica $\max(|p|, |q|) \leq N$ entonces la longitud de la fracción continua es $O(\log(N))$.

Bibliografía

[1] Barenco, A.; Bennet, C. y otros, *Elementary gates for quantum computation*, Phys. Rev. A **52**, 3457–3467 (1995) (arXiv:quant-ph/9503016).

[2] Benioff, P., *Quantum mechanical Hamiltonian models of Turing machines*, J. Statistic. Phys. **29**, 515–546 (1982).

[3] Bennet, C. H., *Quantum Cryptography using any two nonorthogonal states*, Phys. Rev. Lett. **68**, 3121–3124 (1992)

[4] Bennett, C. H.; *Quantum cryptography using any two nonorthogonal states*, Phys. Rev. Lett., **68** (21), (1992).

[5] Bennet, C. H.; Brassard, G., *Quantum Cryptography: public key distribution and coin tossing*, Proc. of IEEE Int. Conf. on Computers, Systems and Signal Processing, 175–179 (1984).

[6] Bennett, C. H.; Brassard, G. and Ekert, A. K., *Quantum cryptography*, Sci. Amer. **267**, 4, 50 (1992).

[7] Boyer, M.; Brassard G.; Hoyer, P. and Tapp, A., *Tight bounds on quantum searching*, Workshop of Phys. Comp. 96, 36–43 (1996) (arXiv:quant-ph/9605034 v1).

[8] Calderbank, A. R.; Shor, P. W., *Good Quantum Error-Correcting Codes Exist*, Phys. Rev. A **54**, 1098–1105 (1996) (arXiv:quant-ph/9512032).

[9] Cilleruelo, J.; Córdoba, A., *La teoría de los números*, Biblioteca Mondadori, 1992.

[10] Cirac, J. I.; Zoller, P., *Quantum Computation with Cold Trapped Ions*, Phys. Rev. Lett. **74**, 4091–4094 (1995).

[11] Cory, D. G.; Price, M. D. and Havel, T. F., *Nuclear Magnetic Resonance Spectroscopy: An Experimentally Accesible Paradigm for Quantum Computing*, PhysComp96, New Englang Complex Systems Institute, 1996 (arXiv:quant-ph/9709001).

[12] Deutsch, D., *Quantum theory, the Church-Turing principle and the universal quantum computer*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. **400**, 97–117 (1985).

[13] Domokos, P.; Raimond, J. M. and Brune, M., *Simple Cavity-QED Two-bit Universal Quantum Logic Gate: The Principle and Expected Performances*, Phys. Rev. A **52**, 3554–3559 (1995).

[14] Ekert, A., *Quantum Cryptography based on Bell's theorem*, Phys. Rev. Lett. **67**, 661–663 (1991)

[15] Ekert, A.; Hayden, P. and Inamori, H., *Basic concepts in quantum computation*, arXiv: quant-ph/0011013 (2001).

[16] Feynman R., *Simulating physics with computers*, Internat. J. Theoret. Phys. **21**, 467–488 (1982).

[17] Gathen, J. and Gerhard, J., *Modern Computer Algebra*, Cambridge University Press, 1999.

[18] Gershenfeld, N. A. and Chuang, I. L., *Bulk Spin-Resonance Quantum Computation*, Science **257**, 350–356 (1997).

[19] Gottesman, D., *Stabilizer Codes and Quantum Error Correction*, Ph.D. thesis, California Institute of Technology, Pasadena, CA, 1997 (arXiv:quant-ph/9705052).

[20] Grover, L. K., *A fast quantum mechanical algorithm for database search*, Proceedings of the 28th ACM Symposium on the Theory of Computing, 212–219 (1996) (arXiv:quant-ph/9605043).

[21] Hirvansalo, M., *Quantum Computing*, Springer, 2001.

[22] Kane, B. E., *A Silicon-Based Nuclear Spin Quantum Computer*, Nature **393**, 133–137 (1998).

[23] Lenstra, A. K.; Lenstra, H. W.; Manasse, M. S. and Pollard, J. M., *The number field sieve*, Proceedings of the 22th ACM Symposium of the Theory of Computing, Baltimore, MD, ACM Press, 564–572 (1990).

[24] Nielsen, M. A.; Chuang, L. I., *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.

[25] Preskill, J., *Robust solutions to hard problems*, *Nature*, **391**, 631–632 (1998). arXiv:quant-ph/9712048 (1997).

[26] Rivest, R. L.; Shamir, A. and Adleman, L. M., *A Method for Obtaining Digital Signatures and Public-Key Cryptosystems*, Comm. ACM **21** (2), 120–126 (1978).

[27] Scarany, V., Acín, A., Ribordy, G. and Gisin, N. *Quantum Cryptography Protocols Robust against Photon Number Splitting Attacks for weak Laser Pulse implementations.* Phys. Rev. A **92** (5), 57901-1–57901-4 (2004).

[28] Shor, P. W., *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*, SIAM J. Comput. **26**, 1484–1509 (1994) (arXiv:quant-ph/9508027).

[29] Shor, P. W., *Scheme for Reducing Decoherence in Quantum Computer Memory*, Phys. Rev. A **52**, 2493–2496 (1995).

[30] Steane, A. M., *Multiple Particle Interference and Quantum Error Correction*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. **452**, 2551–2576 (1996).

[31] Vedral, V.; Barenco, A. and Ekert, A., *Quantum Networks for Elementary Arithmetic Operations*, arXiv:quant-ph/9511018 (1995).

[32] Vrijen, R.; Yablonovitch, E.; Wang, K.; Jiang, H. W.; Baladin, A.; Roychowdhury, V.; Mor, T. and DiVincenzo, D., *Electron Spin Resonance Transistors for Quantum Computing in Silicon-Germanium Hetero-Structures*, arXiv:quant-ph/9905096 (1999).

[33] Wootters, W. K. and Zurek, W. H., *A Single quantum cannot be cloned*, Nature **299**, 802–803 (1982).

[34] Zalka, C., *Grover's quantum searching algorithm is optimal*, arXiv:quant-ph/9711070 v2 (1999).