INFORMACIÓN CUÁNTICA Y APLICACIONES

L. L. Salcedo
Departamento de Física Atómica, Molecular y Nuclear,
Universidad de Granada, E-18071 Granada, Spain
E-mail: salcedo@ugr.es

3 de junio de 2024

Resumen

Apuntes de la asignatura de Información cuántica y aplicaciones.
Versión 2-jun-2024 2022-2024. salcedo@ugr.es

Índice

1. Introducción

La información es física. Por ejemplo, la **entropía de Shannon** (que mide la información en un sistema clásico) coincide con la entropía de Boltzmann (sistema termodinámico)

$$S = -\sum_{\alpha} p_\alpha \log(p_\alpha) = H(p_\alpha) \tag{1.1}$$

(la distribución de Boltzmann a temperatura T es la que maximiza S con $\{H\}$ dado, $1/T$ es el multiplicador de Lagrange.)

O el **principio de Landauer**: para borrar un bit hay que gastar una energía mínima de $kT\ln(2)$. Sin embargo (a raíz de la computación cuántica) se sabe que se puede hacer computación clásica reversible (la cuántica lo es necesariamente).

Los sistemas cuánticos se basan en **amplitud de probabilidad** no sólo **probabilidad** como los clásicos lo cual introduce importantes diferencias. Sólo un sistema físico **cuántico** (fotones, iones, etc) es capaz de hacer una computación con las reglas cuánticas (no simularlas). Un sistema cuántico se puede simular de manera clásica pero no *eficientemente*. El coste en número de qubits es aditivo en el caso cuántico y exponencial 2^n en su emulación clásica. Se sigue del teorema de Bell que implica que no se puede simular un sistema cuántico con variables clásicas locales, deben ser no-locales, de modo que conecten todas las componentes entre sí.

En un momento dado se pensaba que limitaciones cuánticas tales como el principio de incertidumbre arruinarían, o pondrían insalvables impedimentos a la posibilidad de hacer computación cuántica. De hecho no es así. Todo lo que se puede hacer clásicamente se puede reproducir cuánticamente, y por el contrario hay cosas imposibles clásicamente que son viables cuánticamente (por ejemplo, problema de Deutsch). Un punto clave es el llamado **paralelismo cuántico** y **entrelazamiento** (en esencia todo viene de la **superposición** cuántica). Hay un paralelismo clásico basado en probabilidades; éste siempre tiende a **desenfocar** (difundir, igualar) en cambio el cuántico basado en amplitud de probabilidad, gracias a la interferencia y por tanto la posibilidad de **enfocar** (concentrar).

El tratamiento cuántico es mucho más eficiente cuando se trata por ejemplo de hacer transformadas de Fourier (mejora exponencial en la dimensión del espacio, $\log(N)$ frente a $N\log(N)$ de la transformada rápida de Fourier). Y es cuadráticamente más eficiente en otros casos (el algoritmo de búsqueda de Grover, va como $O(\sqrt{N})$ frente a $O(N)$ clásico) haciendo uso del paralelismo cuántico. Un ejemplo paradigmático, basado en transformada de Fourier, es la factorización de números enteros

de Shor (1996):

$$\begin{aligned} T_{\text{clás}} &= O(e^{2(\ln N)^{1/3}(\ln\ln N)^{2/3}}) \qquad \text{(mejor algoritmo clásico conocido)}, \\ T_{\text{Shor}} &= O((\ln N)^3) \end{aligned} \tag{1.2}$$

En la computación clásica hay errores y se arregla con redundancia. Se pensaba que en el caso cuántico (debido al **teorema de no clonación**) los errores producirían pérdida de coherencia cuántica y arruinarían irremediablemente el proceso. De hecho Shor probó que no era así, se puede usar también redundancia en el caso cuántico. El número de qubits extra es grande pero al final converge a una computación cuántica fiable.

También la **comunicación cuántica** tiene aspectos muy novedosos respecto de la clásica. Por ejemplo la posibilidad de teleportación, o criptografía segura, a prueba de espías.

1.1. El qubit

En su versión más simple, la información consiste en indicar, de entre varias posibles alternativas, cuál es la correcta. La unidad básica de información clásica es el **bit**, en presencia de dos alternativas. Las dos opciones se suelen denotar 0 y 1. Su versión cuántica es el **qubit**, [1.1] es el sistema cuántico más simple no trivial y corresponde a un espacio de Hilbert (complejo) de dimensión 2, es decir, admite dos estados ortogonales (una base ortonormal) que suelen denotarse $|0\rangle$ y $|1\rangle$. [1.2] La base física de un qubit puede ser absolutamente cualquier sistema físico que tenga dos estados activos tal como el espín de un ion con espín $1/2$, la polarización de un fotón, el estado de un átomo en un pozo de potencial con dos niveles accesibles, el estado de ocupación o no de un electrón en un nivel atómico, etc, aunque en la práctica sólo se usan los sistemas más fáciles de controlar y manipular sin perder coherencia cuántica.

La diferencia fundamental entre bit y qubit es que el bit sólo puede ser 0 ó 1, mientras que el qubit puede encontrarse en un estado superposición de los dos estados básicos $|0\rangle$ y $|1\rangle$ (que definen la base computacional de un qubit)

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \qquad \alpha,\beta \in \mathbb{C}, \quad |\alpha|^2 + |\beta|^2 = 1. \tag{1.3}$$

[1.1] Más correctamente, cúbit.

[1.2] Aquí "qubit" se está usando para referirse a un sistema físico cuántico –en realidad todos los sistemas físicos lo son– con dos estados posibles. En completa analogía con los bits, también se usa qubit como unidad de información. Así un sistema con 2^n estados puede almacenar hasta n qubits de información.

Se ha elegido $|\psi\rangle$ normalizado. Recuérdese que $|\psi\rangle$ y $\lambda|\psi\rangle$, $\lambda \in \mathbb{C}\setminus\{0\}$ (en particular $e^{i\varphi}|\psi\rangle$) representan el mismo estado físico. $|\psi\rangle \in \mathscr{H}$, $\dim\mathscr{H} = 2$, $\mathscr{H} \approx \mathbb{C}^2$. El isomorfismo depende de la base $\{|0\rangle, |1\rangle\}$ elegida, $|\psi\rangle \to \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$. $\quad |0\rangle = \begin{pmatrix}1\\ 0\end{pmatrix}, \quad |1\rangle = \begin{pmatrix}0\\ 1\end{pmatrix}$.

α, β son las amplitudes de probabilidad (y $|\alpha|^2$, $|\beta|^2$ las probabilidades) de encontrar el sistema en el estado $|\psi\rangle$ en el estado $|0\rangle$ o $|1\rangle$ al hacer una medida de un observable con estados propios $|0\rangle$ y $|1\rangle$. Sin embargo, a diferencia de la situación clásica, $|0\rangle$ y $|1\rangle$ no es una base privilegiada a priori, cualquier otra base ortonormal es igualmente válida y tiene sus observables asociados; cualquier estado $|\psi\rangle$ puede formar parte de una tal base ortonormal. Parecería entonces que hay muchos más estados cuánticos posibles que clásicos. En realidad la situación es más sutil, porque en una sola medida sobre $|\psi\rangle$, digamos de un observable A con la base ortonormal $\{|u_0\rangle, |u_1\rangle\}$ como estados propios, sólo se va a poder obtener uno de los dos resultados $|u_0\rangle$ o $|u_1\rangle$ y no otros estados. No se puede determinar $|\psi\rangle$ con una sola medida. [1.3] Por tanto sí hay más estados cuánticos que clásicos en cierto sentido, pero cada vez sólo se puede elegir medir un observable maximal (una base) que sólo decidirá entre dos opciones.

Observación:

Aquí nos estamos refiriendo a estados con máxima información posible. 0 y 1 serían los "estados puros" del bit. Más generalmente, el bit puede estar en un estado mezcla (estadística) con probabilidad p_0 de ser 0 y p_1 de ser 1, siendo $p_{0,1} \geq 0$, $p_0 + p_1 = 1$. Esta mezcla clásica indica ignorancia subsanable con más información; nosotros no sabemos cuál es exactamente el estado pero alguien lo puede saber, o puede estar registrado en algún sitio. $|\psi\rangle$ representa un estado puro del qubit. Puro quiere decir unívocamente caracterizado por un conjunto maximal de observables compatibles; [1.4] la aleatoriedad observada al hacer una medida sobre $|\psi\rangle$ de otros observables no compatibles con los que definen el estado no se puede subsanar con más información. El qubit puede estar también en un estado mezcla estadística de estados puros $|\psi\rangle$. Las mezclas se estudiarán en el Tema 2.

En cierto sentido un estado puro $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ se parece superficialmente a un estado mezcla clásico que se describe con probabilidades p_0 y p_1, pero con una "probabilidad exótica" que es la amplitud de probabilidad, que a diferencia de la auténtica probabilidad toma valores complejos. Tanto en el caso clásico (estados descritos por el par (p_0, p_1)) como en el cuántico (estados descritos por el par (α, β)) todo el tiempo hay sólo dos opciones, y ambos espacios son lineales. Sin embargo la versión cuántica introduce importantes diferencias con la clásica. Por ejemplo, de entre todos los estados clásicos (p_0, p_1) los casos $(1,0)$ (bit= 0) y $(0,1)$ (bit= 1) son especiales (técnicamente, "extremales") en cambio no hay estados cuánticos (α, β) con

[1.3] A menos que se sepa a priori que $|\psi\rangle$ era o bien $|u_0\rangle$ o bien $|u_1\rangle$, y se quiere saber cuál de los dos es. Eso es un bit clásico.

[1.4] Compatible es que la medida de uno no modifica la del otro. Maximal es que cualquier otro observable compatible es función de éstos.

propiedades especiales (extremales) dentro del espacio de Hilbert. Otra diferencia, por ejemplo es que no es posible representar un grupo de Lie de transformaciones, digamos las rotaciones, en el espacio clásico (p_0, p_1), pero sí en el espacio cuántico (α, β). Probabilidades y amplitudes de probabilidad son estructuras matemáticas compatibles. Las primeras se pueden incluir como un caso particular de las segundas y el sistema cuántico es una extensión del clásico que permite nuevas posibilidades, y que la naturaleza utiliza.

Como es sabido, no todos los vectores no nulos en el espacio de Hilbert $\mathscr{H}$ del sistema cuántico corresponden a estados físicos distintos. Cada estado físico se identifica con un rayo unitario. Esto es, $|\psi\rangle \neq 0$ y $\lambda|\psi\rangle$ ($\lambda \in \mathbb{C},\ \lambda \neq 0$,) representan el mismo estado físico. Se suelen utilizar estados normalizados a 1, $\langle\psi|\psi\rangle = |\alpha|^2 + |\beta|^2 = 1$, usando que la base $\{|0\rangle, |1\rangle\}$ es ortonormal,

$$\langle i|j\rangle = \delta_{ij}, \qquad i, j \in \{0, 1\}, \tag{1.4}$$

pero aún queda la libertad de elegir la fase. Si $\alpha = 0$ se puede elegir $\beta = 1$, es decir, $|\psi\rangle = |1\rangle$. En otro caso se puede elegir $\alpha > 0$,

$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\,\mathrm{sen}(\theta/2)|1\rangle = \begin{pmatrix} \cos(\theta/2) \\ e^{i\phi}\,\mathrm{sen}(\theta/2) \end{pmatrix}, \qquad \begin{array}{l} 0 \leq \theta \leq \pi, \quad 0 \leq \phi < 2\pi \\ \phi = 0 \quad \text{si} \quad \theta = \pi. \end{array} \tag{1.5}$$

También se elige a veces $-\pi < \phi \leq \pi$, y también se usa la parametrización $|\psi\rangle = \begin{pmatrix} e^{-i\phi/2}\cos(\theta/2) \\ e^{i\phi/2}\,\mathrm{sen}(\theta/2) \end{pmatrix}$.

El par (θ, ϕ) se puede identificar con las coordenadas esféricas de un vector unitario $\hat{n}$ en $\mathbb{R}^3$, en el que θ es el ángulo polar y ϕ el acimutal,

$$\hat{n} = (\mathrm{sen}(\theta)\cos(\phi), \mathrm{sen}(\theta)\,\mathrm{sen}(\phi), \cos(\theta)). \tag{1.6}$$

Por tanto cada estado físico distinto de un qubit se corresponde biunívocamente con un punto en una superficie esférica S^2, denominada esfera de Bloch (Fig. 1.1). El estado $|0\rangle$ se corresponde con el vector $\hat{n} = (0, 0, 1)$ (polo norte) y $|1\rangle$ se corresponde con el vector $\hat{n} = (0, 0, -1)$ (polo sur). En general, dos estados ortogonales del qubit se corresponden con puntos de la esfera diametralmente opuestos, como es fácil comprobar. La esfera de Bloch es útil para visualizar los estados puros de un qubit y sus transformaciones, y también los estados mezcla como se verá. Una transformación unitaria en el espacio de Hilbert del qubit equivale a una rotación en la esfera de Bloch. [1.5] La simetría de la esfera visualiza la simetría entre estados cuánticos puros (no hay estados privilegiados, por ejemplo

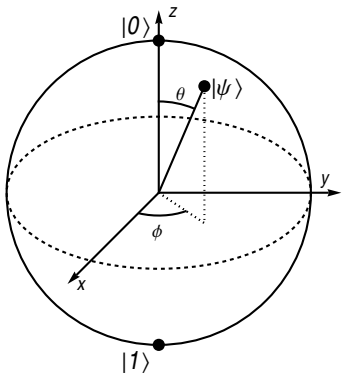


Figura 1.1: Esfera de Bloch.

extremales, en $\mathscr{H}$, a menos que se introduzca estructura adicional).

El espacio de Hilbert de n qubits, $\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$ admite como base la denominada **base computacional**, que es la formada por el producto tensorial de n qubits en estados $|0\rangle$ o $|1\rangle$,

$$\begin{aligned} |0\rangle \otimes \cdots |0\rangle \otimes |0\rangle &= |0\cdots 00\rangle \qquad x = 0 \\ |0\rangle \otimes \cdots |0\rangle \otimes |1\rangle &= |0\cdots 01\rangle \qquad x = 1 \\ &\vdots \\ |1\rangle \otimes \cdots |1\rangle \otimes |1\rangle &= |1\cdots 11\rangle \qquad x = 2^n - 1 \end{aligned} \tag{1.7}$$

Cada $x \in 0, 1, \ldots, N-1$ determina una cadena de n bits que son sus cifras en binario, siendo $N \equiv 2^n$. El espacio de Hilbert de los n qubits tiene dimensión N. El vector-estado más general de los n-qubits es una superposición

$$|\psi\rangle = \sum_{x=0}^{N-1} \psi_x |x\rangle, \qquad \psi_x \in \mathbb{C}. \tag{1.8}$$

1.2. Puertas cuánticas

Todas las transformaciones físicas (por ejemplo, una rotación) sobre un sistema cuántico con espacio de Hilbert $\mathscr{H}$ las realizan operadores unitarios. En particular la **evolución dinámica** en el tiempo la lleva a cabo el operador de evolución $U(t_f, t_i)$ que es unitario; para un sistema conservativo con hamiltoniano H, $U(t_f, t_i) = e^{-i(t_f - t_i)H/\hbar}$, es unitario por ser H autoadjunto.

[1.5] Por ejemplo $\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \to \begin{pmatrix} \alpha \\ e^{i\varphi}\beta \end{pmatrix}$ corresponde a una rotación de ángulo φ alrededor del eje z de la esfera.

Por este motivo los operadores unitarios (los que conservan la estructura del espacio de Hilbert) son fundamentales en la teoría cuántica. La unitaridad expresa la conservación de la probabilidad durante la evolución. Como es sabido, dado un estado $|\psi\rangle$, la probabilidad de encontrarlo en el estado $|\phi\rangle$ al hacer una medida (suponemos ambos normalizados) viene dado por su **solapamiento** $|\langle\phi|\psi\rangle|^2$. El solapamiento se conserva bajo evolución temporal de ambos estados. Así por ejemplo dos estados ortogonales son totalmente distinguibles (son incompatibles) y eso se conserva bajo evolución.

1.2.0.1. Apartado matemático: operadores unitarios

Un operador (lineal) $U : \mathscr{H} \to \mathscr{H}$ es **unitario** cuando es invertible y además es **isométrico**, es decir, conserva el producto escalar (o equivalentemente la norma) del espacio de Hilbert complejo $\mathscr{H}$. [1.6] Equivale a decir que $U^{-1} = U^\dagger$. En efecto,

$$\forall|\psi\rangle, |\phi\rangle \in \mathscr{H} \qquad \langle\psi|\phi\rangle = \langle U\psi|U\phi\rangle = \langle\psi|U^\dagger U\phi\rangle \tag{1.9}$$

implica

$$U^\dagger U = I, \qquad U^\dagger = U^{-1}. \tag{1.10}$$

(Se ha usado que U es invertible. Isométrico implica inyectivo, y eso para $\mathscr{H}$ de dimensión finita ya garantiza que U debe ser invertible, pero no en dimensión infinita.)

Los operadores unitarios son los que transforman una base ortonormal en otra. Si $\{|i\rangle\}$ es una base ortonormal, $\{|i\rangle' = U|i\rangle\}$ también lo es. Los operadores unitarios forman el grupo $\mathrm{U}(\mathscr{H})$, isomorfo a $\mathrm{U}(n)$ para $n \equiv \dim\mathscr{H}$.

En espacios de dimensión finita, la matriz de un operador unitario *en una base ortonormal*

$$U_{ij} \equiv \langle i|U|j\rangle \tag{1.11}$$

es una **matriz unitaria** (y viceversa, U es unitario si y sólo si U_{ij} es unitaria). Es decir, tanto sus columnas como sus filas forman una base ortonormal de $\mathbb{C}^n$:

$$\begin{aligned} U^*_{ki}U_{kj} &= \delta_{ij} \qquad \text{ortonormalidad por columnas,} \\ U^*_{ik}U_{jk} &= \delta_{ij} \qquad \text{ortonormalidad por filas.} \end{aligned} \tag{1.12}$$

(Se sobreentiende suma sobre k.)

[1.6] Más generalmente se pueden definir operadores unitarios entre dos espacios distintos. Un operador unitario es un isomorfismo de espacios de Hilbert.

Las **puertas cuánticas** son operadores unitarios que actúan sobre uno o más qubits. Son unitarios ya que representan transformaciones físicas del sistema de qubits. Unitario implica invertible y eso hace que algunas puertas lógicas clásicas (es decir, que trabajan con bits) no tengan versión cuántica por no ser reversibles.

Este es el caso por ejemplo las puertas AND y OR. Si X es una proposición, su valor x de "verdadero" o "falso" se puede representar por bits 1 o 0, respectivamente. La conjunción $X \wedge Y$ corresponde a la puerta $\mathrm{AND}(x,y) = xy$ (producto numérico de x e y, $x, y \in \{0,1\}$). Igualmente la disyunción $X \vee Y$ es la puerta $\mathrm{OR}(x,y) = x + y - xy$. Claramente estas puertas no son reversibles (las funciones no son inyectivas) ya que el resultado no determinada las entradas x e y.

En cambio la negación $\neg X$, con puerta lógica clásica $\mathrm{NOT}(x) = 1 - x$ sí es invertible, y tiene una versión cuántica. Clásicamente $0 \to 1$ y $1 \to 0$, cuánticamente $|0\rangle \to |1\rangle$ y $|1\rangle \to |0\rangle$. Es una permutación de una base ortonormal y por tanto unitario. Esta puerta cuántica se suele denominar X porque su matriz en la base computacional es la matriz de Pauli $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$,

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \longrightarrow X|\psi\rangle = \alpha|1\rangle + \beta|0\rangle, \qquad \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \longrightarrow \begin{pmatrix} \beta \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{1.13}$$

(El convenio para construir la matriz es que $|0\rangle$ es el primer vector de la base y $|1\rangle$ es el segundo.) También

$$X = |0\rangle\langle 1| + |1\rangle\langle 0|. \tag{1.14}$$

El operador X satisface $X^2 = I$, $X^\dagger = X$.

Las únicas puertas clásicas de un bit en un bit son $f(x) = x$, $f(x) = 1 - x$, $f(x) = 0$ y $f(x) = 1$, y sólo las dos primeras son invertibles. [1.7] También hay puertas cuánticas que no tienen una versión clásica. Una puerta puramente cuántica de un qubit especialmente útil es la puerta de Hadamard, H. Este operador actúa de la siguiente manera:

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \equiv |+\rangle, \qquad H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \equiv |-\rangle. \tag{1.15}$$

Aquí se han definido los estados $|+\rangle$ y $|-\rangle$, que forman otra base ortonormal dado que H es un operador unitario. Matricialmente

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \xrightarrow[H]{} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{1.16}$$

[1.7] Las puertas clásicas pueden procesar de n a m bit, pero $n = m$ si son reversibles. Las cuánticas deben procesar de n a n qubits por ser reversibles.

Obviamente la puerta de Hadamard no es clásica ya que produce superposiciones de estados (de la base computacional). Esta puerta satisface $H^2 = I$, siendo I el operador identidad. Es decir $H = H^{-1} = H^\dagger$.

Una puerta clásica reversible importante es la puerta CNOT, **No-Controlado**, que actúa sobre pares de bits de acuerdo con la regla

$$00 \to 00, \quad 01 \to 01, \quad 10 \to 11, \quad 11 \to 10. \tag{1.17}$$

Es decir en todo caso el primer bit se queda igual, pero el segundo cambia (se voltea $0 \leftrightarrow 1$) si y sólo si el primero es 1. El primer bit se denomina de **control** y el segundo el bit blanco, diana, **objetivo** o controlado. O sea, cuando el bit de control está conectado (1, "on") sobre el segundo actúa NOT. La acción de CNOT se puede expresar como

$$\mathrm{CNOT}(x,y) = (x, x \oplus y). \tag{1.18}$$

Aquí $\oplus$ representa la suma de bits módulo 2,

$$x \oplus y \equiv x + y \pmod 2 \tag{1.19}$$

que equivale XOR (disyunción-exclusiva) que produce 1 (verdadero) si y sólo si una de las proposiciones es verdadera y la otra falsa.

La puerta CNOT es reversible (de hecho es su propio inverso) y por tanto tiene una versión cuántica que actúa sobre dos qubits:

$$\begin{aligned} |0\rangle|0\rangle &\longrightarrow |0\rangle|0\rangle, \qquad |0\rangle|1\rangle \longrightarrow |0\rangle|1\rangle, \\ |1\rangle|0\rangle &\longrightarrow |1\rangle|1\rangle, \qquad |1\rangle|1\rangle \longrightarrow |1\rangle|0\rangle. \end{aligned} \tag{1.20}$$

Puesto que $|i\rangle \otimes |j\rangle$, $i, j \in \{0,1\}$ forman una base del espacio de dos qubits (la base computacional) estas relaciones definen completamente a CNOT como un operador en $\mathbb{C}^2 \otimes \mathbb{C}^2$ que es unitario, ya que es simplemente una permutación de la base. Además satisface $\mathrm{CNOT}^2 = I$. Su matriz es

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{1.21}$$

(De nuevo, la base está ordenada según $00, 01, 10, 11$.) También puede expresarse

$$\mathrm{CNOT} = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes X. \tag{1.22}$$

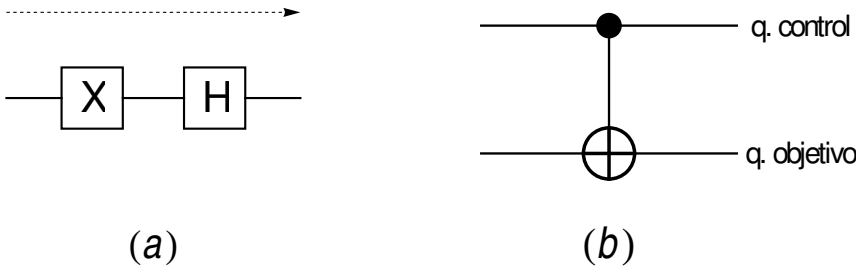


Figura 1.2: (a) Puertas X y H. La flecha (que no forma parte del circuito) indica el orden de operación de las puertas: Los qubits entran por la izquierda y salen procesados por la derecha. (b) Puerta CNOT.

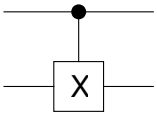
$|j\rangle\langle j|$ es proyector sobre el estado $|j\rangle$, en este caso del primer qubit.

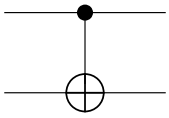
Hay que notar que el qubit de control no cambia en la base computacional, pero sí es afectado por la puerta. Por ejemplo, en un estado inicial $|+\rangle \otimes |0\rangle$ la probabilidad de obtener $|-\rangle$ en una medida del primer qubit es cero, pero deja de serlo después de aplicar CNOT.

Más generalmente una puerta $C(U)$ o U controlado corresponde a $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U$, siendo U un operador que puede actuar o no sobre el qubit controlado, dependiendo del estado del qubit de control.

1.3. Circuitos cuánticos

Combinando puertas lógicas se forman circuitos para procesar los bits o qubits. Cada qubit se representa por una línea horizontal. Es usual ordenar los qubits de arriba a abajo. En el conjunto de líneas se insertan las puertas que van actuando sobre los qubits. Así en Fig. 1.2a se representan las puertas X y H que actúan en ese orden: $|\psi\rangle \to X|\psi\rangle \to HX|\psi\rangle$. (Nótese que el orden de operadores en el circuito $\cdots A \cdots B \cdots$ produce el producto $\cdots B \cdots A \cdots$ sobre los estados.)

La puerta de dos qubits CNOT está representada en la Fig. 1.2b. El qubit superior en la figura es el qubit de control y el inferior el qubit objetivo. De acuerdo con su definición la puerta CNOT se puede representar mediante 

, pero es más usual usar  debido a $\mathrm{CNOT}(x,y) = (x, x \oplus y)$.

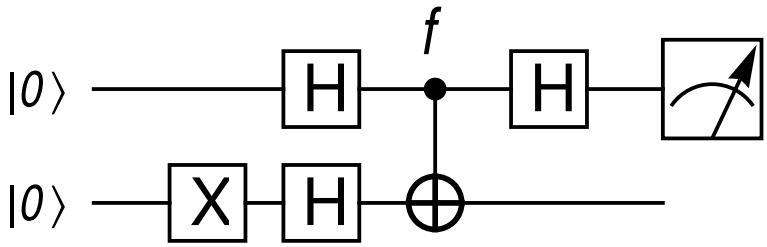


Figura 1.3: Circuito para el algoritmo de Deutsch.

Un circuito cuántico de n qubits equivale a un operador unitario en un espacio de $N = 2^n$ dimensiones. Se puede demostrar que mediante puertas de un qubit y la puerta CNOT, es posible construir cualquier operador unitario en un sistema de n qubits (usando un número finito de puertas y sin usar qubits auxiliares). Por este motivo es esencial en cualquier implementación física del sistema de qubits (por ejemplo iones atrapados) conseguir implementar la puerta CNOT (o similar) y ello siempre requiere interacción entre qubits. Que todos los operadores unitarios sean representables mediante circuitos implica también que cualquier estado cuántico de los n qubits es realizable empezando desde $|00\cdots0\rangle$.

1.4. Algoritmo de Deutsch

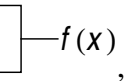
El **problema de Deutsch** es extremadamente simple, pero sirve para demostrar que los qubits, gracias a la posibilidad de superposición, pueden hacer algo que no es posible con bits. El problema de Deutsch es el siguiente. se tiene una función desconocida $f : \{0,1\} \longrightarrow \{0,1\}$. Hay cuatro funciones posibles, dos constantes y dos no:

$$f_{00} : \begin{matrix} 0 \to 0 \\ 1 \to 0 \end{matrix} \qquad f_{01} : \begin{matrix} 0 \to 0 \\ 1 \to 1 \end{matrix} \qquad f_{10} : \begin{matrix} 0 \to 1 \\ 1 \to 0 \end{matrix} \qquad f_{11} : \begin{matrix} 0 \to 1 \\ 1 \to 1 \end{matrix} \tag{1.23}$$

Se trata de determinar *con una sola llamada a la función* si ésta es constante o no. Una forma de visualizar el problema es considerar un sistema interruptor-bombilla: f_{01}, f_{10} corresponden a que la bombilla se enciende y se apaga según el estado del interruptor, en cambio para f_{00}, f_{11} bombilla e interruptor no están conectados.

Clásicamente (bits) se necesitan dos llamadas (dos evaluaciones) de la función para saber de cuál de los dos tipos es, constante o no. Cuánticamente (qubits) se puede hacer con una sola llamada a la función.

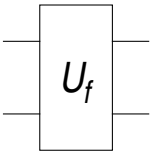
Lo primero es incluir la evaluación de la función en el circuito. Clásicamente se podría usar una

puerta de un bit, 

, sin embargo esta puerta no es reversible en general (concretamente cuando f es constante). Hay un método para codificar la función mediante una puerta clásica reversible, *usando dos bits*, a saber, $(x,y) \mapsto (x, y \oplus f(x))$. Es una puerta f-CNOT: x es el bit de control, y NOT actúa sobre el bit objetivo y sólo cuando $f(x) = 1$. También se denomina U_f. Evidentemente esta puerta es reversible porque si se aplica dos veces queda $(x,y) \mapsto (x,y)$ (coincide con su inversa). Y también es obvio que la puerta identifica unívocamente a la función f. [1.8]

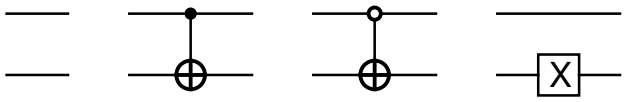
La puerta f-CNOT aparece en la Fig. 1.3. Esta puerta está definida automáticamente para qubits,

$$U_f|x\rangle \otimes |y\rangle = |x\rangle \otimes |y \oplus f(x)\rangle, \qquad x,y \in \{0,1\}. \tag{1.24}$$

Lo más frecuente es representar f-CNOT en circuitos como 

. Un circuito en el que aparezca una vez U_f cuenta como llamar f una vez, eso vale en el caso clásico y en el cuántico. Hay que tener en cuenta que si una puerta aparece una vez en un circuito cuántico esa puerta va a actuar a lo sumo una vez (puede no actuar si está controlada) ya que en circuitos cuánticos no hay bucles (DO-loop's).

El circuito para el algoritmo de Deutsch está en la Fig. 1.3. El estado inicial es $|0\rangle|0\rangle$, esta es una elección estándar (se podría simplificar el circuito empezando de otro estados y ahí hacer actuar U_f). El símbolo  representa medir el estado del qubit en la base computacional.

[1.8] Circuitos explícitos que realizan U_f para $f_{00}, f_{01}, f_{10}, f_{11}$: 

.

$$
\begin{aligned}
&|0\rangle|0\rangle \xrightarrow[X_2]{} |0\rangle|1\rangle \\
&\xrightarrow[H_1H_2]{} |+\rangle|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle) \\
&= \frac{1}{2}|0\rangle(|0\rangle-|1\rangle)+\frac{1}{2}|1\rangle(|0\rangle-|1\rangle) \\
&\xrightarrow[U_f]{} \frac{1}{2}|0\rangle(|0\oplus f(0)\rangle-|1\oplus f(0)\rangle)+\frac{1}{2}|1\rangle(|0\oplus f(1)\rangle-|1\oplus f(1)\rangle) \\
&= \frac{(-1)^{f(0)}}{\sqrt{2}}|0\rangle|-\rangle+\frac{(-1)^{f(1)}}{\sqrt{2}}|1\rangle|-\rangle \\
&\xrightarrow[H_1]{} \frac{(-1)^{f(0)}}{\sqrt{2}}|+\rangle|-\rangle+\frac{(-1)^{f(1)}}{\sqrt{2}}|-\rangle|-\rangle \\
&= \frac{(-1)^{f(0)}+(-1)^{f(1)}}{2}|0\rangle|-\rangle+\frac{(-1)^{f(0)}-(-1)^{f(1)}}{2}|1\rangle|-\rangle,
\end{aligned}
\tag{1.25}
$$

donde se ha usado la propiedad

$$
\frac{1}{\sqrt{2}}(|0\oplus y\rangle-|1\oplus y\rangle)=\left\{\begin{array}{cccc} |-\rangle & \text{si} & y=0 \\ -|-\rangle & \text{si} & y=1 \end{array}\right\}=(-1)^y|-\rangle. \tag{1.26}
$$

Es claro que, salvo fase, el estado final es

$$
\begin{array}{lll} |0\rangle|-\rangle & \text{si} & f(0)=f(1) \\ |1\rangle|-\rangle & \text{si} & f(0)\neq f(1) \end{array} \tag{1.27}
$$

por tanto al medir el primer qubit en la base $\{|0\rangle,|1\rangle\}$ se determina si la función es constante o no, y se hace con una sola llamada a la función.

La clave para determinar la clase de la función con una sola llamada es el denominado **paralelismo cuántico**: al aplicar la puerta de Hadamard H sobre el primer qubit se produce una superposición $|0\rangle+|1\rangle$ de modo que al actuar la puerta U_f, la función f se evalúa a la vez para $x=0$ y para $x=1$.

$$
|0\rangle \longrightarrow |0\rangle+|1\rangle \longrightarrow (-1)^{f(0)}|0\rangle+(-1)^{f(1)}|1\rangle. \tag{1.28}
$$

El segundo qubit es un qubit auxiliar para que U_f actúe así sobre el primer qubit. La puerta H final concentra el resultado sobre $|0\rangle$ o $|1\rangle$ según sea f constante o no. Es un proceso de interferencia. El

qubit se ha desdoblado y vuelto a juntar, si no hubiera U_f en medio volvería al estado inicial: interferencia constructiva para $|0\rangle$ y destructiva para $|1\rangle$. U_f puede modificar esa interferencia e invertir el resultado.

Lo que no es posible es usar el paralelismo cuántico para determinar completamente f con una sola llamada. Por ejemplo, si se envía el estado $|+\rangle|0\rangle$ a la puerta f-CNOT, se obtiene $\frac{1}{\sqrt{2}}(|0\rangle|f(0)\rangle + |1\rangle|f(1)\rangle)$, que contiene los dos valores de f etiquetados por el primer qubit. Sin embargo si se mide este estado en la base computacional se obtendrá al azar **uno** de los estados $|i\rangle|j\rangle$ que proporcionará $j = f(i)$, pero no el otro valor $f(1-i)$.

Nótese que la normalización final tiene que ser 1. Por ese motivo los factores **globales** que van apareciendo en cada paso se podrían omitir. Por otra parte guardarlos puede usarse como comprobación del cálculo.

La computación cuántica puede hacer todo lo que hace la clásica: basta no salirse nunca de la base computacional, usando exclusivamente la versión cuántica de puertas clásicas reversibles. El algoritmo de Deutsch demostró que la computación cuántica puede hacer cosas vetadas para la clásica. Pero también queda claro que el procedimiento para conseguirlo puede no ser trivial, por el contrario, puede ser considerablemente sofisticado.

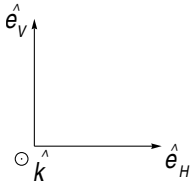
También hay que notar que después de utilizar el circuito, el circuito "sabe" el estado final, pero nosotros no. El paso final es siempre medir los qubits en una base (sin pérdida de generalidad se puede tomar la computacional) [1.9] y nos dará exactamente uno de los estados de esa base. La función de onda se borra al hacer la medida. Esto es análogo en un problema estocástico clásico: después de múltiples manipulaciones aleatorias de un cubilete con dados cada cara tiene cierta probabilidad calculable, pero al levantar el cubilete se obtiene exactamente un resultado y la distribución de probabilidad de antes de levantar el cubilete se borra (deja de ser relevante).

[1.9] Si quiero medir en la base $|\pm\rangle$ basta aplicar H antes de medir para que esa base se transforme en $|0\rangle, |1\rangle$

1.5. Fotones

1.5.1. Ondas planas electromagnéticas y fotones

En una solución de las ecs. de Maxwell, tipo onda plana con número de ondas $\boldsymbol{k}$, los vectores $\boldsymbol{k}, \boldsymbol{E}, \boldsymbol{B}$ son ortogonales y $\boldsymbol{B}$ está determinado por $\boldsymbol{k}$ y $\boldsymbol{E}$. Tomando una base adaptada a $\boldsymbol{k}$, $(\hat{e}_H, \hat{e}_V, \boldsymbol{k})$ (triedro ortonormal positivo, H, V corresponden a polarización horizontal y vertical respectivamente)



$$\begin{aligned}\boldsymbol{E}(\boldsymbol{x},t) &= |E_H|\cos(\boldsymbol{k}\boldsymbol{x}-\omega t+\varphi_H)\hat{e}_H + |E_V|\cos(\boldsymbol{k}\boldsymbol{x}-\omega t+\varphi_V)\hat{e}_V \\ &= \mathrm{Re}(\boldsymbol{E}_c e^{-i\omega t+i\boldsymbol{k}\boldsymbol{x}})\end{aligned} \tag{1.29}$$

siendo $\omega = c|\boldsymbol{k}|$ la frecuencia angular y

$$\boldsymbol{E}_c = |E_H|e^{i\varphi_H}\hat{e}_H + |E_V|e^{i\varphi_V}\hat{e}_V \tag{1.30}$$

$\boldsymbol{E}_c$ es un vector complejo ortogonal a $\boldsymbol{k}$, es decir $\boldsymbol{E}_c \in \mathbb{C}^2$

$$\boldsymbol{E}_c = E_H\hat{e}_H + E_V\hat{e}_V = \begin{pmatrix} E_H \\ E_V \end{pmatrix} \tag{1.31}$$

Esto describe un haz de luz clásica, pero vale igualmente para la *función de onda* de un fotón (cuando se atenúa el haz lo suficiente se obtiene el estado de un fotón). Por tanto un fotón con $\boldsymbol{k}$ dado constituye un qubit con respecto de su estado de polarización.

$|H\rangle = |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ fotón linealmente polarizado horizontalmente

$|V\rangle = |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ fotón linealmente polarizado verticalmente

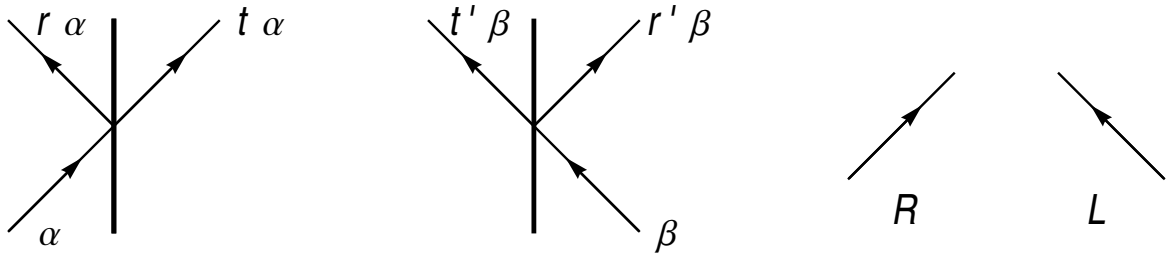
Las polarizaciones con $E_V/E_H \in \mathbb{R}$ son polarizaciones lineales. Una rotación de ángulo θ del plano de polarización corresponde a $\begin{pmatrix} \cos(\theta) & -\mathrm{sen}(\theta) \\ \mathrm{sen}(\theta) & \cos(\theta) \end{pmatrix}$. Esta matriz es $e^{-i\theta\sigma_y}$ (que sería una rotación $R(\hat{e}_y, 2\theta)$ para una partícula de espín $1/2$.)

$\begin{pmatrix} 1 \\ \pm i \end{pmatrix}$ son las polarizaciones circulares dextrógira y levógira, respectivamente (con el convenio más extendido).

Con elementos ópticos adecuados (cristales) se puede realizar cualquier operador U de $\mathrm{U}(2)$ (grupo de matrices unitarias 2×2) sobre el qubit asociado al estado polarización

1.5.2. Elementos ópticos

Un **espejo semirreflectante** o **separador de haz** (que suponemos no absorbente) es un elemento óptico independiente de la polarización. Un haz con amplitud α y polarización cualquiera, se separa en una componente transmitida $t\alpha$ y otra reflejada $r\alpha$ que salen con la misma polarización inicial. t y r son las **amplitudes de transmisión** y **reflexión** respectivamente.



Debe cumplirse cumplirse $|t|^2 + |r|^2 = 1$. Clásicamente (esto es para haz de luz) esta relación se sigue de conservación de la energía (ya que la energía del haz es proporcional $\boldsymbol{E}^2$). Cuánticamente, para amplitud del fotón se sigue de conservación de la probabilidad: si se ponen detectores en los haces salientes se encontrará el fotón en uno y sólo uno de ellos (suponiendo eficiencia perfecta del detector) con probabilidades $|t|^2$ y $|r|^2$, respectivamente (condicionadas a que el fotón llegue al separador). Sólo hay un fotón en todo momento. Se puede denotar

$$\alpha|R\rangle_1 \mapsto t\alpha|R\rangle_2 + r\alpha|L\rangle_2 \tag{1.32}$$

$|R\rangle$, $|L\rangle$, denotan los estados del fotón moviéndose hacia la derecha o la izquierda. Equivalentemente

$$|0\rangle_1 \mapsto t|0\rangle_2 + r|1\rangle_2 \tag{1.33}$$

Nótese que la identificación de $|0\rangle$ a ambos lados es completamente arbitraria, se podían haber intercambiado los nombres de las etiquetas 0, 1 en $|\,\rangle_2$. [1.10] Lo importante es que los dos estados son

[1.10] Otro convenio sería considerar que el fotón se queda en "el mismo estado" cuando se queda al mismo lado del espejo y cambia de estado cuando lo cruza.

incompatibles, es decir ortogonales: si el fotón está en uno de ellos la probabilidad de encontrarlo en el otro al hacer una medida es cero.

Igualmente el fotón puede incidir desde la derecha (moviéndose hacia la izquierda) y se tendrá

$$|L\rangle_1 \mapsto t'|L\rangle_2 + r'|R\rangle_2 \tag{1.34}$$

Los estados $|R\rangle, |L\rangle$ del fotón definen un qubit, esta vez asociado al sentido del movimiento y no a la polarización. En el caso más general el fotón tiene una amplitud de probabilidad α de incidir desde la izquierda y otra β de hacerlo desde la derecha

$$\alpha|R\rangle_1 + \beta|L\rangle_1 \mapsto (t\alpha + r'\beta)|R\rangle_2 + (r\alpha + t'\beta)|L\rangle_2 \tag{1.35}$$

Matricialmente

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \mapsto \begin{pmatrix}t & r'\\ r & t'\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \tag{1.36}$$

La intensidad del haz (en el caso clásico) o la probabilidad (en el caso de fotones) debe conservarse cualquiera que sea el estado $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$. Eso requiere que la matriz $B = \begin{pmatrix}t & r'\\ r & t'\end{pmatrix}$ sea unitaria, $B \in \mathrm{U}(2)$, es decir, $B^\dagger B = I$ de modo que sus columnas forman una base ortonormal de $\mathbb{C}^2$:

$$|t|^2 + |r|^2 = 1, \qquad |t'|^2 + |r'|^2 = 1, \qquad t^* r' + r^* t' = 0. \tag{1.37}$$

La forma más general de B puede expresarse como

$$B = \begin{pmatrix}t & -e^{i\varphi} r^*\\ r & e^{i\varphi} t^*\end{pmatrix} \qquad |t|^2 + |r|^2 = 1, \quad \varphi \in \mathbb{R} \ \ (\text{mód}\, 2\pi)\,. \tag{1.38}$$

Si se impone la condición adicional de que el separador sea simétrico respecto derecha/izquierda (simetría espacial) $t = t'$, $r = r'$, unitaridad requiere que sea de la forma [1.11]

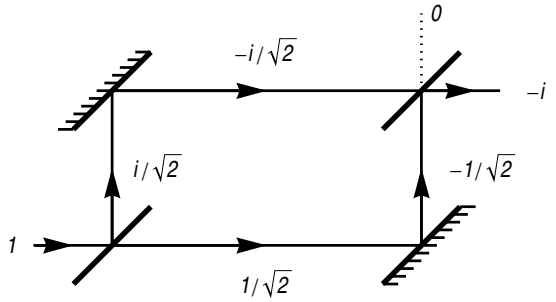
$$B = e^{i\gamma}\begin{pmatrix}\tau & \pm i\rho\\ \pm i\rho & \tau\end{pmatrix}, \qquad \tau, \rho \geq 0, \quad \gamma \in \mathbb{R} \ \ (\text{mód}\, 2\pi), \qquad \tau^2 + \rho^2 = 1. \tag{1.39}$$

La simetría espacial no es un requerimiento obligado para un separador. A menudo se modela como la puerta de Hadamard $H = \dfrac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$ (la matriz H es simétrica pero $t = -t'$ de modo que no es espacialmente simétrica).

[1.11] En efecto, $0 = t^* r + r^* t \implies r/t = -(r/t)^*$ y r/t debe ser imaginario puro. Como consecuencia B también es automáticamente invariante bajo inversión temporal.

1.5.3. Interferómetro de Mach-Zehnder

Por ejemplo en un interferómetro de Mach-Zehnder, con $B = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} = \frac{1}{\sqrt{2}}(I + i\sigma_x)$. En este separador simétrico $|t|^2 = |r|^2 = \dfrac{1}{2}$.



Con el convenio $|0\rangle =$ "dirección horizontal" y $|1\rangle =$ "dirección vertical", y como en todo momento la función de onda del fotón está a lo sumo en dos ramas, el montaje se puede tratar dentro de $\mathbb{C}^2$, es decir, como un qubit.

$$|0\rangle \mapsto \frac{1}{\sqrt{2}}\big(|0\rangle + i|1\rangle\big) \mapsto \frac{1}{\sqrt{2}}\big(-|1\rangle - i|0\rangle\big) \mapsto \frac{1}{2}\Big(-\big(|1\rangle + i|0\rangle\big) - i\big(|0\rangle + i|1\rangle\big)\Big) = -i|0\rangle \quad (1.40)$$

Si se ponen fotodetectores en los dos extremos finales sólo se detectará un fotón en el estado $|0\rangle$ y ninguno en $|1\rangle$. Una descripción clásica probabilística de la acción de los separadores de haz daría una detección de 50 % en cada estado final. [1.12] Es importante notar que la interferencia sólo involucra un único fotón. Si se ponen detectores intermedios, en cada una de las dos ramas antes de llegar al segundo espejo semirreflectante, se detectará el fotón en una y sólo una de ellas.

Los espejos cambian la fase de la onda en media longitud de onda (cambio de signo) su matriz es por tanto

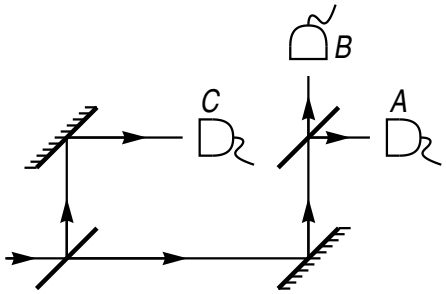
$$M = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \qquad \text{(espejos)} \quad (1.41)$$

El operador unitario del circuito es

$$U = BMB = \frac{1}{\sqrt{2}}(I + i\sigma_x)(-\sigma_x)\frac{1}{\sqrt{2}}(I + i\sigma_x) = -iI \quad (1.42)$$

[1.12] Las fases i oscurecen la idea, que se ve mejor usando espejos semirreflectantes de tipo Hadamard, $|0\rangle \underset{H}{\to} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \underset{H}{\to} |0\rangle$. La interferencia entre caminos es imprescindible para que no reaparezca el estado $|1\rangle$ al final, y eso requiere signos menos en H, que no existen en una descripción probabilística.

Es decir el fotón sale en el mismo estado (con un desfase $-\pi/2$).



La interferencia entre las dos ramas es crucial: si se interpone un fotodetector C en el camino de arriba, el estado final será $\begin{pmatrix} -i/2 \\ -1/2 \\ -i/\sqrt{2} \end{pmatrix}$ (amplitudes en A, B y C)

- $1/4$ de las veces el fotón será detectado en A
- $1/4$ de las veces el fotón será detectado en B
- $1/2$ de las veces el fotón será detectado en C

Supongamos que no se sabe si realmente el detector C está presente o no, sólo tenemos acceso a los fotodetectores A y B.

- Si no se detecta nada en A y B, es que el fotón ha sido absorbido en C.
- Si se detecta en B, es que C está presente. Llegar a B es imposible si no está C para destruir la interferencia de los dos caminos.
- Si se detecta en A no podemos concluir nada sobre la presencia de C.

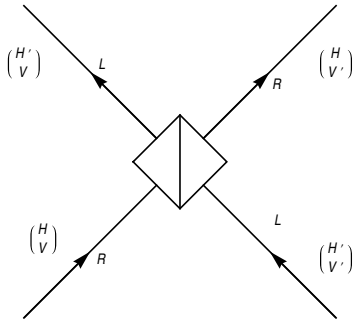
Lo realmente notable es que en algunos casos (cuando el fotón se detecta en B) podemos saber que C está presente incluso aunque este fotodetector no se dispare. Nótese que todos los detectores se suponen perfectos, si les llega un fotón lo detectan seguro. Clásicamente o el fotón va por arriba y C se dispara, o va por abajo y C no se dispara, pero ese caso nos quedamos sin saber si si C está o no. Cuánticamente en un 25 % de los casos se puede saber que está sin que se dispare. [1.13]

[1.13] Ésta es la idea del *Detector de bombas de Elitzur–Vaidman*.

También se pueden ver patrones de interferencia entre los caminos poniendo un elemento desfasador $\begin{pmatrix} 1 & 0 \\ 0 & e^{i\varphi} \end{pmatrix}$, que desfasa uno de los caminos respecto del otro.

1.5.4. Separador por polarización

Otro elemento óptico es el **separador por polarización**. Por ejemplo transmite totalmente la amplitud con polarización horizontal y refleja totalmente la polarizada verticalmente.



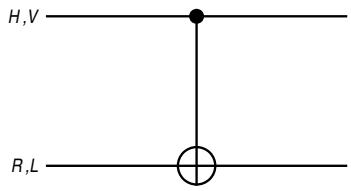
Aquí tenemos un qubit asociado al estado de polarización $\mathscr{H}_p \cong \mathbb{C}^2$, y otro asociado al de dirección del fotón $\mathscr{H}_d \cong \mathbb{C}^2$. En conjunto es el espacio de dos qubits $\mathscr{H} = \mathscr{H}_p \otimes \mathscr{H}_d$. El separador por polarización actúa sobre la base

$$|HR\rangle, |HL\rangle, |VR\rangle, |VL\rangle, \tag{1.43}$$

(equivalentemente $|00\rangle, |01\rangle, |10\rangle, |11\rangle$) con la matriz

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad \begin{matrix} HR \to HR \\ HL \to HL \\ VR \to VL \\ VL \to VR \end{matrix} \tag{1.44}$$

Técnicamente es una puerta CNOT en la que el qubit de polarización es el de control y el de dirección es el qubit objetivo.

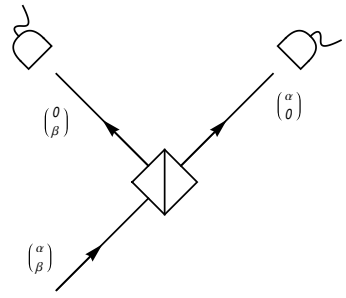


Puesto que todo circuito cuántico se puede construir con CNOT y puertas de un qubit, esto nos dice que con haces de luz clásica se puede hacer computación cuántica. De hecho es así pero es que con ordenadores clásicos se puede hacer computación cuántica (basta codificar la matriz unitaria del circuito). La cuestión es que para manipular $\dim \mathscr{H} = N = 2^n$ (n qubits) se requerirían del orden de N haces clásicos [1.14] Eso es una simulación clásica de un circuito cuántico. El tratamiento cuántico propiamente dicho es usar (del orden de) n fotones. El problema es que la puerta CNOT produce entrelazamiento, por ejemplo,

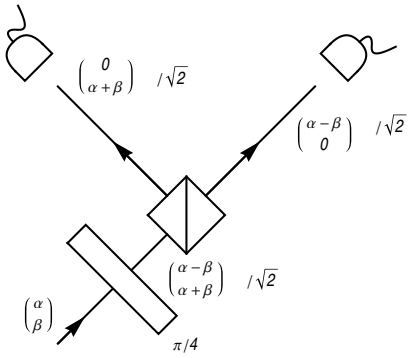
$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle \mapsto \frac{1}{\sqrt{2}}|0\rangle \otimes |0\rangle + \frac{1}{\sqrt{2}}|1\rangle \otimes |1\rangle \tag{1.45}$$

El entrelazamiento requiere interacción entre los fotones y no se puede hacer con elementos ópticos pasivos/lineales como los vistos aquí, que actúan sobre cada fotón por separado [1.15] Sí se puede hacer con cristales no lineales o bien probabilísticamente.

Finalmente, para medir en la base computacional, $|0\rangle, |1\rangle$, el qubit de polarización, $|\psi\rangle = \alpha|H\rangle + \beta|V\rangle$, se puede usar el esquema



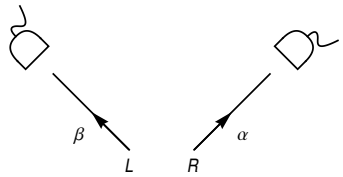
Igualmente se puede medir en la base $|+\rangle$, $|-\rangle$ rotando el plano de polarización



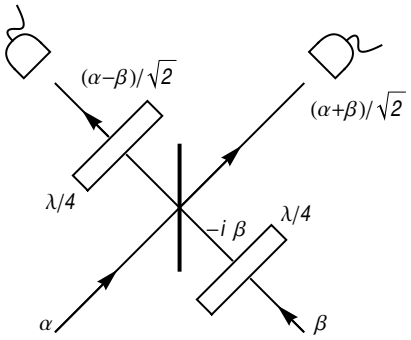
[1.14] Algo menos, se pueden usar pulsos luminosos bien separados.

[1.15] En la terminología de teoría de muchos cuerpos, son operadores de un cuerpo. CNOT requiere operadores de dos cuerpos.

El qubit asociado a la dirección, $|\psi\rangle = \alpha|R\rangle + \beta|L\rangle$, se puede medir en la base computacional, simplemente colocando los detectores adecuadamente



Para medir ese qubit en la base $|\pm\rangle = \frac{1}{\sqrt{2}}(|R\rangle \pm |L\rangle)$ el método es rotar esos estados $|\pm\rangle$ a la base computacional y aplicar la medida anterior. Los cristales $\lambda/4$ introducen una fase $-i$.



1.6. Apéndice: Paralelismo clásico

ADVERTENCIA: La discusión que se presenta aquí es completamente prescindible y el enfoque utilizado podría causar confusión si no se tienen las ideas claras. Se recomienda dejar su lectura para más adelante (o incluso *sine die*).

Como se ha visto en el problema de Deutsch y su solución cuántica, el *paralelismo cuántico* desempeña un papel esencial en la solución; el circuito cuántico produce como estado intermedio una superposición de $|0\rangle$ y $|1\rangle$ del tipo

$$\frac{1}{\sqrt{2}}\Big((-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle\Big) \tag{1.46}$$

que es $|+\rangle$ o $|-\rangle$ según $f(0) = f(1)$ o $f(0) \neq f(1)$.

A nivel clásico también se puede considerar un concepto de paralelismo. Si tenemos un sistema clásico que puede estar en n estados $x = 0, \ldots, n-1$, la situación más general corresponde a permitir una descripción estadística en la que cada valor de x tiene cierta probabilidad p_x y $\sum_x p_x = 1$. Los casos de certeza absoluta $x = x_0$ son un caso particular con $p_x = \delta_{xx_0}$. Por analogía con la notación cuántica, se puede expresar el estado mezcla (estadística) como [1.16]

$$|\rho) = p_0|0) + \cdots + p_{n-1}|n-1) = \sum_{x=0}^{n-1} p_x|x), \qquad p_x \geq 0\,, \qquad \sum_{x=0}^{n-1} p_x = 1 \tag{1.47}$$

o bien $|\rho) = \begin{pmatrix} p_0 \\ \vdots \\ p_{n-1} \end{pmatrix}$.

Si el sistema que está en un estado y cualquiera puede evolucionar (dinámicamente) a cualquiera de los otros estados x con cierta probabilidad condicionada $P(x|y) \equiv P_{xy}$ (que fija la dinámica) la evolución de la mezcla será

$$|\rho) \mapsto |\rho'), \qquad p'_x = \sum_y P_{xy} p_y \tag{1.48}$$

o $|\rho') = \hat{P}|\rho)$. Por supuesto las probabilidades condicionadas deben cumplir

$$P_{xy} \geq 0 \qquad \sum_x P_{xy} = 1 \tag{1.49}$$

[1.16] Debe quedar claro que $|\rho)$ no es un estado cuántico como $|\psi\rangle$. Sí se puede identificar con la matriz densidad $\sum_x p_x |x\rangle\langle x|$.

Es decir la suma de cada columna de la matriz debe ser 1.

Una matriz P que cumpla estas dos propiedades es una **matriz estocástica** por la izquierda. Es una versión clásica del operador cuántico de evolución unitaria. Nótese que si U_{xy} es una matriz unitaria ($|x\rangle$ representa una base ortonormal del sistema anterior en versión cuántica) la matriz con elementos de matriz $P_{xy} = |U_{xy}|^2$ es estocástica, de hecho doblemente estocástica $\sum_x |U_{xy}|^2 = \sum_y |U_{xy}|^2 = 1$. Hay que notar que en general una matriz estocástica P_{xy} no es invertible y ni siquiera tiene que ser cuadrada.

P_{xy} describe una evolución de tipo estocástico (concretamente markoviana) e incluye el caso determinista como caso particular (a saber, P_{xy} se anula a menos que $x = f(y)$ para cierta función f que dicta la dinámica). Si se aplican dos evoluciones sucesivas $\rho \underset{P}{\mapsto} \rho' \underset{Q}{\mapsto} \rho''$, siendo P y Q matrices estocásticas, el efecto es el de aplicar la matriz QP que también es una matriz estocástica. Todo lineal y bastante parecido a la evolución cuántica.

Por otro lado, igual que en el caso cuántico, si se tiene un estado mezcla ρ (por ejemplo se prepara un cubilete con varios dados) se pueden conocer las probabilidades de cada configuración x pero no la configuración concreta. Al hacer una medida (levantar el cubilete y mirar) el estado colapsará a un ρ_0 de tipo $p_x = \delta_{xx_0}$. Se obtendrá uno y sólo uno de los valores con probabilidad no nula, y ρ pasará a estar sólo en ese estado (hasta que la evolución posterior posiblemente introduzca nueva aleatoriedad).

Esta analogía permite plantear un circuito clásico, pero no necesariamente determinista, en el que, aunque el estado inicial sea conocido, eventualmente se pasa a un estado mezcla donde todas las posibilidades x pueden aparecer en estados ρ intermedios con cierta probabilidad, y se calculan las distintas posibilidades en paralelo. Eso es lo que hace $\rho \mapsto P\rho \mapsto QP\rho \mapsto \cdots$.

Tendríamos un paralelismo clásico. Todos los operadores actúan linealmente. La diferencia crucial entre este paralelismo y el cuántico es que en el caso clásico los pesos p_x son números reales positivos o cero (probabilidades) mientras que en el caso cuántico los pesos son números complejos (amplitudes de probabilidad).

Podemos entonces preguntar si el paralelismo clásico se puede utilizar para resolver el problema de Deutsch con una sola llamada a la función f. Veamos que no.

Primero, es fácil ver que el oráculo U_f, $(x,y) \mapsto (x, y \oplus f(x))$ (actúa sobre dos bits) se puede construir usando la puerta $x \mapsto f(x)$ (actúa sobre un bit) una sola vez. Podemos entonces formular el problema como intentar determinar si la función es constante o no usando esta puerta f una sola vez. La puerta actúa linealmente como una matriz sobre el estado ρ que le llega.

Como hay cuatro funciones posibles $f:\{0,1\}\to\{0,1\}$, habrá cuatro puertas asociadas $f_{00},f_{01},f_{10},f_{11}$ (f_{ij} siendo $i=f(0),j=f(1)$). Estas funciones cumplen la identidad $\forall x\ f_{00}(x)+f_{11}(x)=f_{01}(x)+f_{10}(x)$, por tanto también sus puertas (matrices) asociadas cumplen

$$f_{00}+f_{11}=f_{01}+f_{10} \tag{1.50}$$

y también los cuatro estados *finales* de la mezcla estadística (dado que todo es lineal)

$$\rho_{00}+\rho_{11}=\rho_{01}+\rho_{10} \tag{1.51}$$

Es decir, las mezclas de las dos funciones constantes suman lo mismo que las dos equilibradas.

Después de aplicar el circuito se hace una medida en la "base computacional" es decir, se mira el valor que realmente tiene x al final. No se ha prejuzgado cómo es el circuito, puede tener cualquier número de bits iniciales y finales, sólo importa que la puerta f actúa linealmente exactamente una vez.

Aquí se ve el problema: si un cierto resultado $|x)$ tiene una probabilidad *no nula* para la puerta f_{00} o f_{11} también lo tendrá para f_{01} o f_{10} (ya que suman lo mismo). Entonces ese resultado no permitirá distinguir inequívocamente un caso del otro. Clásicamente, no es posible distinguir el tipo de función con una sola llamada.

En la versión cuántica se cumple la misma relación: también los cuatro operadores U_f (oráculos, matrices 4×4) cumplen que la suma de los constantes iguala a la de los equilibrados

$$U_{f_{00}}+U_{f_{11}}=U_{f_{01}}+U_{f_{10}} \tag{1.52}$$

y consecuentemente también para los estados finales de dos qubits en la solución de Deutsch

$$|\psi\rangle=\frac{1}{2}\left(\left((-1)^{f(0)}+(-1)^{f(1)}\right)|0\rangle+\left((-1)^{f(0)}-(-1)^{f(1)}\right)|1\rangle\right)\otimes|-\rangle \tag{1.53}$$

se cumple

$$|\psi_{00}\rangle+|\psi_{11}\rangle=|\psi_{01}\rangle+|\psi_{10}\rangle \tag{1.54}$$

Explícitamente (poniendo los estados como columnas)

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1&0&0&-1\\-1&0&0&1\\0&1&-1&0\\0&-1&1&0\end{pmatrix} \tag{1.55}$$

Sin embargo en este tratamiento cuántico se elude la conclusión negativa del tratamiento clásico. Los resultados posibles para f constante o equilibrada son disjuntos los cual permite distinguir ambos casos; la relación $|\psi_{00}\rangle + |\psi_{11}\rangle = |\psi_{01}\rangle + |\psi_{10}\rangle$ no implica que un resultado no nulo en un caso implique un resultado no nulo en el otro porque de hecho $|\psi_{00}\rangle + |\psi_{11}\rangle = 0$. Esa cancelación se puede conseguir con amplitudes de probabilidad, pero nunca usando sólo probabilidades.

[Caveat: esta demostración no es concluyente. En los argumentos clásico o cuántico se ha supuesto que la puerta f o U_f actúa insertada en el proceso, pero podría actuar en la forma C^n-U_f (un U_f controlado). Habría que extender la demostración a ese caso, aunque parece obvio: la parte de control no puede depender de f, entonces cuando f actúe se aplicará la demostración anterior.].

2. Matriz densidad

La descripción más general de un sistema cuántico incluye no sólo amplitudes de probabilidad sino también probabilidades. Éstos son los estados mezcla que pueden surgir mediante una mezcla estadística de varios estados puros o cuando se estudia un sistema abierto, es decir, un subsistema del sistema cuántico completo.

2.1. Colectividades y subsistemas

2.1.1. Mezcla estadística

Si se tiene una preparación que produce sistemáticamente un estado puro $|\psi\rangle$, al medir un observable A se encontrarán en general distintos valores (correspondientes a su espectro) con un valor esperado $\langle A\rangle_\psi = \langle\psi|A|\psi\rangle$ ($|\psi\rangle$ normalizado). Si en lugar de eso se tienen varias preparaciones de estados puros $|\psi_j\rangle$, $j = 1,\ldots,N$, y se utiliza cada una aleatoriamente con probabilidad p_j (o equivalentemente se forma una colectividad de estados con frecuencias relativas p_j) el valor esperado pasará a ser

$$\langle A\rangle = \sum_{j=1}^{N} p_j \langle A\rangle_{\psi_j} = \sum_{j=1}^{N} p_j \langle\psi_j|A|\psi_j\rangle. \tag{2.1}$$

Se dice que se tiene una **mezcla estadística** de los estados puros $|\psi_j\rangle$ con pesos p_j. Por supuesto debe cumplirse

$$p_j \geq 0\,, \qquad \sum_j p_j = 1\,. \tag{2.2}$$

A la mezcla estadística se le asocia un operador **matriz densidad**, mediante

$$\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j| \tag{2.3}$$

De modo que

$$\langle A\rangle = \mathrm{Tr}(\rho A) \tag{2.4}$$

2.1.1.1. Apartado matemático: traza

Si $\{|j\rangle\}$ es una base de un espacio vectorial V y $A : V \to V$ un operador lineal, la matriz de A se define mediante $A|j\rangle = \sum_k A^k{}_j |k\rangle$. La **traza** de A es [2.1]

$$\mathrm{Tr}(A) = \sum_j A^j{}_j. \tag{2.5}$$

La traza es lineal en A y es una propiedad del operador, es decir, no depende de la base. La traza satisface la denominada **propiedad cíclica**

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA) \qquad \text{(Propiedad cíclica de la traza).} \tag{2.6}$$

En efecto

$$\mathrm{Tr}(AB) = \sum_k (AB)^k{}_k = \sum_{k,j} A^k{}_j B^j{}_k = \sum_{k,j} B^j{}_k A^k{}_j = \sum_j (BA)^j{}_j = \mathrm{Tr}(BA) \tag{2.7}$$

Aquí sólo se requiere que AB sea una matriz cuadrada aunque A y B no lo sean. A puede ser $n \times m$ y B $m \times n$. [2.2]

Si V es un espacio de Hilbert y la base es *ortonormal*, $A^k{}_j = \langle k|A|j\rangle$ y entonces

$$\mathrm{Tr}(A) = \sum_j \langle j|A|j\rangle \qquad \text{(base ortonormal).} \tag{2.8}$$

En este caso

$$\mathrm{Tr}(AB) = \sum_n \langle n|AB|n\rangle = \sum_{n,m} \langle n|A|m\rangle\langle m|B|n\rangle = \sum_{n,m} \langle m|B|n\rangle\langle n|A|m\rangle = \mathrm{Tr}(BA) \tag{2.9}$$

donde se ha usado la descomposición de la identidad $I = \sum_m |m\rangle\langle m|$.

Si se calcula la traza en otra base ortonormal $\overline{|n\rangle} = U|n\rangle$, siendo U unitario,

$$\sum_n \overline{\langle n|} A \overline{|n\rangle} = \sum_n \langle n|U^\dagger A U|n\rangle = \mathrm{Tr}(U^\dagger A U) = \mathrm{Tr}(AUU^\dagger) = \mathrm{Tr}(A). \tag{2.10}$$

También, para el operador $|\psi\rangle\langle\phi|$, que actúa sobre un estado $|\chi\rangle$ cualquiera mediante

$$(|\psi\rangle\langle\phi|)\,|\chi\rangle = \langle\phi|\chi\rangle|\psi\rangle, \tag{2.11}$$

se tiene

$$\mathrm{Tr}(|\psi\rangle\langle\phi|) = \sum_n \langle n|\psi\rangle\langle\phi|n\rangle = \sum_n \langle\phi|n\rangle\langle n|\psi\rangle = \langle\phi|\psi\rangle \tag{2.12}$$

(de nuevo la propiedad cíclica).

[2.1] Si no se dice otra cosa los espacios vectoriales referidos son de dimensión finita. No hay problemas de convergencia.

[2.2] Bajo un cambio de base la matriz A pasa a $A' = SAS^{-1}$, una **transformación de semejanza**. De ahí que la traza no dependa de la base, $\mathrm{Tr}(A') = \mathrm{Tr}(SAS^{-1}) = \mathrm{Tr}(S^{-1}SA) = \mathrm{Tr}(A)$.

Entonces

$$\mathrm{Tr}(\rho A) = \sum_j p_j \mathrm{Tr}\left(|\psi_j\rangle\langle\psi_j|A\right) = \sum_j p_j \langle\psi_j|A|\psi_j\rangle = \langle A\rangle \tag{2.13}$$

Es importante notar que ρ queda invariante si se cambian las fases de los estados normalizados $|\psi_j\rangle$. La mezcla descrita por la matriz densidad es una mezcla estadística o **incoherente** de estados puros, es diferente de una superposición $\sum_j \alpha_j|\psi_j\rangle$ o superposición **coherente** de estados, con pesos α_j complejos. Los p_j son probabilidades, los α_j amplitudes de probabilidad.

A todos los efectos prácticos (medibles) el estado mezcla queda caracterizado por el conjunto de valores esperados de los observables, y eso sólo depende de ρ. La matriz densidad caracteriza completamente el estado físico del sistema. Distintas mezclas que den el mismo ρ son físicamente equivalentes. [2.3]

Un estado mezcla es el estado más general de un sistema cuántico y está unívocamente descrito por su matriz densidad. Generalmente cuando se dice "estado" nos referimos a un estado mezcla. (Los estados puros son un caso particular con $\rho = |\psi\rangle\langle\psi|$.) Una preparación experimental produce un estado mezcla (que puede ser puro) aunque estemos hablando de un único sistema físico. Siempre se puede pensar como que se produce una colectividad y se elige un miembro de esa colectividad al azar.

2.1.2. Sistema abierto, subsistema

Otra forma en la que una matriz densidad puede aparecer es en la descripción del estado de un subsistema dentro de un sistema físico mayor. El sistema completo está compuesto por los subsistemas A y B, y tiene espacio de Hilbert $\mathscr{H} = \mathscr{H}_A \otimes \mathscr{H}_B$. Cuando nos centramos en el subsistema A y B es el ambiente (el resto del universo) nos referimos a A como un **sistema abierto** ya que puede intercambiar energía, materia, e información con B. A y B representan distintos grados de libertad.

2.1.2.1. Apartado matemático: proyección y traza parcial

[2.3] Las medidas sobre el sistema también se expresan mediante valores esperados y por tanto tampoco pueden distinguir entre distintas mezclas con la misma matriz densidad.

Si $\{|n\rangle_A\}$, $\{|m\rangle_B\}$ son bases ortonormales de $\mathscr{H}_A$ y $\mathscr{H}_B$ respectivamente, $\{|n\rangle_A \otimes |m\rangle_B = |n,m\rangle_{AB}\}$ es una base ortonormal de $\mathscr{H}_A \otimes \mathscr{H}_B$, por la propiedad

$$({}_A\langle\psi_1| \otimes {}_B\langle\phi_1|)\,(|\psi_2\rangle_A \otimes |\phi_2\rangle_B) = \langle\psi_1|\psi_2\rangle_A \langle\phi_1|\phi_2\rangle_B. \tag{2.14}$$

Para un estado cualquiera de $\mathscr{H}_A \otimes \mathscr{H}_B$

$$|\psi\rangle_{AB} = \sum_{n,m} \psi_{n,m}\, |n\rangle_A \otimes |m\rangle_B \tag{2.15}$$

las componentes se obtienen como siempre con

$$\psi_{n,m} = ({}_A\langle n| \otimes {}_B\langle m|)\,|\psi\rangle_{AB} \tag{2.16}$$

pero también se puede hacer una proyección parcial

$${}_B\langle\phi_1|\,(|\psi\rangle_A \otimes |\phi_2\rangle_B) = \langle\phi_1|\phi_2\rangle_B |\psi\rangle_A \tag{2.17}$$

Así,

$$|\psi\rangle_{A,B} = \sum_m |\phi_m\rangle_A \otimes |m\rangle_B \qquad \text{con} \qquad |\phi_m\rangle_A = {}_B\langle m|\psi\rangle_{AB} = \sum_n \psi_{n,m}|n\rangle_A \in \mathscr{H}_A \tag{2.18}$$

Y también

$$\||\psi\rangle_{AB}\|^2 = \sum_{n,m} |\psi_{nm}|^2 = \langle\psi|\psi\rangle_{AB} = \sum_m \langle\phi_m|\phi_m\rangle = \sum_m \||\phi_m\rangle\|^2. \tag{2.19}$$

Si X es un operador en $\mathscr{H}_A \otimes \mathscr{H}_B$

$$X = \sum_{n,m,n',m'} X_{nm,n'm'}\, |n,m\rangle\langle n',m'| \tag{2.20}$$

$$\mathrm{Tr}(X) = \sum_{n,m} \langle n,m|X|n,m\rangle = \sum_{n,m} X_{nm,nm} \tag{2.21}$$

Se puede definir una traza parcial, respecto de uno de los dos espacios

$$\mathrm{Tr}_B(X) = \sum_m {}_B\langle m|X|m\rangle_B = \sum_{n,m,n'} X_{nm,n'm}\, |n\rangle_{A\,A}\langle n'| \qquad \text{(operador en } \mathscr{H}_A) \tag{2.22}$$

y lo mismo para $\mathrm{Tr}_A(X)$. De modo que

$$\mathrm{Tr}(X) = \mathrm{Tr}_A(\mathrm{Tr}_B(X)) = \mathrm{Tr}_B(\mathrm{Tr}_A(X)) \tag{2.23}$$

y también

$$\mathrm{Tr}_B\left(|\psi\rangle_A \otimes |\phi\rangle_{B\,A}\langle\psi'| \otimes {}_B\langle\phi'|\right) = \langle\phi'|\phi\rangle_B\, |\psi\rangle_{A\,A}\langle\psi'| \tag{2.24}$$

Un operador de $\mathscr{H}_A \otimes \mathscr{H}_B$ del tipo $A \otimes B$ actúa así

$$A \otimes B\,|\psi\rangle_A \otimes |\phi\rangle_B = (A|\psi\rangle_A) \otimes (B|\phi\rangle_B) \tag{2.25}$$

Entonces

$$\mathrm{Tr}_B(A \otimes B) = \mathrm{Tr}(B)A\,, \qquad \mathrm{Tr}_B(XA \otimes I_B) = \mathrm{Tr}_B(X)A\,, \qquad \mathrm{Tr}_B(A \otimes I_B X) = A\,\mathrm{Tr}_B(X) \qquad \text{(operadores en } \mathscr{H}_A\text{)} \tag{2.26}$$

siendo I_B el operador identidad en $\mathscr{H}_B$. Nótese que $\mathrm{Tr}_B(X)$ y A son dos operadores en $\mathscr{H}_A$ y no conmutan en general.

Supongamos que sólo nos interesan los observables X_A del sistema A. Éstos actúan en el espacio completo como $X_A \otimes I_B$. [2.4] Si el sistema compuesto está en un estado puro $|\psi\rangle_{AB}$, el valor esperado es

$$\langle X_A \rangle = \langle \psi | X_A \otimes I_B | \psi \rangle_{AB} \tag{2.27}$$

Puede expresarse como el valor esperado de una mezcla en $\mathscr{H}_A$,

$$\langle X_A \rangle = \mathrm{Tr}_A(\rho_A X_A) \tag{2.28}$$

siendo

$$\rho_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|_{AB}) \tag{2.29}$$

En efecto,

$$\begin{aligned} \mathrm{Tr}_A(X_A\rho_A) &= \mathrm{Tr}_A\left(X_A\,\mathrm{Tr}_B(|\psi\rangle\langle\psi|)\right) = \mathrm{Tr}_A\left(\mathrm{Tr}_B\left(X_A \otimes I_B|\psi\rangle\langle\psi|\right)\right) \\ &= \mathrm{Tr}\left(X_A \otimes I_B|\psi\rangle\langle\psi|\right) = \langle\psi|X_A \otimes I_B|\psi\rangle_{AB}. \end{aligned} \tag{2.30}$$

ρ_A es la **matriz densidad reducida** del subsistema A.

Aunque el estado AB es puro, desde el punto de vista del subsistema A (es decir, para sus observables) es equivalente a una mezcla estadística de estados (de $\mathscr{H}_A$) con matriz densidad ρ_A. En efecto

$$\rho_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|) = \sum_m {}_B\langle m|\psi\rangle\langle\psi|m\rangle_B = \sum_m |\tilde{\psi}_m\rangle\langle\tilde{\psi}_m|_A \tag{2.31}$$

[2.4] Por ejemplo, para una partícula con espín j, con espacio de Hilbert $L^2(\mathbb{R}^3) \otimes \mathbb{C}^{2j+1}$, el operador momento $\boldsymbol{p}$ es realmente $\boldsymbol{p} \otimes I_{\text{espín}}$ y el operador de espín es $I_{\text{espacio}} \otimes \boldsymbol{S}$, aunque no se escriba.

Los estados $|\tilde{\psi}_m\rangle_A$ no están normalizados. $|\tilde{\psi}_m\rangle_A = \sqrt{p_m}|\psi_m\rangle$ con $p_m = \||\tilde{\psi}_m\rangle\|^2$,

$$\rho_A = \sum_m p_m |\psi_m\rangle\langle\psi_m|_A. \tag{2.32}$$

Nótese que la mezcla concreta equivalente depende de la base ortonormal escogida en $\mathscr{H}_B$, pero no así ρ_A.

Más generalmente, si el estado en AB es ya una mezcla con matriz densidad $\rho_{AB} = \sum_j q_j |\psi_j\rangle\langle\psi_j|_{AB}$, la matriz densidad reducida en el sistema A es

$$\rho_A = \mathrm{Tr}_B(\rho_{AB}) \tag{2.33}$$

ya que reproduce correctamente los valores esperados de operadores X_A de $\mathscr{H}_A$:

$$\mathrm{Tr}_A(X_A\rho_A) = \mathrm{Tr}_A\left(X_A\,\mathrm{Tr}_B(\rho_{AB})\right) = \mathrm{Tr}(X_A\otimes I_B\,\rho_{AB}) = \langle X_A\otimes I_B\rangle. \tag{2.34}$$

La matriz densidad es la descripción más general de un estado cuántico, incluye a los estados puros como caso especial (cuando en la mezcla sólo hay un estado). Contiene a la vez probabilidades y amplitudes de probabilidad. Como veremos *distintas mezclas de estados puros pueden dar lugar a la misma matriz densidad* y toda la información física (valores esperados, medidas) está en ρ independientemente de los estados puros concretos utilizados para construir la mezcla. [2.5]

También muy importante es que la descripción del estado (sea puro o propiamente mezcla) en términos de la matriz densidad es única, a diferencia de la descripción basada en la función de onda que tiene la ambigüedad de una fase global. Por este motivo las formulaciones matemáticas rigurosas de la mecánica cuántica se hacen preferentemente en términos de la matriz densidad.

Como ejemplo de mezcla estadística, consideremos un qubit en un estado mezcla con igual probabilidad para los estados $|0\rangle$ y $|1\rangle$

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}I. \tag{2.35}$$

Por ejemplo un rayo de luz no polarizada, o átomos de espín $1/2$, inicialmente en estado $|+\rangle$ después de pasar por un Stern-Gerlach (en la dirección z) en una medida no selectiva.

[2.5] Esto sugiere que la función de onda no es una magnitud física (una cierta onda material): dado un ρ que no sea estado puro no hay forma de determinar cuál es la mezcla usada para formarlo, sólo ρ mismo (una combinación de probabilidades y amplitudes de probabilidad) es accesible experimentalmente.

Si se quiere calcular el valor esperado de σ_z en esta mezcla, teniendo en cuenta que $\sigma_z|0\rangle = +|0\rangle$ y $\sigma_z|1\rangle = -|1\rangle$, es decir, $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$,

$$\langle \sigma_z \rangle_\rho = \mathrm{Tr}(\sigma_z \rho) = \frac{1}{2}\mathrm{Tr}(\sigma_z) = 0. \tag{2.36}$$

Otro ejemplo, ahora de matriz densidad reducida, es el de un sistema de dos qubits en un estado puro [2.6]

$$|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle_A \otimes |+\rangle_B + \frac{e^{i\varphi}}{\sqrt{2}}|1\rangle_A \otimes |-\rangle_B. \tag{2.37}$$

Podemos calcular el valor esperado del observable σ_z del qubit A, $\sigma_{z,A}$, en este estado

$$\begin{aligned} \langle \sigma_{zA} \rangle &= \langle \psi | \sigma_z \otimes I | \psi \rangle = \frac{1}{2}\big(\langle 0,+| + e^{-i\varphi}\langle 1,-|\big)\sigma_z \otimes I\,\big(|0,+\rangle + e^{i\varphi}|1,-\rangle\big) \\ &= \frac{1}{2}\big(\langle 0|\sigma_z|0\rangle\langle +|+\rangle + e^{i\varphi}\langle 0|\sigma_z|1\rangle\langle +|-\rangle + e^{-i\varphi}\langle 1|\sigma_z|0\rangle\langle -|+\rangle + \langle 1|\sigma_z|1\rangle\langle -|-\rangle\big) = 0. \end{aligned} \tag{2.38}$$

El mismo resultado se obtiene con la matriz densidad reducida

$$\begin{aligned} \rho_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|) &= \frac{1}{2}\mathrm{Tr}_B\Big(\big(|0,+\rangle + e^{i\varphi}|1,-\rangle\big)\big(\langle 0,+| + e^{-i\varphi}\langle 1,-|\big)\Big) \\ &= \frac{1}{2}\Big(\langle +|+\rangle|0\rangle\langle 0| + e^{-i\varphi}\langle +|-\rangle|1\rangle\langle 0| + e^{i\varphi}\langle -|+\rangle|0\rangle\langle 1| + \langle -|-\rangle|1\rangle\langle 1|\Big) \\ &= \frac{1}{2}\big(|0\rangle\langle 0| + |1\rangle\langle 1|\big) = \frac{1}{2}I_A. \end{aligned} \tag{2.39}$$

E igualmente $\rho_B = \frac{1}{2}I_B$. Entonces

$$\langle \sigma_{z,A} \rangle = \mathrm{Tr}_A(\rho_A \sigma_z) = \frac{1}{2}\mathrm{Tr}(\sigma_z) = 0. \tag{2.40}$$

La fase relativa $e^{i\varphi}$, que es relevante para el estado puro $|\psi\rangle$ en el sistema AB, no lo es para observables del subsistema A ya que se cancela en la matriz densidad reducida ρ_A. Pero eso se debe a que los estados $|\pm\rangle_B$ son ortogonales. Entonces cambiar φ equivale a una transformación unitaria U_B sólo en $\mathscr{H}_B$, que no afecta a $\mathscr{H}_A$.

[2.6] Aquí, y en contextos similares, se sobreentiende que φ es real, de modo que $e^{i\varphi}$ es una fase, un número complejo de módulo 1.

2.2. Propiedades de la matriz densidad

Por definición, un operador (lineal) ρ es una matriz densidad cuando puede escribirse como una mezcla,

$$\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|, \qquad \||\psi_j\rangle\| = 1, \qquad p_j \geq 0 \qquad \sum_j p_j = 1 \tag{2.41}$$

Los estados puros no tienen que se ortogonales entre sí (pero sí normalizados). La mezcla puede incluir cualquier número de estados (si $\dim \mathscr{H} \geq 2$) el número puede ser mayor que la dimensión del espacio.

Una matriz densidad satisface ciertas propiedades:

1. $\mathrm{Tr}(\rho) = 1$. Esta propiedad se sigue de

$$\mathrm{Tr}(\rho) = \mathrm{Tr}\left(\sum_j p_j |\psi_j\rangle\langle\psi_j|\right) = \sum_j p_j \langle\psi_j|\psi_j\rangle = \sum_j p_j = 1. \tag{2.42}$$

   Como se ha dicho, al tomar traza parcial, $\rho_A = \mathrm{Tr}_B(\rho_{AB})$ se retiene sólo información sobre A. La expresión $\mathrm{Tr}(\rho) = 1$ retiene la mínima información, a saber, que el estado existe con probabilidad 1. [2.7]

2. Una matriz densidad es hermítica $\rho^\dagger = \rho$. Es inmediato por p_j real y $|\psi\rangle^\dagger = \langle\psi|$.

3. Una matriz densidad es un **operador positivo**, $\rho \geq 0$. Por definición un operador A es (semi-definido) positivo cuando cumple

$$\langle\psi|A|\psi\rangle \geq 0 \qquad \forall|\psi\rangle \in \mathscr{H} \tag{2.43}$$

   Se cumple para la matriz densidad por

$$\langle\psi|\rho|\psi\rangle = \sum_j p_j \langle\psi|\psi_j\rangle\langle\psi_j|\psi\rangle = \sum_j p_j |\langle\psi_j|\psi\rangle|^2 \geq 0. \tag{2.44}$$

   Nótese que un operador positivo es automáticamente hermítico. En efecto, cualquier operador $X : \mathscr{H} \to \mathscr{H}$ puede expresarse en la forma

$$X = A + iB, \qquad A = \frac{1}{2}(X + X^\dagger), \qquad B = \frac{1}{2i}(X - X^\dagger), \tag{2.45}$$

[2.7] Un espacio de dimensión 1 sólo tiene un estado físico y no tiene dinámica, es trivial. Cualquier espacio de Hilbert $\mathscr{H}$ es isomorfo a $\mathscr{H} \otimes \mathscr{H}_0$ siendo $\dim \mathscr{H}_0 = 1$. De modo que la traza respecto de $\mathscr{H}$ se puede ver como una traza parcial. Es análogo a $0! = 1$, cuando no hay ningún factor todavía hay un factor 1 presente.

A es su parte hermítica e iB su parte antihermítica. Si $X \geq 0$ debe cumplirse $B = 0$ y X es puramente hermítico. Por tanto la propiedad 2 es consecuencia de la 3.

También es obvio que el espectro de un operador positivo es no negativo ($\lambda_i \geq 0$). Un operador es positivo si y sólo si es hermítico con espectro no negativo. [2.8]

Por otro lado, cualquier operador ρ positivo y con traza 1 es una matriz densidad, ya que puede llevarse a su forma diagonal [2.9]

$$\rho = \sum_k \lambda_k |k\rangle\langle k| \tag{2.46}$$

siendo $|k\rangle$ una base ortonormal por ρ hermítico, los autovalores $\lambda_k \geq 0$ por ρ positivo y $\sum_k \lambda_k = 1$ por $\mathrm{Tr}(\rho) = 1$. Se concluye que un operador ρ es una matriz densidad si y sólo si es positivo y con traza 1.

Una matriz densidad ρ se puede escribir en la forma (**descomposición espectral** existe para un **operador normal**, es decir, $[A, A^\dagger] = 0$.) [2.10]

$$\rho = \sum_j \lambda_j P_j \tag{2.47}$$

donde λ_j es el espectro y P_j el proyector ortogonal sobre el subespacio con autovalor λ_j,

$$P_j = P_j^\dagger, \qquad P_j P_k = \delta_{jk} P_j, \qquad \sum_j P_j = I. \tag{2.48}$$

La condición $\mathrm{Tr}(\rho) = 1$ equivale a

$$1 = \sum_j n_j \lambda_j, \qquad n_j = \mathrm{rank}(P_j) = \dim(P_j \mathscr{H}) = 1, 2, \ldots \tag{2.49}$$

Se suele definir la **pureza** del estado mezcla ρ como $\mathrm{Tr}(\rho^2)$.

Teorema $\mathrm{Tr}(\rho^2)$ está entre 0 y 1 y es 1 si y sólo si ρ es un estado puro (su mezcla sólo contiene un estado). [2.11]

[2.8] Un operador como $\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$ tiene espectro positivo (a saber, $\{1, 2\}$) pero no es un operador positivo al no ser hermítico.

[2.9] Aquí los autovalores pueden estar repetidos.

[2.10] Aquí cada autovalor distinto aparece sólo una vez.

[2.11] La mínima pureza corresponde a $\rho = \frac{1}{d} I$, por tanto $\frac{1}{d} \leq \mathrm{Tr}(\rho^2) \leq 1$.

Demostración: En efecto, en términos de su descomposición espectral

$$\rho^2 = \sum_j \lambda_j^2 P_j, \qquad \mathrm{Tr}(\rho^2) = \sum_j n_j \lambda_j^2 \tag{2.50}$$

dado que $0 \leq \lambda_j \leq 1$, se deduce que $\lambda_j^2 \leq \lambda_j$, por lo que $\sum_j n_j \lambda_j = 1$ implica $\sum_j n_j \lambda_j^2 \leq 1$.

Por otro lado, si $\mathrm{Tr}(\rho^2) = \sum_j n_j \lambda_j^2 = 1$ se tendrá

$$0 = \sum_j n_j \lambda_j - \sum_j n_j \lambda_j^2 = \sum_j n_j \lambda_j (1-\lambda_j) \implies \forall j \quad \lambda_j(1-\lambda_j) = 0 \implies \lambda_j \in \{0,1\}. \tag{2.51}$$

Para que $\sum_j n_j \lambda_j = 1$, la única posibilidad (salvo ordenación) es $\lambda_1 = 1$ con $n_1 = 1$ y $\lambda_2 = 0$ con $n_2 = \dim(\mathscr{H}) - 1$. Entonces $\rho = P_1 = |\psi\rangle\langle\psi|$ y el estado es puro. □

Como se ha dicho, distintas mezclas $\{p_j, |\psi_j\rangle\}$ pueden producir la misma matriz densidad ρ. De hecho, el número de estados en una mezcla puede ser mucho mayor que la dimensión del espacio, ya que los $|\psi_j\rangle$ no tienen que se ortogonales entre sí. Todas esas mezclas con la misma ρ corresponden al mismo estado mezcla físico.

Como ejemplo, consideremos dos mezclas físicamente equivalentes,

$$\rho = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2}(|+\rangle\langle +| + |-\rangle\langle -|) = \frac{1}{2} I \tag{2.52}$$

Es una mezcla al 50 % de estados $|0\rangle$ y $|1\rangle$, y también de $|+\rangle$ y $|-\rangle$. En este caso, las dos mezclas diagonalizan ρ, esto es posible porque el espectro de $\rho = I/2$ es degenerado. Corresponde a una mezcla máxima, con mínima pureza, todos los estados son equiprobables.

Otro ejemplo, también para un qubit y al 50 %,

$$|a\rangle = \sqrt{\frac{2}{3}}|0\rangle + \sqrt{\frac{1}{3}}|1\rangle, \qquad |b\rangle = \sqrt{\frac{2}{3}}|0\rangle - \sqrt{\frac{1}{3}}|1\rangle, \qquad p_a = p_b = \frac{1}{2} \tag{2.53}$$

Estos estados están normalizados pero no son ortogonales, $\langle a|b\rangle = 1/3$. En este caso

$$\rho = \frac{1}{2}|a\rangle\langle a| + \frac{1}{2}|b\rangle\langle b| = \frac{2}{3}|0\rangle\langle 0| + \frac{1}{3}|1\rangle\langle 1| \tag{2.54}$$

(casualmente diagonaliza en la base computacional). Diagonalizar ρ proporciona una mezcla canónica (única si el espectro es no degenerado).

Una mezcla de matrices densidad, definida como una **combinación lineal convexa** de matrices densidad,

$$\rho = \sum_j p_j \rho_j , \qquad p_j \geq 0 \qquad \sum_j p_j = 1 \tag{2.55}$$

proporciona otra matriz densidad, ya que ρ es manifiestamente positivo y de traza 1. Esta propiedad nos dice que el conjunto de matrices es un **conjunto convexo** dentro del espacio vectorial real de los operadores hermíticos.

En un conjunto convexo, los puntos llamados **extremales** son aquellos que *no pueden* expresarse en la forma [2.12]

$$\rho = \theta\rho_1 + (1-\theta)\rho_2, \qquad 0 < \theta < 1, \qquad \rho_1 \neq \rho_2 \tag{2.56}$$

Los estados puros son los estados extremales en el conjunto de matrices densidad: Es claro que los estados mezcla con más de un estado son no extremales (por definición de no extremal). Menos obvio es que un estado puro es necesariamente extremal, es decir, que nunca se puede obtener mezclando estados distintos. Esto se demostrará más adelante como corolario del teorema de la pág. 18. Según ese teorema, sólo estados que estén en la imagen de ρ pueden formar parte de una mezcla que produzca ρ.

2.3. Estados puros y mezcla de un qubit

La matriz densidad más general para un qubit puede expresarse como[2.13]

$$\rho = \frac{1}{2}(I + \boldsymbol{n}\cdot\boldsymbol{\sigma}). \tag{2.57}$$

donde $\boldsymbol{n}$ es un cierto vector. Se ha usado que $\sigma_\mu = (I, \boldsymbol{\sigma})$ es una base de las matrices 2×2. Teniendo en cuenta que $\mathrm{Tr}(\boldsymbol{\sigma}) = 0$, el coeficiente $1/2$ en I garantiza que $\mathrm{Tr}(\rho) = 1$. Como $\boldsymbol{\sigma}^\dagger = \boldsymbol{\sigma}$, la condición $\rho^\dagger = \rho$ requiere que $\boldsymbol{n}$ sea un vector real.

La propiedad $\sigma_i\sigma_j = \delta_{ij} + i\varepsilon_{ijk}\sigma_k$ (sumado sobre k) implica que $\mathrm{Tr}(\sigma_\mu\sigma_\nu) = 2\delta_{\mu\nu}$, $\mu,\nu = 0,1,2,3$. Entonces

$$\boldsymbol{n} = \mathrm{Tr}(\rho\boldsymbol{\sigma}). \tag{2.58}$$

[2.12] En la bola de Bloch los puntos extremales son los de la esfera. Sin embargo en un cubo macizo en $\mathbb{R}^3$ los puntos extremales son los vértices pero no las caras. Para un punto extremal estar en la frontera es una condición necesaria pero no suficiente.

[2.13] Toda matriz 2×2 se puede expresar como $a_0 I + \boldsymbol{a}\cdot\boldsymbol{\sigma}$, $a_\mu \in \mathbb{C}$.

Falta ver qué condiciones impone $\rho \geq 0$. Explícitamente

$$\rho = \frac{1}{2}\begin{pmatrix} 1+n_z & n_x - in_y \\ n_x + in_y & 1-n_z \end{pmatrix} \tag{2.59}$$

y se deduce $\det(\rho) = (1 - \|\boldsymbol{n}\|^2)/4$. Si los autovalores de ρ son λ_+ y $\lambda_- = 1 - \lambda_+$, con $\lambda_\pm \geq 0$ por $\rho \geq 0$, el determinante $\lambda_+\lambda_-$ debe ser no negativo. Eso requiere $\|\boldsymbol{n}\| \leq 1$.

Podemos entonces representar $\boldsymbol{n}$ de manera unívoca como puntos interiores o sobre la esfera de Bloch del qubit. Los estados puros corresponden a puntos de la superficie $\|\boldsymbol{n}\| = 1$. En efecto, ρ es un estado puro si y sólo si su espectro es $\{0,1\}$ lo que equivale a $\det(\rho) = 0$ y a $\|\boldsymbol{n}\| = 1$. Además el vector unitario $\boldsymbol{n}$ asociado a un qubit es el mismo que se introdujo en el Tema 1.1. Como puede comprobarse fácilmente, para ρ construido con

$$\boldsymbol{n} = (\cos(\phi)\,\mathrm{sen}(\theta), \mathrm{sen}(\phi)\,\mathrm{sen}(\theta), \cos(\theta)) \tag{2.60}$$

y $|\psi\rangle = \begin{pmatrix} \cos(\theta/2) \\ e^{i\phi}\,\mathrm{sen}(\theta/2) \end{pmatrix}$ se cumple

$$\rho|\psi\rangle = +|\psi\rangle \tag{2.61}$$

El estado más general de un bit es una mezcla estadística, permitiendo probabilidades p_0 y p_1 para los valores 0 y 1. Ese estado clásico se puede introducir en el formalismo cuántico, correspondiendo a un estado mezcla $p_0|0\rangle\langle 0| + p_1|1\rangle\langle 1|$. Entonces el estado más general de un qubit corresponde a un punto en la bola de Bloch mientras que el estado más general de un bit se puede identificar con un punto del diámetro de la esfera de Bloch que une el polo norte con el sur ($|0\rangle$ con $|1\rangle$).

2.4. Descomposición de Schmidt

Sea $\mathscr{H}_A \otimes \mathscr{H}_B$ un sistema bipartito y $\{|i\rangle_A\}$ y $\{|j\rangle_B\}$ bases ortonormales de $\mathscr{H}_A$ y $\mathscr{H}_B$ respectivamente. El estado puro más general será de la forma

$$|\psi\rangle = \sum_{i=1}^{d_A}\sum_{j=1}^{d_B} \psi_{ij}\,|i\rangle_A \otimes |j\rangle_B \tag{2.62}$$

La descomposición de Schmidt consiste en que (para cada $|\psi\rangle$ normalizado dado) es posible elegir sendas bases ortonormales en $\mathscr{H}_A$ y $\mathscr{H}_B$ de modo quede una expresión diagonal

$$|\psi\rangle = \sum_{\alpha=1}^{r} \sqrt{\lambda_\alpha}|u_\alpha\rangle_A \otimes |v_\alpha\rangle_B \qquad \lambda_\alpha > 0, \qquad \sum_{\alpha=1}^{r}\lambda_\alpha = 1\,. \tag{2.63}$$

2.4.0.1. Apartado matemático: Descomposición en valores singulares

Toda matriz compleja M $m \times n$ puede expresarse como

$$M = UDV \tag{2.64}$$

siendo U unitaria $m \times m$, V unitaria $n \times n$ y D diagonal no negativa $m \times n$, es decir, $D_{ij} = 0$ si $i \neq j$ y $D_{ii} \geq 0$. Estos D_{ii} son los denominados **valores singulares** de la matriz M. D es única ordenando las entradas de la diagonal en orden decreciente. Se cumple que $r := \text{rank}(M) = \text{rank}(D)$ es el número de D_{ii} no nulos, $r \leq \text{mín}(m,n)$.

Si M es cuadrada, se puede escribir $M = WH$ siendo W unitaria y H semidefinida positiva, basta tomar $H = V^{\dagger}DV$ y $W = UV$. (Análogamente $M = H'W$). Esa es la **descomposición polar** de M.

Utilizando la descomposición en valores singulares de ψ_{ij} visto como una matriz $d_A \times d_B$, [2.14]

$$\psi_{ij} = \sum_{\alpha=1}^{r} \sqrt{\lambda_\alpha} U_{i\alpha} V_{j\alpha}, \qquad \lambda_\alpha > 0, \quad U, V \text{ unitarias} \tag{2.65}$$

se obtiene la forma

$$|\psi\rangle = \sum_{\alpha=1}^{r} \sqrt{\lambda_\alpha} |u_\alpha\rangle_A \otimes |v_\alpha\rangle_B \tag{2.66}$$

denominada **descomposición de Schmidt** de $|\psi\rangle$. Aquí

$$|u_\alpha\rangle_A = \sum_{i=1}^{d_A} U_{i\alpha} |i\rangle_A, \qquad |v_\alpha\rangle_B = \sum_{j=1}^{d_B} V_{j\alpha} |j\rangle_B, \tag{2.67}$$

Estas son bases ortonormales. Nótese que estas bases no son fijas, dependen del estado ψ.

r es el rango de la matriz ψ_{ij}. $r \leq \text{mín}(d_A, d_B)$. Supongamos que $d_A \leq d_B$. La descomposición de Schmidt implica que por muy grande que sea $\mathscr{H}_B$, un estado $|\psi\rangle$ dado sólo puede acceder a un subespacio de $\mathscr{H}_B$ no mayor que $\mathscr{H}_A$.

Si se considera la matriz densidad reducida de cualquiera de los espacios se obtiene

$$\rho_A = \text{Tr}_B(|\psi\rangle\langle\psi|) = \sum_{\alpha=1}^{r} \lambda_\alpha |u_\alpha\rangle\langle u_\alpha|_A, \qquad \rho_B = \text{Tr}_A(|\psi\rangle\langle\psi|) = \sum_{\alpha=1}^{r} \lambda_\alpha |v_\alpha\rangle\langle v_\alpha|_B \,. \tag{2.68}$$

Esto implica:

[2.14] La matriz V aquí es la traspuesta de la matriz V en (2.64)

i) Las bases ortonormales son precisamente las que diagonalizan las matrices densidad reducidas. Los estados $|u_\alpha\rangle$ (ídem $|v_\alpha\rangle$) son únicos salvo fase si los autovalores λ_α son no degenerados. Si hay degeneración distintas elecciones están relacionadas unitariamente.

ii) Las matrices reducidas de los dos subsistemas *tienen los mismos autovalores no nulos* (y con la misma multiplicidad). [2.15] El número de autovalores nulos diferirá si las dimensiones de $\mathscr{H}_A$ y $\mathscr{H}_B$ son distintas.

Para sistemas multipartitos $\mathscr{H}_1 \otimes \cdots \otimes \mathscr{H}_n$ con $n \geq 3$ ya no existe una descomposición diagonal del tipo

$$|\psi\rangle = \sum_{\alpha=1}^{r} \sqrt{\lambda_\alpha} |u_\alpha^{(1)}\rangle_1 \otimes \cdots \otimes |u_\alpha^{(n)}\rangle_n \tag{2.69}$$

para todos los estados porque el conjunto de estados diagonales (incluso eligiendo las bases ortonormales como se desee en cada $\mathscr{H}_j$) es demasiado pequeño para llenar todo el espacio total. Sin embargo, eligiendo un orden en los factores $\mathscr{H}_j$, es posible hacer una descomposición canónica, aplicando Schmidt al sistema bipartito $\mathscr{H}_1 \otimes (\mathscr{H}_2 \otimes \cdots \otimes \mathscr{H}_n)$. Luego, en cada vector del segundo espacio, a $\mathscr{H}_2 \otimes (\mathscr{H}_3 \otimes \cdots \otimes \mathscr{H}_n)$, y así sucesivamente.

2.5. Purificación, matrices densidad reducidas, subsistemas

Sea una matriz densidad ρ_A en un espacio $\mathscr{H}_A$. Un estado $|\Phi\rangle_{AB} \in \mathscr{H}_A \otimes \mathscr{H}_B$ es una **purificación** de ρ_A, cuando ρ_A es la matriz reducida de $|\Phi\rangle_{AB}$, es decir

$$\rho_A = \mathrm{Tr}_B \left(|\Phi\rangle\langle\Phi|_{AB} \right). \tag{2.70}$$

Dado un ρ_A, no es difícil construir una purificación eligiendo $\mathscr{H}_B$ suficientemente grande. Dada cualquier mezcla concreta que produzca ρ_A:

$$\rho_A = \sum_{j=1}^{N} p_j |\psi_j\rangle\langle\psi_j| \tag{2.71}$$

(los $|\psi_j\rangle$ están normalizados pero no tienen que ser ortogonales) tomamos $\dim \mathscr{H}_B \geq N$ y $\{|v_j\rangle_B\}$ una base ortonormal arbitraria. Entonces el estado

$$|\Phi\rangle_{AB} = \sum_{j=1}^{N} \sqrt{p_j} |\psi_j\rangle_A \otimes |v_j\rangle_B \tag{2.72}$$

[2.15] Esto ya no es necesariamente cierto para estados mezcla ρ_{AB}, así por ejemplo, para dos qubits y $\rho_{AB} = |0\rangle\langle 0| \otimes \frac{1}{2} I_B$.

es una purificación de ρ_A, como se comprueba inmediatamente.

Un mismo ρ_A admite muchas purificaciones. Dadas dos purificaciones $|\Phi_1\rangle_{AB}$ y $|\Phi_2\rangle_{AB}$, en sendos espacios auxiliares $\mathscr{H}_{B1}$ y $\mathscr{H}_{B2}$, siempre se puede considerar que usan el mismo espacio auxiliar $\mathscr{H}_B$ (basta considerar $\mathscr{H}_B = \mathscr{H}_{B1} + \mathscr{H}_{B2}$).

Teorema (Hughston, Jozsa, Wooters (HJW)) Dos estados $|\Phi_1\rangle_{AB}$ y $|\Phi_2\rangle_{AB}$ en $\mathscr{H}_A \otimes \mathscr{H}_B$ son purificaciones de una misma matriz densidad ρ_A si y sólo si $\exists U_B$ unitario en $\mathscr{H}_B$ tal que

$$|\Phi_2\rangle_{AB} = I_A \otimes U_B |\Phi_1\rangle_{AB} \tag{2.73}$$

Demostración: Es inmediato que si $|\Phi_2\rangle_{AB} = I_A \otimes U_B|\Phi_1\rangle_{AB}$ entonces sus matrices reducidas son iguales:

$$\mathrm{Tr}_B(|\Phi_2\rangle\langle\Phi_2|) = \mathrm{Tr}_B(U_B|\Phi_1\rangle\langle\Phi_1|U_B^\dagger) = \mathrm{Tr}_B(U_B^\dagger U_B|\Phi_1\rangle\langle\Phi_1|) = \mathrm{Tr}_B(|\Phi_1\rangle\langle\Phi_1|). \tag{2.74}$$

Supongamos ahora que $|\Phi_1\rangle_{AB}$ y $|\Phi_2\rangle_{AB}$ son estados tales que $\mathrm{Tr}_B(|\Phi_1\rangle\langle\Phi_1|) = \mathrm{Tr}_B(|\Phi_2\rangle\langle\Phi_2|)$. Entonces, llevándolos a su forma de Schmidt:

$$\begin{aligned} |\Phi_1\rangle_{AB} &= \sum_{j=1}^{r} \sqrt{\lambda_j}|u_j\rangle_A \otimes |v_j\rangle_B, \\ |\Phi_2\rangle_{AB} &= \sum_{j=1}^{r} \sqrt{\lambda_j}|u_j\rangle_A \otimes |w_j\rangle_B \end{aligned} \tag{2.75}$$

donde

$$\rho_A = \sum_{j=1}^{r} \lambda_j |u_j\rangle\langle u_j|_A\,, \tag{2.76}$$

y $|v_j\rangle_B$ y $|w_j\rangle_B$ son bases ortonormales de $\mathscr{H}_B$. Dado que dos bases ortonormales están unívocamente relacionadas por un operador unitario, se tendrá

$$|w_j\rangle_B = U_B|v_j\rangle_B \qquad U_B \text{ operador unitario en } \mathscr{H}_B \tag{2.77}$$

En consecuencia $|\Phi_2\rangle_{AB} = (I_A \otimes U_B)|\Phi_1\rangle_{AB}$. □

Nota: En la demostración se ha usado que la base $\{|u_j\rangle\}$ de vectores propios de ρ_A es la misma (quizá salvo fase) en las dos purificaciones. Si el espectro de ρ_A es degenerado hay una ambigüedad en la elección de los vectores propios, pero si $|\Phi_2\rangle$ usara una base distinta $\{|u'_j\rangle\}$ (para un autovalor dado) estaría relacionada unitariamente con $\{|u_j\rangle\}$ y es fácil ver que equivale a usar la misma base redefiniendo los $|w_j\rangle$.

En particular, el teorema implica que si ρ_A se desarrolla como una cierta mezcla concreta, $\rho_A = \sum_{j=1}^{N} p_j |\psi_j\rangle\langle\psi_j|$, cualquier purificación $|\Phi\rangle_{AB}$ de ρ_A admite la forma

$$|\Phi\rangle_{AB} = \sum_{j=1}^{N} \sqrt{p_j} |\psi_j\rangle_A \otimes |v_j\rangle_B \tag{2.78}$$

con unos $|v_j\rangle_B$ ortonormales adecuados. Esto es así porque $|\Phi'\rangle_{AB} = \sum_{j=1}^{N} \sqrt{p_j} |\psi_j\rangle_A \otimes |j\rangle_B$ es una purificación, siendo $|j\rangle_B$ una base ortonormal fija, y entonces $|\Phi\rangle_{AB} = I_A \otimes U_B |\Phi'\rangle_{AB}$ para cierto U_B, y $|v_j\rangle_B = U_B |j\rangle_B$ produce el resultado pedido.

2.6. Descomposición en estados puros e interpretación de colectividades

Teorema Sean dos conjuntos de N vectores cualesquiera $\{|\tilde{\psi}_i\rangle\}_{i=1}^{N}$ y $\{|\tilde{\phi}_j\rangle\}_{j=1}^{N}$ en un espacio de Hilbert complejo. Entonces la igualdad

$$\sum_{i=1}^{N} |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i| = \sum_{j=1}^{N} |\tilde{\phi}_j\rangle\langle\tilde{\phi}_j| \tag{2.79}$$

se satisface si y sólo si los vectores están relacionados por una transformación unitaria, es decir, existe una matriz unitaria U, $N \times N$, tal que

$$|\tilde{\psi}_i\rangle = \sum_{j=1}^{N} U_{ij} |\tilde{\phi}_j\rangle. \tag{2.80}$$

Nótese que: 1) Los vectores no tienen que estar normalizados ni ser ortogonales. 2) Algunos de los vectores pueden ser 0, por tanto si la sumas tienen distinto número de sumandos se pueden añadir los vectores nulos necesarios para igualar el número en ambos lados de (2.79).

Demostración: .

($\Longleftarrow$) Si $|\tilde{\psi}_i\rangle = \sum_{j=1}^{N} U_{ij} |\tilde{\phi}_j\rangle$, usando $\sum_i U_{ij} U_{ij'}^* = \delta_{jj'}$

$$\sum_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i| = \sum_i \left(\sum_j U_{ij} |\tilde{\phi}_j\rangle \right) \left(\sum_{j'} U_{ij'}^* \langle\tilde{\phi}_{j'}| \right) = \sum_j \sum_{j'} \delta_{jj'} |\tilde{\phi}_j\rangle\langle\tilde{\phi}_{j'}| = \sum_j |\tilde{\phi}_j\rangle\langle\tilde{\phi}_j| \tag{2.81}$$

($\Longrightarrow$) Sea $\rho \equiv \sum_{i=1}^{N} |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|$. Si $\mathscr{H}_A$ denota el espacio de Hilbet de ρ, purificamos este estado en un espacio $\mathscr{H}_A \otimes \mathscr{H}_B$ usando una base ortonormal $\{|j\rangle_B\}$. Tenemos dos purificaciones

$$|\Psi\rangle_{AB} = \sum_i |\tilde{\psi}_i\rangle_A \otimes |i\rangle_B \qquad |\Phi\rangle_{AB} = \sum_j |\tilde{\phi}_j\rangle_A \otimes |j\rangle_B \,. \tag{2.82}$$

Por el teorema HJW, $|\Psi\rangle_{AB} = I_A \otimes U_B |\Phi\rangle_{AB}$, para cierto U_B unitario y

$$U_B|j\rangle_B = \sum_i U_{ij}|i\rangle_B \tag{2.83}$$

donde U_{ij} es una matriz unitaria. Entonces

$$|\Psi\rangle_{AB} = I_A \otimes U_B \sum_j |\tilde{\phi}_j\rangle_A \otimes |j\rangle_B = \sum_j |\tilde{\phi}_j\rangle_A \otimes \sum_i U_{ij}|i\rangle_B \implies |\tilde{\psi}_i\rangle = \sum_i U_{ij}|\tilde{\phi}_j\rangle \,. \tag{2.84}$$

□

Podemos verificar el teorema con el ejemplo visto anteriormente en la pág. 10 para la mezcla

$$|a\rangle = \begin{pmatrix} \sqrt{\frac{2}{3}} \\ \sqrt{\frac{1}{3}} \end{pmatrix} \qquad |b\rangle = \begin{pmatrix} \sqrt{\frac{2}{3}} \\ -\sqrt{\frac{1}{3}} \end{pmatrix} \tag{2.85}$$

con $p_a = p_b = \frac{1}{2}$,

$$|\tilde{a}\rangle = \sqrt{p_a}|a\rangle = \begin{pmatrix} \sqrt{\frac{1}{3}} \\ \sqrt{\frac{1}{6}} \end{pmatrix} \qquad |\tilde{b}\rangle = \sqrt{p_b}|b\rangle = \begin{pmatrix} \sqrt{\frac{1}{3}} \\ -\sqrt{\frac{1}{6}} \end{pmatrix} \tag{2.86}$$

$\rho = |\tilde{a}\rangle\langle\tilde{a}| + |\tilde{b}\rangle\langle\tilde{b}| = |\tilde{0}\rangle\langle\tilde{0}| + |\tilde{1}\rangle\langle\tilde{1}|$ con

$$|\tilde{0}\rangle = \begin{pmatrix} \sqrt{\frac{2}{3}} \\ 0 \end{pmatrix} \qquad |\tilde{1}\rangle = \begin{pmatrix} 0 \\ \sqrt{\frac{1}{3}} \end{pmatrix} \tag{2.87}$$

ambos conjuntos está relacionados unitariamente:

$$|\tilde{a}\rangle = \sqrt{\frac{1}{2}}|\tilde{0}\rangle + \sqrt{\frac{1}{2}}|\tilde{1}\rangle, \qquad |\tilde{b}\rangle = \sqrt{\frac{1}{2}}|\tilde{0}\rangle - \sqrt{\frac{1}{2}}|\tilde{1}\rangle. \tag{2.88}$$

En este caso la matriz unitaria es $U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, la puerta de Hadamard.

Teorema Sea ρ una matriz densidad expresada como dos mezclas concretas distintas:

$$\rho = \sum_{i=1}^{N} p_i |\psi_i\rangle\langle\psi_i| = \sum_{j=1}^{M} q_j |\phi_j\rangle\langle\phi_j|, \qquad p_i, q_j > 0. \tag{2.89}$$

Entonces los dos conjuntos de estados subtienden el mismo subespacio lineal:

$$\text{span}\{|\psi_i\rangle\}_{i=1}^{N} = \text{span}\{|\phi_j\rangle\}_{j=1}^{M} = \text{ran}(\rho). \tag{2.90}$$

Aquí $\text{ran}(\rho) = \rho\mathscr{H}$ denota la imagen del operador ρ. El teorema implica $N, M \geq \text{rank}(\rho) \equiv \dim\text{ran}(\rho)$ (el rango del operador ρ o de su matriz en una base).

Demostración: Diagonalizamos ρ

$$\rho = \sum_{k=1}^{r} \lambda_k |k\rangle\langle k|, \qquad \lambda_k > 0, \qquad \langle k|k'\rangle = \delta_{kk'}, \qquad r = \text{rank}(\rho). \tag{2.91}$$

Es claro que $\text{ran}(\rho) = \text{span}\{|k\rangle\}_{k=1}^{r}$ ya que $\rho|k\rangle = \lambda_k|k\rangle$ y $\lambda_k \neq 0$.

Se trata entonces de probar que $|k\rangle$ es combinación lineal de los $|\psi_i\rangle$ y viceversa (y entonces igual para los $|\phi_j\rangle$). Pero eso es consecuencia inmediata del teorema previo. □

Nota: Si en una suma $A \equiv \sum_{i=1}^{N} c_i |\psi_i\rangle\langle\psi_i|$ se permiten pesos *reales pero negativos* (A no es una matriz densidad) ya no es necesariamente cierto que $|\psi_i\rangle \in \text{ran}(A)$. [2.16]

Corolario Si $\rho = |\psi\rangle\langle\psi|$ y se expresa como mezcla de $|\psi_j\rangle$ debe cumplirse $|\psi_j\rangle \propto |\psi\rangle$ y ρ es extremal.

2.7. Fidelidad

[2.17]

2.7.1. Definición de fidelidad

Entre estados puros, el **solapamiento** $|\langle\phi|\psi\rangle|^2$ indica hasta qué punto $|\phi\rangle$ y $|\psi\rangle$ se parecen, y a esta magnitud se le denomina **fidelidad** (de un estado como copia del otro). [2.18] Nótese que en

[2.16] Por ejemplo $A = 2|0\rangle\langle 0| + |1\rangle\langle 1| - |+\rangle\langle +| - |-\rangle\langle -| = |0\rangle\langle 0|$, no contiene $|1\rangle$ ni $|\pm\rangle$ en la imagen.

[2.17] Elaborado a partir de las notas de Preskill, 2.6.1

[2.18] A veces se define sin el cuadrado.

mecánica cuántica dos estados puros con vectores-estado distintos son parcialmente iguales (no se pueden distinguir inequívocamente) a menos que sean ortogonales.

Para estados mezcla, la fidelidad se suele definir como

$$F(\rho,\sigma) \equiv \left(\mathrm{Tr}\left[\sqrt{\rho^{1/2}\sigma\rho^{1/2}} \right] \right)^2 . \tag{2.92}$$

[Dado un operador A no negativo, $A = \sum_a a|a\rangle\langle a|$, con $a \geq 0$, $A^{1/2} \equiv \sum_a \sqrt{a}|a\rangle\langle a|$, siendo $\sqrt{a}$ la raíz no negativa.]

Esta definición es tal que cuando $\rho = |\psi\rangle\langle\psi|$, $F = \langle\psi|\sigma|\psi\rangle$, y por tanto cuando además $\sigma = |\phi\rangle\langle\phi|$, $F = |\langle\phi|\psi\rangle|^2$. También satisface que $F = 1 \iff \rho = \sigma$, y $F = 0 \iff \mathrm{ran}(\rho) \perp \mathrm{ran}(\sigma)$.

Además, aunque la definición (2.92) no es manifiestamente simétrica, $F(\rho,\sigma) = F(\sigma,\rho)$.

En efecto, notemos primero que si $\{\lambda_n\}$ es el espectro de $\rho^{1/2}\sigma\rho^{1/2}$, $F = (\sum_n \sqrt{\lambda_n})^2$. Por otro lado si A es una matriz cuadrada, las matrices $AA^\dagger$ y $A^\dagger A$ tienen el mismo espectro: Usando la descomposición en valores singulares $A = U_1 D U_2$ y $A^\dagger = U_2^\dagger D U_1^\dagger$ (D diagonal no negativa y $U_{1,2}$ unitarias). Implica que

$$AA^\dagger = U_1 D^2 U_1^\dagger, \qquad A^\dagger A = U_2^\dagger D^2 U_2, \tag{2.93}$$

son matrices semejantes con el mismo espectro que D^2. Este resultado implica que $\rho^{1/2}\sigma\rho^{1/2} = (\rho^{1/2}\sigma^{1/2})(\sigma^{1/2}\rho^{1/2})$ tiene el mismo espectro que $(\sigma^{1/2}\rho^{1/2})(\rho^{1/2}\sigma^{1/2}) = \sigma^{1/2}\rho\sigma^{1/2}$, y por tanto, $F(\rho,\sigma) = F(\sigma,\rho)$.

2.7.2. Teorema de Uhlmann

Teorema (Uhlmann) La fidelidad entre ρ y σ es el máximo de las fidelidades entre purificaciones de ρy σ.

Demostración: En efecto, el espacio de ρ y σ es $\mathscr{H}_A$, y $\rho = \sum_j \lambda_j |u_j\rangle\langle u_j|_A$. Los vectores propios de ρ, $\{|u_j\rangle_A\}$, forman una base ortonormal de $\mathscr{H}_A$. La purificación más general de ρ es ($\mathscr{H}_A \to \mathscr{H}_A \otimes \mathscr{H}_B$)

$$|\Phi_\rho\rangle = \sum_j \sqrt{\lambda_i}|u_j\rangle_A \otimes |v_j\rangle_B \tag{2.94}$$

siendo $\{|v_j\rangle_B\}$ una cierta base ortonormal de $\mathscr{H}_B$ (que depende de la purificación). Entonces, también

$$|\Phi_\rho\rangle = \rho^{1/2} \otimes I \sum_i |u_i\rangle_A \otimes |v_i\rangle_B . \tag{2.95}$$

Si $\{|i\rangle_A\}$ y $\{|i\rangle_B\}$ son bases ortonormales cualesquiera (fijas),

$$|u_i\rangle_A = U_\rho |i\rangle_A, \qquad |v_i\rangle_B = V|i\rangle_B, \tag{2.96}$$

siendo U_ρ y V operadores unitarios en $\mathscr{H}_A$ y $\mathscr{H}_B$, respectivamente. Así

$$|\Phi_\rho\rangle = \rho^{1/2} U_\rho \otimes V|\tilde{\Phi}\rangle, \qquad |\tilde{\Phi}\rangle \equiv \sum_i |i\rangle_A \otimes |i\rangle_B. \tag{2.97}$$

Nótese que el estado $|\tilde{\Phi}\rangle$ no está normalizado a 1, y que es independiente de ρ y de la purificación. Del mismo modo

$$|\Phi_\sigma\rangle = \sigma^{1/2} U_\sigma \otimes W|\tilde{\Phi}\rangle. \tag{2.98}$$

La purificación más general se obtiene al variar V y W. Ahora se puede usar la propiedad [2.19]

$$I \otimes V|\tilde{\Phi}\rangle = \sum_i |i\rangle_A \otimes \left(\sum_j V_{ji}|j\rangle_B\right) = \sum_j \left(\sum_i (V^T)_{ij}|i\rangle_A\right) \otimes |j\rangle_B = V^T \otimes I|\tilde{\Phi}\rangle, \tag{2.99}$$

que implica

$$|\Phi_\rho\rangle = \left(\rho^{1/2} U_\rho \otimes I\right)(I \otimes V)|\tilde{\Phi}\rangle = \left(\rho^{1/2} U_\rho \otimes I\right)\left(V^T \otimes I\right)|\tilde{\Phi}\rangle = \rho^{1/2} U_\rho V^T \otimes I|\tilde{\Phi}\rangle. \tag{2.100}$$

Igualmente

$$|\Phi_\sigma\rangle = \sigma^{1/2} U_\sigma W^T \otimes I|\tilde{\Phi}\rangle. \tag{2.101}$$

La fidelidad entre las dos purificaciones es

$$\begin{aligned} |\langle\Phi_\sigma|\Phi_\rho\rangle|^2 &= |\langle\tilde{\Phi}|W^{T\dagger} U_\sigma^\dagger \sigma^{1/2}\rho^{1/2} U_\rho V^T \otimes I|\tilde{\Phi}\rangle|^2 \\ &= \left|\mathrm{Tr}(W^{T\dagger} U_\sigma^\dagger \sigma^{1/2}\rho^{1/2} U_\rho V^T)\right|^2 = \left|\mathrm{Tr}(U\sigma^{1/2}\rho^{1/2})\right|^2, \end{aligned} \tag{2.102}$$

siendo $U \equiv U_\rho V^T W^{T\dagger} U_\sigma^\dagger$ (un operador unitario en $\mathscr{H}_A$).

Por otro lado, todo operador A se puede descomponer como $A = U'H$ siendo U' unitario y H semidefinido positivo (descomposición polar) $H = \sqrt{A^\dagger A}$. Entonces,

$$|\langle\Phi_\sigma|\Phi_\rho\rangle|^2 = |\,\mathrm{Tr}(UU'\sqrt{\rho^{1/2}\sigma\rho^{1/2}}|^2 = |\sum_n \sqrt{\lambda_n}\langle n|UU'|n\rangle|^2, \tag{2.103}$$

[2.19] El operador V^T está definido por $\langle i|V^T|j\rangle = \langle j|V|i\rangle$. Al igual que V^*, V^T depende de la base, pero en todo caso si V es unitario y la base $\{|i\rangle\}$ es ortonormal, V^T y V^* son también unitarios.

siendo λ_n y $|n\rangle$ los autovalores y autoestados de $\rho^{1/2}\sigma\rho^{1/2}$. El máximo se obtiene eligiendo V o W de modo que $UU' = I$, y corresponde a $|\sum_n \sqrt{\lambda_n}|^2 = F(\rho,\sigma)$.

En conclusión

$$F(\rho,\sigma) = \max_{V,W} |\langle\Phi_\sigma|\Phi_\rho\rangle|^2. \tag{2.104}$$

Corolario $F(\rho_{AB},\sigma_{AB}) \leq F(\rho_A,\sigma_A)$, siendo $\rho_A = \mathrm{Tr}_B(\rho_{AB})$, ídem σ.

En efecto: toda purificación de ρ_{AB} lo es de ρ_A (ídem σ) pero no al revés. El conjunto de purificaciones de ρ_A es mayor (o no menor) que el de ρ_{AB}, por tanto, el máximo al pasar de AB a A sólo puede aumentar.

Obviamente, al perder información descartando B $(AB \to A)$ se pierde poder de discriminación y es más difícil distinguir entre los estados ρ y σ y la fidelidad aumenta.

3. Entrelazamiento

3.1. Definición de entrelazamiento

3.1.1. Entrelazamiento en estados puros

Consideremos un sistema bipartito $\mathscr{H} = \mathscr{H}_A \otimes \mathscr{H}_B$. Se dice que un estado puro $|\psi\rangle_{AB}$ (que podemos suponer normalizado) es **separable** cuando admite una factorización

$$|\psi\rangle_{AB} = |\phi\rangle_A \otimes |\chi\rangle_B. \tag{3.1}$$

En este caso es inmediato que los estados normalizados $|\phi\rangle_A$ y $|\chi\rangle_B$ son únicos salvo fase. Por definición, cualquier estado puro que no sea separable está **entrelazado**. La definición se extiende de modo análogo al caso multipartito. Los estados puros entrelazados son superposición de separables (ya que existe una base separable $|j\rangle_A \otimes |k\rangle_B$).

Por ejemplo, en el caso de dos qubits, los estados

$$\begin{aligned} &|0\rangle_A|0\rangle_B, \\ &\frac{1}{2}(|0\rangle_A|0\rangle_B - |0\rangle_A|1\rangle_B + |1\rangle_A|0\rangle_B - |1\rangle_A|1\rangle_B) = |+\rangle_A|-\rangle_B \end{aligned} \tag{3.2}$$

son separables.

Por otro lado los denominados **estados de Bell**:

$$\begin{aligned} |\Psi_\pm\rangle_{AB} &= \frac{1}{\sqrt{2}}(|00\rangle_{AB} \pm |11\rangle_{AB}) \\ |\Phi_\pm\rangle_{AB} &= \frac{1}{\sqrt{2}}(|01\rangle_{AB} \pm |10\rangle_{AB}) \end{aligned} \tag{3.3}$$

están entrelazados. Los estados de Bell forman una base ortonormal del sistema de dos qubits, alternativa a la base computacional que es separable.

Si se identifica $|0\rangle = |\uparrow\rangle$ y $|1\rangle = |\downarrow\rangle$ para partículas de espín $1/2$, $|\Phi_-\rangle$ es el estado singlete de espín $|J = 0, M = 0\rangle$. Que $|\Phi_-\rangle$ sea singlete implica que

$$U \otimes U|\Phi_-\rangle \propto |\Phi_-\rangle \qquad \forall U \in \mathrm{U}(2). \tag{3.4}$$

(El mismo operador U en ambos sectores A y B, es decir, con la misma matriz en sendas bases). El factor de proporcionalidad es $\det(U)$ y es una fase que es 1 cuando $U \in \mathrm{SU}(2)$. Equivalentemente, si $\{|\alpha\rangle, |\beta\rangle\}$ es cualquier base ortonormal de $\mathbb{C}^2$,

$$\frac{1}{\sqrt{2}}(|\alpha\beta\rangle - |\beta\alpha\rangle) \propto |\Phi_-\rangle. \tag{3.5}$$

Producir un estado de Bell a partir de la base computacional es fácil usando la puerta CNOT.

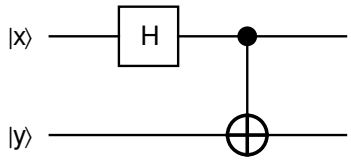


Figura 3.1: Circuito para producir estados de Bell.

Para un estado puro bipartito hay un criterio simple para ver si es separable o entrelazado: basta llevarlo a su forma canónica de Schmidt. Es separable si y sólo si hay exactamente un sumando ($r = 1$ en ec. (2.66)). El criterio se puede extender a estados puros multipartitos. Los estados de Bell en (3.3) ya están en su forma de Schmidt.

Otro criterio (que se deduce del anterior): el estado puro bipartito es separable sí y sólo si las matrices densidad reducidas ρ_A y ρ_B son estados puros [3.1] Por ejemplo para los estados de Bell es inmediato que

$$\rho_A = \frac{1}{2} I_A, \qquad \rho_B = \frac{1}{2} I_B. \tag{3.6}$$

Por el teorema HJW (ec. (2.73)) al ser los estados de Bell purificaciones de un mismo ρ_A (ídem ρ_B) se deduce que se puede pasar de uno a otro mediante una transformación unitaria en uno de los espacios, por ejemplo

$$|\Psi_-\rangle = I \otimes X |\Phi_-\rangle\,. \tag{3.7}$$

En un sistema bipartito tal que $\dim \mathscr{H}_A = \dim \mathscr{H}_B =: d < \infty$, un estado puro se dice que está **máximamente entrelazado** cuando $\rho_A = \frac{1}{d} I_A$ (y entonces $\rho_B = \frac{1}{d} I_B$). [3.2] Los estados de Bell están máximamente entrelazados.

[3.1] Basta que lo sea una, una implica la otra para estados AB puros.

[3.2] Se deduce que no hay estados máximamente entrelazados en espacios de Hilbert de dimensión infinita, ya que $\rho_A \propto I$ sería una matriz densidad impropia (en el mismo sentido que las ondas planas o las deltas de Dirac son estados impropios, no normalizables).

Un estado $\rho \propto I$ corresponde a mínima información, es la mezcla uniforme sobre todos los estados de $\mathscr{H}$. Entonces, el entrelazamiento máximo implica que valores esperados $\langle O_A \rangle$ son un promedio de O_A sobre todo $\mathscr{H}_A$ (ídem B) sin ninguna otra información específica sobre el estado $|\psi\rangle_{AB}$. Así cualquier información (un observable o una medida) que involucre *exclusivamente* al primer qubit de un estado de Bell (ídem el segundo qubit, más generalmente observables separables $O_A \otimes O_B$) no será capaz de aportar ninguna información para distinguir un estado de Bell de otro (un estado máximamente entrelazado de otro). Esto significa que cuando el entrelazamiento es máximo también la información está máximamente repartida (o compartida) en el estado puro entre los dos sectores A y B. Por separado no tienen ninguna información. Esto es todo lo contrario de lo que ocurre para estados separables.

Observación:

Hay que notar que ser separable no es una propiedad intrínseca de un estado $|\psi\rangle$ en un espacio $\mathscr{H}$. Un espacio de Hilbert es homogéneo, todos los estados normalizados son equivalentes a priori. Es la factorización de $\mathscr{H}$ como $\mathscr{H}_A \otimes \mathscr{H}_B$, que es una estructura adicional, la que define que un estado sea separable.

Así, si permito cualquier operador unitario U_{AB} en el espacio de dos qubits, no hay impedimento para obtener $U_{AB}|\Psi_+\rangle = |0\rangle|1\rangle$, en cambio el entrelazamiento no se puede eliminar (reducir) mediante operadores separables, $U_A \otimes U_B$, tales operadores no cambian la matrices densidad reducidas de estados máximamente entrelazados.

Matemáticamente, un $\mathscr{H}$ cuya dimensión sea un número compuesto admite múltiples factorizaciones inequivalentes. Físicamente la factorización debe venir de la naturaleza física de $\mathscr{H}_A$ y $\mathscr{H}_B$, por ejemplo dos grados de libertad distintos, o que A y B sean sistemas dinámica o localmente distinguibles.

Otra observación es que aunque los estados entrelazados son simplemente superposición de separables, sus propiedades pueden ser muy distintas, como ya se ha visto en el caso de máximo entrelazamiento.

3.1.2. Entrelazamiento en estados mezcla

Por definición, un estado mezcla ρ_{AB} es **separable** cuando *admite* la descomposición

$$\rho_{AB} = \sum_j p_j \rho_{Aj} \otimes \rho_{Bj}, \qquad p_j \geq 0, \quad \sum_j p_j = 1. \tag{3.8}$$

En otro caso el estado mezcla está **entrelazado**.

Las definiciones de separable y entrelazado para estados puros y mezcla son consistentes: un estado puro es separable si y sólo si es separable como estado mezcla.

Un caso especial de estado mezcla separable es un estado mezcla **factorizable** o **tipo producto**, es decir, de la forma $\rho_A \otimes \rho_B$. Los estados puros separables son ejemplos de estados mezcla factorizables. [3.3]

Un estado separable es una mezcla estadística de factorizables. [3.4] En un estado mezcla factorizable las medidas en A y B no están correlacionadas. En un estado mezcla separable hay correlaciones estadísticas (el estado ρ_{Aj} está correlacionado con el ρ_{Bj}). Estas correlaciones son clásicas, basadas en probabilidades. En un estado entrelazado hay correlaciones cuánticas, esto es, basadas en amplitudes de probabilidad, que pueden ser más fuertes que las alcanzables clásicamente, como se verá. En sentido vago, el entrelazamiento es la parte de la información de un estado que no se puede describir clásicamente (esto es, con correlaciones y probabilidades).

Los estados mezcla separables son mezcla estadística de estados puros separables (basta desarrollar ρ_{Aj} y ρ_{Bj} como mezcla de estados puros en A y en B respectivamente).

También es evidente que los estados separables forman un conjunto convexo: mezcla de separables da separable.

No es así para los entrelazados (el conjunto complementario a uno convexo generalmente no es convexo). Al mezclar estados entrelazados puede obtenerse un estado separable. Por ejemplo una mezcla por igual de los cuatro estados de Bell produce $\rho_{AB} = \frac{1}{4} I_{AB}$, que es separable: en efecto el mismo resultado se obtiene al mezclar por igual cualquier base ortonormal, sea separable o no, por tanto es mezcla de estados puros separables y es separable. Al mezclar la sutil información asociada al entrelazamiento puede borrarse.

De hecho determinar si un estado mezcla está entrelazado o no, o en qué medida lo está, es un problema matemático considerablemente más complicado que para estados puros. [3.5] Así por ejemplo un estado de dos qubits

$$\rho_{AB} = p|\Psi_+\rangle\langle\Psi_+| + (1-p)\frac{1}{4} I_{AB}, \qquad 0 \leq p \leq 1 \tag{3.9}$$

está entrelazado si $p > 0$ y es separable si $p = 0$. Debe concluirse (correctamente) que pasa de "poco

[3.3] Quizá una nomenclatura más correcta sería llamar separables a los factorizables y "no entrelazados" a los separables, pero aquí usamos las denominación más extendida.

[3.4] Mezcla estadística quiere decir basada en probabilidades, y no en amplitudes de probabilidad.

[3.5] De hecho es un problema NP-difícil.

entrelazado" a "muy entrelazado" cuando p pasa de 0 a 1. Sin embargo

$$\rho_A = \frac{1}{2} I_A, \qquad \rho_B = \frac{1}{2} I_B, \quad \forall p \tag{3.10}$$

lo cual indica que, a diferencia de los estados puros, la matrices densidad reducidas de estados mezcla no tienen suficiente información sobre el grado de entrelazamiento.

3.2. Desigualdades de Bell

El entrelazamiento o "correlación cuántica" explota que la descripción cuántica se basa en amplitudes de probabilidad, además de probabilidades. Al ser una extensión de la descripción puramente clásica que sólo usa probabilidades, debe esperarse que cuánticamente se puedan obtener correlaciones más fuertes que las que se podrían obtener a nivel puramente clásico. Eso es correcto y una de las formas de manifestarse es a través de la violación de las **desigualdades de Bell**.

3.2.1. Argumento EPR

El teorema de Bell está motivado por el **argumento EPR** (Einstein-Podolsky-Rosen) que aquí vemos en su versión Bohm con qubits. Tenemos dos partículas A y B de espín $1/2$, con espacio de Hilbert $\mathscr{H}_A \otimes \mathscr{H}_B$, en estado singlete de espín

$$|\psi\rangle = \frac{1}{2}\left(|0\rangle_A \otimes |1\rangle_B - |1\rangle_A \otimes |0\rangle_B\right). \tag{3.11}$$

Las partículas se han producido juntas (digamos por desintegración de otra partícula de espín 0) y luego se han separado, y su estado de espín se mide en puntos del espacio tiempo x_A^μ y x_B^μ espacialmente separados, $(\boldsymbol{x}_A - \boldsymbol{x}_B)^2 > c^2(t_A - t_b)^2$.

En x_A se mide el espín en la dirección $\hat{\boldsymbol{a}}$, es decir, se mide el observable $\hat{A} = \boldsymbol{\sigma}_A \cdot \hat{\boldsymbol{a}}$, y en x_B se mide $\hat{B} = \boldsymbol{\sigma}_B \cdot \hat{\boldsymbol{b}}$. Estos observables tienen espectro $\{\pm 1\}$. Como son operadores en espacios distintos ($\mathscr{H}_A$ y $\mathscr{H}_B$) conmutan y se pueden medir a la vez sobre $|\psi\rangle$.

Sean $a, b \in \{\pm 1\}$ los valores obtenidos al medir $\hat{A}$ y $\hat{B}$. Estos valores varían de sistema a sistema preparado en el mismo estado $|\psi\rangle$. Según la mecánica cuántica $\langle a\rangle = \langle b\rangle = 0$ (en el estado singlete) es decir sale ± 1 al 50 %, pero los resultados a y b están correlacionados. El cálculo de $\langle\psi|\hat{A}\hat{B}|\psi\rangle$ da

$$\langle ab\rangle = -\hat{\boldsymbol{a}} \cdot \hat{\boldsymbol{b}} \tag{3.12}$$

En particular, si ambas partes miden el espín en la misma dirección, $\hat{\boldsymbol{a}} = \hat{\boldsymbol{b}}$, se debe obtener siempre $ab = -1$, es decir, $(a,b) = (+1,-1)$ o $(a,b) = (-1,+1)$. Esto no está en discusión y se verifica empíricamente.

El argumento EPR es que las medidas en x_A y x_B no pueden influirse mutuamente ya que esos puntos están separados espacialmente y las interacciones físicas son locales, no hay acción a distancia. Entonces la única forma de entender la correlación entre los resultados a y b es que sus valores estén ya codificados en el estado físico antes de hacer las medidas. Pero esa información no está en la función de onda (una misma función de onda es compatible con distintos valores (a,b) que pueden resultar al hacer las medidas) y por tanto la mecánica cuántica (la función de onda) sería una descripción **incompleta** del sistema físico. Se podría haber argumentado de entrada que la función de onda no da una descripción completa porque cada vez que se mide $\hat{A}$ sale un resultado aleatorio, pero eso se podría explicar por interacción y perturbación incontrolable del espín con el aparato de medida. El punto clave del argumento EPR es que ya no se puede invocar esa perturbación incontrolable porque hay claramente una correlación (no aleatoria) y no se puede atribuir a una influencia mutua después de (o causada por) las medidas, ya que son medidas locales, espacialmente separadas.

Otro aspecto del argumento EPR, en el mismo sentido de incompletitud de la mecánica cuántica, es que el espín de A se puede medir en la dirección z, y obtener digamos $+1$, entonces se sabe que si el espín de B midiera también en la dirección z se obtendría -1 con seguridad. Si en vez de eso se mide en la dirección x, B "tendrá a la vez" (para un mismo sistema físico) valores definidos del espín en las direcciones z y x (el valor que tendría con seguridad si se midiera σ_z y el que se obtiene de hecho al medir σ_x).[3.6] La mecánica cuántica no puede predecir ambos al no ser observables compatibles y en consecuencia sería incompleta.

Cuando se prepara un estado físico en un estado cuántico puro $|\psi\rangle$, cada vez se obtienen valores distintos en las medidas de los observables (con distribución de probabilidad dependiente del estado y del observable). El argumento EPR sugiere entonces que las distintas realizaciones físicas obtenidas con el mismo $|\psi\rangle$ no son realmente idénticas (la mecánica cuántica supone que sí lo son) sino que difieren por variables ocultas que son las que explicarían los distintos resultados cada vez. Las variables ocultas fluctuarían de caso a caso de acuerdo con una distribución de probabilidad, dependiente de $|\psi\rangle$.

El problema es que en ese caso todos los valores de todos los observables, compatibles o no, estarían definidos a la vez para un sistema físico, una vez fijados los valores de las variables ocultas. Como vamos a ver, esa hipótesis implica restricciones entre los valores esperados, las desigualdades de Bell. Estas desigualdades no se satisfacen en mecánica cuántica y de hecho no se verifican experimentalmente, al contrario, los experimentos confirman las predicciones de la mecánica cuántica.

[3.6] Este es un denominado argumento **contrafáctico**: se argumenta usando escenarios que realmente no han llegado a producirse.

3.2.2. Teorema de Bell

Establecemos primero la desigualdad de Bell. Lo vemos en su versión CHSH (Clauser, Horne, Shimony, Holt)

Para un cierto sistema físico (no necesariamente el experimento EPR) [3.7] tenemos cuatro observables A_1, A_2, B_1 y B_2, que sólo pueden tomar valores $\{\pm 1\}$. Se tiene un montaje experimental que permite medir a la vez uno de los dos A_i y uno de los dos B_j (a elegir, hay cuatro elecciones) sobre un estado físico (posiblemente mezcla). [3.8] Con esa medida se construye el observable producto A_iB_j. Haciendo repetidas copias del estado físico y repetidas medidas de los cuatro observables A_iB_j, se pueden determinar los cuatro valores esperados $\langle A_iB_j\rangle$ en ese estado.

Proposición En una hipotética teoría de *variables ocultas locales*, debe cumplirse la denominada **desigualdad de Bell**

$$-2 \leq \langle A_1B_1\rangle + \langle A_1B_2\rangle + \langle A_2B_1\rangle - \langle A_2B_2\rangle \leq 2\,. \tag{3.13}$$

Demostración: En una teoría de variables ocultas locales, en cada copia del estado físico los observables tomarán ciertos valores concretos $a_1, a_2, b_1, b_2 \in \{\pm 1\}$, que pueden variar de copia a copia y que son los que se observan al hacer medidas. Esto es así aunque en cada medida sólo se tome nota de uno de los a_i y uno de los b_j. Los cuatro valores tendrán una cierta distribución de probabilidad conjunta $P(a_1, a_2, b_1, b_2)$. Así por ejemplo

$$\langle A_1B_2\rangle = \sum_{a_1,a_2,b_1,b_2} a_1b_2\,P(a_1,a_2,b_1,b_2). \tag{3.14}$$

Nótese que puede ocurrir que los observables A_1 y A_2 (ídem B_1 y B_2) no sean compatibles. Eso implicaría que si después de medir A_1 se mide A_2 se puede obtener un valor a_2' distinto del que se habría obtenido si se midiera A_2 directamente, pero eso no invalida el razonamiento que se está haciendo. Se puede entonces definir la variable aleatoria

$$S = A_1B_1 + A_1B_2 + A_2B_1 - A_2B_2, \tag{3.15}$$

y la desigualdad de Bell a probar equivale a

$$-2 \leq \langle S\rangle \leq 2. \tag{3.16}$$

[3.7] No presuponemos que la teoría cuántica es válida, por lo que no se usa aquí jerga cuántica. En todo caso el estado puede ser mezcla estadística de otros estados.

[3.8] En realidad en un mundo descrito por variables ocultas locales, los estados puros cuánticos son ellos mismos estados mezcla de estados superpuros, en los que las variables ocultas y por tanto todos los observables, tienen valores bien definidos.

En cada copia del estado físico S sólo puede tomar valores

$$s = a_1 b_1 + a_1 b_2 + a_2 b_1 - a_2 b_2 \in \{\pm 2\}. \tag{3.17}$$

Eso se puede ver reescribiendo s en la forma

$$s = a_1(b_1 + b_2) + a_2(b_1 - b_2) \tag{3.18}$$

se ve s sólo puede tomar los valores ± 2. En efecto, si $b_1 = b_2$, $s = 2a_1 b_1 = \pm 2$, y si $b_1 = -b_2$, $s = 2a_2 b_1 = \pm 2$, también. Entonces al hacer un promedio sobre distintos casos se obtendrá $-2 \leq \langle S \rangle \leq 2$. □

Esta demostración supone que los observables tienen unos valores que son los que quedan registrados al medir. La demostración se puede extender al caso en que la medida misma introduzca cierta aleatoriedad en el resultado, pero de manera local: Sea $P(a,b|A_i,B_j)$ la distribución de probabilidad de obtener (a,b) al medir (A_i,B_j). Hay cuatro tales distribuciones de probabilidad al variar i, j y

$$\langle A_i B_j \rangle = \sum_{a,b} ab\, P(a,b|A_i,B_j)\,. \tag{3.19}$$

Que la medida de A_i introduzca cierta aleatoriedad en el valor a (ídem B_j y b) pero medir B_j no afecte al valor de a (ídem A_i y b) es lo que define que las **variables ocultas** sean **locales**. Corresponde a que las cuatro distribuciones sean de la forma

$$P(a,b|A_i,B_j) = \int P(a|A_i,\lambda) P(b|B_j,\lambda) \rho(\lambda) d\lambda\,. \tag{3.20}$$

Aquí $\rho(\lambda) \geq 0$ y $\int \rho(\lambda) d\lambda = 1$, es la distribución de probabilidad de las variables ocultas λ. [3.9] Las variables ocultas son locales porque la distribución de probabilidad de a depende sólo de A_i y λ pero no de B_j (ídem cambiando A por B). [3.10]

No todos los conjuntos de cuatro distribuciones de probabilidad son compatibles con la forma (3.20) (de hecho la teoría cuántica no la satisface). Si las funciones $P(a,b|A_i,B_j)$ satisfacen (3.20) entonces los valores esperados (3.19) satisfacen la desigualdad de Bell. De hecho tales valores esperados coinciden con los que se obtendrían de la distribución de probabilidad conjunta

$$P(a_1,a_2,b_1,b_2) = \int P(a_1|A_1,\lambda) P(a_2|A_2,\lambda) P(b_1|B_1,\lambda) P(b_2|B_2,\lambda)\, \rho(\lambda) d\lambda\,, \tag{3.21}$$

[3.9] Aquí λ puede representar varios parámetros. Por simplificar la notación lo tratamos como si fuera sólo uno.

[3.10] Que a no esté totalmente fijado por A_i y λ (ídem b y B_j) sino que hay una distribución de probabilidad indica que hay más variables ocultas locales, pero no afecta al argumento.

y por tanto se aplica la demostración previa.

Teorema (de Bell) Una teoría cuántica con espacio $\mathscr{H}$ tal que $\dim \mathscr{H} \geq 4$, no puede describirse mediante variables ocultas locales.

La demostración se basa en probar que existen estados ρ y observables $\hat{A}_1$, $\hat{A}_2$, $\hat{B}_1$, y $\hat{B}_2$ con espectro $\{\pm 1\}$, y tales que $[\hat{A}_i, \hat{B}_j] = 0 \; \forall i, j$, para los que se viola la desigualdad de Bell.

Se puede definir el operador

$$\hat{S} = \hat{A}_1\hat{B}_1 + \hat{A}_1\hat{B}_2 + \hat{A}_2\hat{B}_1 - \hat{A}_2\hat{B}_2. \tag{3.22}$$

Según la desigualdad de Bell se debería cumplir $|\langle \hat{S} \rangle| \leq 2$ en cualquier estado, es decir, $\hat{S}$ debe tener espectro en $[-2, 2]$.

Basta encontrar contraejemplos de la desigualdad en el caso de estados puros de dos qubits, ya que el espacio $\mathbb{C}^2 \otimes \mathbb{C}^2$ está contenido en $\mathscr{H}$ si su dimensión es al menos 4. Veamos un par de contraejemplos en espacios de dos qubits:

Para el estado $|\psi\rangle = |\Phi_-\rangle$ (singlete de espín) como se ha dicho

$$\langle ab \rangle = -\hat{\boldsymbol{a}} \cdot \hat{\boldsymbol{b}} \tag{3.23}$$

entonces para las elecciones

$$\hat{\boldsymbol{a}}_1 = \hat{\boldsymbol{b}}_1 = (0,0,1), \qquad \hat{\boldsymbol{a}}_2 = \frac{1}{\sqrt{2}}(1,0,1), \qquad \hat{\boldsymbol{b}}_2 = \frac{1}{\sqrt{2}}(-1,0,1), \tag{3.24}$$

$$-\langle \hat{S} \rangle = \hat{\boldsymbol{a}}_1 \cdot \hat{\boldsymbol{b}}_1 + \hat{\boldsymbol{a}}_1 \cdot \hat{\boldsymbol{b}}_2 + \hat{\boldsymbol{a}}_2 \cdot \hat{\boldsymbol{b}}_1 - \hat{\boldsymbol{a}}_2 \cdot \hat{\boldsymbol{b}}_2 = 1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} - 0 = 1 + \sqrt{2} > 2 \tag{3.25}$$

Otro ejemplo

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + \omega|11\rangle), \qquad \omega = e^{i\pi/4} = \frac{1+i}{\sqrt{2}}. \tag{3.26}$$

Tomamos $\hat{A}_1 = \sigma_x^A,\ \hat{A}_2 = \sigma_y^A,\ \hat{B}_1 = \sigma_x^B,\ \hat{B}_2 = \sigma_y^B$:

$$\begin{aligned}
\langle \hat{A}_1 \hat{B}_1 \rangle &= \frac{1}{2}(\langle 00| + \omega^* \langle 11|)(|11\rangle + \omega|00\rangle) = \frac{1}{2}(\omega + \omega^*) = \mathrm{Re}(\omega) = \frac{1}{\sqrt{2}} \\
\langle \hat{A}_1 \hat{B}_2 \rangle &= \frac{1}{2}(\langle 00| + \omega^* \langle 11|)(i|11\rangle - i\omega|00\rangle) = \frac{1}{2}(-i\omega + i\omega^*) = \mathrm{Im}(\omega) = \frac{1}{\sqrt{2}} \\
\langle \hat{A}_1 \hat{B}_2 \rangle &= \langle \hat{A}_2 \hat{B}_1 \rangle = \frac{1}{\sqrt{2}} \\
\langle \hat{A}_2 \hat{B}_2 \rangle &= \frac{1}{2}(\langle 00| + \omega^* \langle 11|)(-|11\rangle - \omega|00\rangle) = \frac{1}{2}(-\omega - \omega^*) = -\mathrm{Re}(\omega) = -\frac{1}{\sqrt{2}}
\end{aligned} \tag{3.27}$$

Por tanto, en este estado $\langle \hat{S} \rangle = 4\frac{1}{\sqrt{2}} = 2\sqrt{2} > 2$. La violación de la desigualdad es mayor que antes, de hecho como veremos éste es el valor máximo de $|\langle \hat{S} \rangle|$ (desigualdad de Tsirelson).

Observaciones:

1) Para ver una contradicción de la teoría cuántica con variables ocultas locales, quizá se podría haber considerado simplemente la identidad $(\sigma_x \sigma_y)^2 = -1$, en el espacio de un qubit (ya que cualquiera que sean los "auténticos valores" de σ_x y σ_y el resultado debería ser positivo). Esto no es correcto porque σ_x y σ_y no conmutan y se puede aducir que la medida de uno afecta al otro (σ_x y σ_y podrían tomar distintos valores cada vez). Dos observables $\hat{A}_1$ y $\hat{A}_2$ sólo se pueden medir a la vez (esto es, sobre una misma copia del sistema) si son compatibles [3.11] En la teoría cuántica compatible equivale a que conmuten. El operador producto $\hat{A}_1 \hat{A}_2$ no es un observable (no es hermítico) a menos que $\hat{A}_1$ y $\hat{A}_2$ conmuten.

2) Típicamente la desigualdad se va a aplicar al caso $\mathscr{H} = \mathscr{H}_A \otimes \mathscr{H}_B$, siendo los observables $\hat{A}_i$ del primer espacio y $\hat{B}_j$ del segundo (lo cual asegura que conmutan) además los subsistemas A y B están espacialmente separados. En ese caso $\hat{S}$ como tal no se puede medir (excepto en el caso trivial de que los dos $\hat{A}_i$ sean compatibles y lo mismo los dos $\hat{B}_j$): dado que las interacciones son locales y no hay acción a distancia no se pueden hacer medidas conjuntas en sistemas espacialmente separados de operadores entrelazados, sólo de operadores separables, tal como $\hat{A}_i \otimes \hat{B}_j$, a saber, midiendo $\hat{A}_i$ en el sistema A y $\hat{B}_j$ en B. En cada copia del sistema sólo se puede medir uno de los A_i y uno de los $\hat{B}_j$ (por tanto tampoco se puede medir $\hat{S}$ midiendo sus componentes sueltas). Sin embargo la desigualdad de Bell (3.13) se refiere a medidas que sí son factibles, lo cual permite su verificación (o más bien su refutación) experimental.

[3.11] Esto es para un estado genérico. Si un estado de momento angular es $J = 0$, $J_x = J_y = 0$ aunque J_x y J_y no conmuten en el espacio completo.

3) En mecánica cuántica se pueden medir observables compatibles, por lo que existe una distribución de probabilidad para los resultados (A_i, B_j). Así, existe por ejemplo $P(a,b|A_1,B_1)$, pero no $P(a_1,a_2|A_1,A_2)$ ya que los valores (A_1,A_2) no pueden obtenerse en una misma medida sobre un mismo estado físico. La distribución $P(a,b|A_i,B_j)$ cuántica es simplemente [3.12]

$$P(a,b|A_i,B_j) = \left| {}_{\hat{A}_i}\langle a| \otimes {}_{\hat{B}_j}\langle b| \,|\, \psi\rangle \right|^2. \tag{3.28}$$

Por otro lado, en una teoría con variables ocultas locales todos los valores están definidos cada vez y existe la distribución de probabilidad conjunta $P(a_1,a_2,b_1,b_2)$. Por supuesto

$$P(a_1,a_2,b_1,b_2) \geq 0, \qquad \sum_{a_1=\pm1}\sum_{a_2=\pm1}\sum_{b_1=\pm1}\sum_{b_2=\pm1} P(a_1,a_2,b_1,b_2) = 1\,. \tag{3.29}$$

Las distribuciones $P(a,b|A_i,B_j)$ serían distribuciones marginales de ésta, así por ejemplo

$$P(a,b|A_1,B_1) = \sum_{a'=\pm1}\sum_{b'=\pm1} P(a,a',b,b')\,. \tag{3.30}$$

Lo que implica el teorema de Bell es que la mecánica cuántica proporciona distribuciones de probabilidad para conjuntos de observables compatibles entre sí, pero no existe una *superdistribución* de probabilidad para todos los observables (compatibles y no compatibles entre sí) de la cual todas las distribuciones cuánticas sean sus distribuciones marginales. [3.13] La mecánica cuántica sólo da probabilidades de experimentos factibles.

4) Todo el tiempo hemos dicho que las variables ocultas tratadas son *locales*. Si λ son las variables ocultas, y $\hat{\boldsymbol{a}}_i$, $\hat{\boldsymbol{b}}_j$ las direcciones de medida, la hipótesis es que

$$a_i = A(\hat{\boldsymbol{a}}_i,\lambda), \qquad b_j = B(\hat{\boldsymbol{b}}_j,\lambda) \tag{3.31}$$

para ciertas funciones A y B que toman valores en $\{\pm1\}$. El teorema de Bell equivale a decir que

$$\langle a_i b_j\rangle = \int A(\hat{\boldsymbol{a}}_i,\lambda) B(\hat{\boldsymbol{b}}_j,\lambda)\,\rho(\lambda) d\lambda \tag{3.32}$$

no puede reproducir $\langle \hat{A}_i \hat{B}_j\rangle$ cuántico (por ejemplo en estado singlete) para todas las elecciones de $\hat{\boldsymbol{a}}_i$, $\hat{\boldsymbol{b}}_j$, no importa cómo se elijan las funciones A y B y la distribución de probabilidad $\rho(\lambda)$. [3.14]

[3.12] Aquí se usa $|a\rangle_{\hat{A}_i}$ para indicar el vector propio normalizado de $\hat{A}_i$ con valor propio $a = \pm1$. Ídem B.

[3.13] La distribución de un observable tal como la posición $\hat{\boldsymbol{x}}$ se puede expresar como $\rho(\boldsymbol{x}) = \langle \delta(\boldsymbol{x}-\hat{\boldsymbol{x}})\rangle$ y lo mismo para el momento $\hat{\boldsymbol{p}}$, pero una distribución conjunta tal como $\rho(\boldsymbol{x},\boldsymbol{p}) = \langle \frac{1}{2}\{\delta(\boldsymbol{x}-\hat{\boldsymbol{x}}),\delta(\boldsymbol{p}-\hat{\boldsymbol{p}})\}\rangle$ no es definida positiva. En general si A y B son operadores positivos, el operador $\{A,B\}$ puede no serlo (AB en general no es hermítico) por ejemplo $A = \sigma_x + I$ y $B = \sigma_y + I$.

[3.14] A veces se dice que una forma más general de (3.32) es $\langle a_i b_j\rangle = \sum_{a,b=\pm1} ab \int \mathrm{Prob}(a|\hat{\boldsymbol{a}}_i,\lambda)\,\mathrm{Prob}(b|\hat{\boldsymbol{b}}_j,\lambda)\,\rho(\lambda) d\lambda$. En realidad no es así, ambas expresiones son igual de generales, se puede pasar de una a otra aumentando el número de parámetros λ en (3.32).

a_i depende sólo de las variables ocultas y de la dirección $\hat{\boldsymbol{a}}_i$ de la medida en A, y lo mismo para b_j. No se permite $A(\hat{\boldsymbol{a}}_i, \hat{\boldsymbol{b}}_j, \lambda)$ ni $B(\hat{\boldsymbol{b}}_j, \hat{\boldsymbol{a}}_i, \lambda)$, es decir, que el resultado de la medida en A dependa de lo que decide medir B o viceversa. Eso serían **variables ocultas no locales**. Para tales variables no se deriva la desigualdad de Bell, (y de hecho en la versión no local no es difícil elegir las funciones A y B tales que sí reproducen por ejemplo los promedios cuánticos en el estado de dos qubits en singlete). Técnicamente el motivo de que la desigualdad (3.13) no se derive es que en "$a_1 b_1$" y "$a_1 b_2$", en variables locales a_1 es la misma variable en ambos casos, a saber $A(\hat{\boldsymbol{a}}_1, \lambda)$, mientras que en variables ocultas no locales no: en "$a_1 b_1$" $a_1 = A(\hat{\boldsymbol{a}}_1, \hat{\boldsymbol{b}}_1, \lambda)$ mientras que en "$a_1 b_2$" $a_1 = A(\hat{\boldsymbol{a}}_1, \hat{\boldsymbol{b}}_2, \lambda)$, y la demostración no se aplica.

Esto quiere decir que los resultados cuánticos se pueden simular clásicamente si se permite no localidad pero no se pueden simular de manera local. Implica que si hay un número n creciente de qubits el cálculo cuántico crece como n pero la simulación clásica, que tiene que correlacionar todos los qubits entre sí, requiere una cantidad de variables clásicas que crece exponencialmente. *Un sistema cuántico grande general no puede ser simulado clásicamente de manera eficiente.* [3.15]

Aparte de variables ocultas no locales, otra forma de eludir el teorema de Bell es suponiendo que $\rho(\lambda)$ depende de la elección de $\hat{\boldsymbol{a}}$ y $\hat{\boldsymbol{b}}$. Es decir, se tiene una distribución de probabilidad condicionada $\rho(\lambda|\hat{\boldsymbol{a}}, \hat{\boldsymbol{b}})$, las elecciones de medidas de A y B afectan a λ, lo cual ciertamente no es el espíritu de las variables ocultas. Además, implicaría a su vez una probabilidad condicionada no trivial para $\text{Prob}(\hat{\boldsymbol{a}}, \hat{\boldsymbol{b}}|\lambda)$, es decir, las elecciones de A y B estarían influidas por las variables ocultas, lo cual contradice el principio de libre albedrío. Ese sería el caso por ejemplo en un mundo completamente determinista (solución que también se ha propuesto). [3.16]

Proseguimos la discusión en el Tema 3.2.5

3.2.3. Violación de las desigualdades de Bell y entrelazamiento

Los ejemplos que hemos visto de violación de desigualdades de Bell son estados puros entrelazados. Todo estado puro entrelazado viola alguna desigualdad de Bell (eligiendo las direcciones de medida $\hat{\boldsymbol{a}}_i$, $\hat{\boldsymbol{b}}_j$ adecuadamente) [3.17] En cambio los estados puros o mezcla separables no violan ninguna desigualdad de Bell.

[3.15] R.P. Feynman, *Simulating physics with computers,* Int. J. Theor. Phys. **21**(1982)467 [1]

[3.16] El libre albedrío significa la posibilidad de elegir. No existe en un mundo determinista y tampoco en uno aleatorio donde no se puede elegir el resultado. Por otro lado la posibilidad de libre albedrío no aparece por ningún lado en las ecuaciones de las teorías físicas.

[3.17] Una afirmación similar no se verifica para estados mezcla entrelazados. [Notas de Preskill, Tema 4.3.4].

En efecto, supongamos que se tiene un estado puro separable $|\psi\rangle$, entonces $\langle \hat{A}_i \hat{B}_j \rangle = \langle \hat{A}_i \rangle \langle \hat{B}_j \rangle$. Dado que $\hat{A}_i, \hat{B}_j$ toman valores ± 1,

$$x_i \equiv \langle \hat{A}_i \rangle, \qquad y_j \equiv \langle \hat{B}_j \rangle, \qquad \langle \hat{A}_i \hat{B}_j \rangle = x_i y_j, \qquad |x_i| \leq 1, \quad |y_i| \leq 1, \tag{3.33}$$

Podemos escribir

$$\langle \hat{S} \rangle = x_1(y_1 + y_2) + x_2(y_1 - y_2) \tag{3.34}$$

y también

$$|\langle \hat{S} \rangle| \leq |x_1||y_1 + y_2| + |x_2||y_1 - y_2| \leq |y_1 + y_2| + |y_1 - y_2| \tag{3.35}$$

Si $y_1 - y_2$ tiene el mismo signo que $y_1 + y_2$ se tiene

$$|\langle \hat{S} \rangle| \leq |(y_1 + y_2) + (y_1 - y_2)| = 2|y_1| \leq 2 \tag{3.36}$$

mientras que si tiene signo opuesto

$$|\langle \hat{S} \rangle| \leq |(y_1 + y_2) - (y_1 - y_2)| = 2|y_2| \leq 2 \tag{3.37}$$

En todo caso se satisface la desigualdad

$$-2 \leq \langle \hat{S} \rangle \leq 2 \qquad \text{(estado separable)} \tag{3.38}$$

Si se trata de un estado mezcla separable, la conclusión es la misma ya que $\langle \hat{S} \rangle$ será un promedio sobre estados puros que cumplen la desigualdad.

En realidad el resultado era obvio ya que para estados separables realmente sí existen variables ocultas locales que pueden reproducir los promedios cuánticos. Las variables aleatorias a_i y b_j se pueden simular Monte Carlo de manera independiente. Si $|\psi\rangle = |\phi\rangle_A |\chi\rangle_B$, basta hacer $a_i \sim |\langle a_i|\phi\rangle|^2$ y $b_j \sim |\langle b_j|\chi\rangle|^2$.

3.2.4. Desigualdad de Tsirelson

Veamos que la cota $|\langle \hat{S} \rangle| \leq 2\sqrt{2}$ no se puede superar para ningún estado. De nuevo suponemos sólo que $\hat{A}_i$ conmuta con $\hat{B}_j$ y que sus espectros son $\{\pm 1\}$ (no hace falta que sean qubits) y $\hat{S} = \hat{A}_1\hat{B}_1 + \hat{A}_1\hat{B}_2 + \hat{A}_2\hat{B}_1 - \hat{A}_2\hat{B}_2$. Por un lado, el cuadrado de un operador hermítico es positivo

$$\begin{aligned} 0 &\leq \left(\hat{A}_1 - \frac{1}{\sqrt{2}}(\hat{B}_1 + \hat{B}_2) \right)^2 \\ &= \hat{A}_1^2 + \frac{1}{2}\left(\hat{B}_1 + \hat{B}_2\right)^2 - \sqrt{2}\hat{A}_1\left(\hat{B}_1 + \hat{B}_2\right) \\ &= \hat{A}_1^2 + \frac{1}{2}\left(\hat{B}_1^2 + \hat{B}_2^2\right) + \frac{1}{2}\{\hat{B}_1, \hat{B}_2\} - \sqrt{2}\hat{A}_1\hat{B}_1 - \sqrt{2}\hat{A}_1\hat{B}_2 \end{aligned} \tag{3.39}$$

Igualmente con las sustituciones $\hat{A}_1 \leftrightarrow \hat{A}_2$ y $\hat{B}_2 \to -\hat{B}_2$,

$$0 \leq \hat{A}_2^2 + \frac{1}{2}\left(\hat{B}_1^2 + \hat{B}_2^2\right) - \frac{1}{2}\{\hat{B}_1, \hat{B}_2\} - \sqrt{2}\hat{A}_2\hat{B}_1 + \sqrt{2}\hat{A}_2\hat{B}_2 \tag{3.40}$$

Sumando ambas desigualdades se obtiene

$$0 \leq \hat{A}_1^2 + \hat{A}_2^2 + \hat{B}_1^2 + \hat{B}_2^2 - \sqrt{2}\hat{S} \tag{3.41}$$

Teniendo en cuenta que $\hat{A}_i^2 = \hat{B}_j^2 = 1$, se obtiene [3.18]

$$\hat{S} \leq 2\sqrt{2} \tag{3.42}$$

Igualmente las sustituciones $\hat{B}_j \to -\hat{B}_j$, producen $\hat{S} \to -\hat{S}$, por tanto $-\hat{S} \leq 2\sqrt{2}$, y en conjunto

$$-2\sqrt{2} \leq \hat{S} \leq 2\sqrt{2}\,. \tag{3.43}$$

La cota es óptima ya que hemos visto que se satura con el ejemplo en (3.27)

3.2.5. Apéndice: Paradoja EPR y mecánica cuántica

[Nota: estas consideraciones son prescindibles.]

El argumento EPR parece bastante sólido, sin embargo su implicación más obvia (la existencia de variables ocultas locales) no se verifica experimentalmente, en favor de la predicción cuántica. ¿Qué falla en el argumento EPR o quizá en su aparente implicación? En realidad no hay una respuesta que satisfaga a todo el mundo. Supongamos un experimento (clásico) en el que dos bolas, blanca (0) y negra (1) son sorteadas (sin mirar) al 50 % y se meten en sendas cajas cerradas A y B que son enviadas a puntos separados x_A y x_B del espacio-tiempo. Podríamos describir el estado del sistema como (todo es clásico pero usamos una notación sugerente de tipo cuántico)

$$\rho = \frac{1}{2}|0\rangle_A|1\rangle_B + \frac{1}{2}|1\rangle_A|0\rangle_B \tag{3.44}$$

Cuando A abre la caja el sistema colapsará a $|0\rangle_A|1\rangle_B$ o $|1\rangle_A|0\rangle_B$ con igual probabilidad. Si en A está la bola negra en B estará la blanca y viceversa, y no hay necesidad de atribuirlo a una influencia supralumínica entre A y B.

En el caso cuántico se tiene algo parecido con

$$|\Psi\rangle = \frac{1}{\sqrt{2}}|0\rangle_A|1\rangle_B - \frac{1}{\sqrt{2}}|1\rangle_A|0\rangle_B \tag{3.45}$$

[3.18] Dados dos operadores hermíticos, A y B, $A \geq B$ quiere decir $A - B \geq 0$, equivalentemente, $\forall|\psi\rangle\ \langle A\rangle_\psi \geq \langle B\rangle_\psi$.

Al medir y encontrar 0 en A se sabe que va a ser 1 en B y viceversa. Pero, también hay importantes diferencias:

i) La descripción clásica más general es mediante **probabilidades** (los casos de certeza están incluidos ahí). En mecánica cuántica hay **amplitudes de probabilidad** además de probabilidades. La mecánica cuántica es una extensión del concepto clásico de probabilidad. De hecho, una expresión como (3.44) aparece de manera natural en el formalismo cuántico: si x son los estados posibles $|x)$ representa $|x\rangle\langle x|$ siendo $|x\rangle$ una base ortonormal (privilegiada, digamos computacional) y ρ es un estado mezcla de estados $|x\rangle$ sin permitir superposición.

ii) Para una misma base $|x\rangle$, en mecánica cuántica hay muchos más estados y observables que en la versión clásica. Mientras que un estado clásico $p_0|0) + p_1|1)$ es meramente una mezcla de estados especiales (extremales) $|0)$ y $|1)$, en el caso cuántico la superposición $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ es un estado genuino, ya que puede obtenerse por ejemplo aplicando una rotación sobre $|0\rangle$ (pensando en el caso de partículas de espín $1/2$). [3.19] En un espacio de Hilbert (por tanto coeficientes arbitrarios) no hay estados "extremales", todos los estados de norma 1 son equivalentes (a falta de más condiciones sobre ese espacio [3.20]). Cualquier vector unitario se puede expresar como combinación de otros y cualquier vector unitario se puede completar para formar una base ortonormal. En consecuencia, clásicamente los observables son medir x (es decir, sólo la base computacional) en cuántica se puede medir cualquier superposición (estado puro).

iii) En clásica se pueden medir estados de la base computacional. Todos esos observables son siempre compatibles y se pueden medir a la vez. En cuántica se pueden medir muchos más observables pero no conjuntos arbitrarios, han de ser conjuntos de observables compatibles (deben conmutar como operadores). La mecánica cuántica proporciona predicciones (en forma de probabilidades) para cualquier conjunto de observables compatibles, pero no para conjuntos generales. Sólo proporciona predicciones (que además son experimentalmente exitosas) para medidas que realmente se pueden hacer experimentalmente. El argumento EPR está hábilmente diseñado para hacer una medida y a la vez casi tener el resultado de otra medida que realmente no se llega a hacer (todo sobre una misma copia del sistema físico preparado en cierto estado puro cuántico). La violación de las desigualdades de Bell indica que en cuántica no es equivalente saber lo que saldría al hacer la medida que realmente hacer la medida. Puede parecer paradójico pero no es lógicamente inconsistente. La mecánica cuántica aprovecha todos los recursos lógicos permitidos bordeando la contradicción pero sin caer en ella. La violación de las desigualdades de Bell no implica una inconsistencia en el sentido matemático, sólo un conflicto

[3.19] Como es fácil comprobar, el espacio $p_0|0) + p_1|1)$, o cualquier espacio de dimensión finita pero permitiendo sólo pesos no negativos, no puede llevar una representación de un grupo de Lie, y en particular no puede representar rotaciones.

[3.20] Por ejemplo, si el espacio es bipartito distinguimos entre estados separables y entrelazados.

con la idea que nos habíamos formado de que cómo funcionan las cosas. El conflicto desaparece cuando se admite una descripción más general que una basada sólo en probabilidades y se permiten también las amplitudes de probabilidad.

iv) La descripción clásica sólo usa estados de la base computacional, si se pone la misma restricción en un tratamiento cuántico (mezclar sólo estados de la base computacional y no otros estados puros superposición de éstos) se obtiene exactamente lo mismo y sólo habrá aleatoriedad debida a la mezcla. Tal aleatoriedad representa una ignorancia (sobre el resultado de la medida cuando se haga) pero es una ignorancia que se puede reducir con más información (alguien puede saber cuál es realmente el estado concreto, $|0)_A|1)_B$ o $|1)_A|0)_B$, aunque nosotros no lo sepamos). En cambio la aleatoriedad en el resultado de una medida de un observable cuántico cualquiera en un estado puro (conocido) cualquiera representa una ignorancia intrínseca (no subsanable) ya que no se puede reducir con más información (intentar obtener más información haciendo una medida modificaría el estado). [3.21] En mecánica cuántica dos sistemas físicos preparados en el mismo estado puro son idénticos. La versión cuántica permite medir muchos más observables que la clásica pero sólo garantiza resultados estadísticos para esos nuevos observables.

En resumen, la mecánica cuántica implica una estructura matemática que es una extensión de la teoría de probabilidades clásicas, al permitir asignar amplitudes de probabilidad a distintos estados x. Tal extensión es perfectamente consistente en sentido matemático. Una vez que una construcción se revela lógicamente consistente, existe la posibilidad lógica de que la naturaleza haga uso de ella. [3.22] Además de enormemente exitosa, la teoría cuántica es extremadamente elegante: de manera compacta y completamente transparente un estado como en (3.45) produce predicciones, con unas reglas muy simples, sobre cualquier conjunto compatible de observables del sistema. Eso es a comparar con teorías de variables ocultas (necesariamente no locales o en contradicción con la hipótesis del libre albedrío) que requieren sofisticados métodos completamente ad hoc para reproducir los mismos resultados cuánticos. [3.23] La simplicidad de la mecánica cuántica en la naturaleza podría ser un fenómeno emergente (esto es, aparentemente simple a escala microscópica o mayor, pero complicado a escalas mucho menores) pero parece más verosímil la posibilidad de que la naturaleza hace uso de un recurso matemático disponible básico y lógicamente consistente. A nosotros nos cuesta asimilarlo porque estamos hechos a un mundo clásico dado que la riqueza de posibilidades de la coherencia cuántica tiende a disiparse rápidamente cuando hay muchos grados de libertad involucrados, es decir,

[3.21] Por poner una imagen, la función de onda se asemeja a un horizonte de sucesos (de un agujero negro clásico) a través del cual no hay observación posible, en lugar de censura cósmica sería censura cuántica.

[3.22] Ejemplos son el grupo de Lorentz, o la existencia de partículas de espín semientero. Presumiblemente también dimensiones extra o supersimetría.

[3.23] Cualquier distribución de variables aleatorias se puede reproducir siempre mediante un método Monte Carlo, que es exactamente lo mismo que usar variables ocultas, lo que no es trivial es que esas variables sean locales.

en el mundo macroscópico. [3.24]

Otra observación es que sólo los **sistemas cuánticos naturales** pueden realizar auténtica computación o información cuántica. Si consideramos un promedio de una magnitud A pesado con pesos no negativos, $\langle A\rangle = \sum_x w_x A_x$, es factible hacer una simulación Monte Carlo de ese promedio haciendo un muestreo de x con distribución w_x (es decir, generar x con frecuencia relativa w_x). Si los pesos pueden ser negativos o complejos la suma o integral sigue estando bien definida pero ya no se puede hacer tal muestreo y hay que recurrir a métodos más complicados (y crecientemente ineficientes en general al aumentar el tamaño del sistema). Del mismo modo, sólo los sistemas físicos saben calcular cuánticamente, por ejemplo producir estados entrelazados. Aunque el esquema matemático de la mecánica cuántica (que incluya un postulado para medidas ideales) se entiende teóricamente, no se sabe construir un sistema artificial que lo implemente. Hay que hacerlo siempre sobre la base de grados de libertad físicos presentes en la naturaleza ya que sólo ellos saben hacer el truco cuántico.

[3.24] Debe notarse también que la propia teoría de probabilidades (clásicas, no amplitudes de probabilidad) está lejos de tener una interpretación universalmente aceptada.

3.3. Algunas aplicaciones del entrelazamiento

3.3.1. Codificación densa

Tenemos dos estaciones (posiciones localizadas espacialmente) A y B que pueden intercambiar qubits físicos entre sí, por ejemplo mediante fotones. Típicamente las dos estaciones son etiquetadas como Alice y Bob, o Andrea y Benito.

Podría pensarse que Andrea puede enviar mucha información a Benito usando un qubit. El qubit tendrá dos parámetros (θ, ϕ) (ángulos esféricos en la esfera de Bloch) y en esos números reales se puede codificar un mensaje de 0 y 1 (bits) arbitrariamente grande. En realidad no es así, si Andrea envía un qubit $|\psi\rangle$ a Benito sin ningún tipo de acuerdo previo, todo lo que puede hacer Benito es hacer una medida en cierta base a elegir por él, digamos la computacional, y obtendrá como resultado $|0\rangle$ o $|1\rangle$ y el estado $|\psi\rangle$ que tuviera el qubit enviado por Andrea se pierde. Si por ejemplo sale $|1\rangle$ todo lo que se puede deducir es que la amplitud de probabilidad del estado $|1\rangle$ en $|\psi\rangle$ no era 0.

Si previamente se ha acordado entre las dos partes que Andrea va a enviar exactamente uno de los dos estados $|0\rangle$ o $|1\rangle$, entonces sí puede transmitir un bit de información. Benito mide el qubit recibido y lee 0 o 1.

El protocolo de **codificación densa** (también llamada superdensa) permite enviar *dos bits* de información enviando un solo qubit. El protocolo es como sigue:

1) Andrea tiene un qubit A y Benito tiene otro B. Los dos qubits están entrelazados formando un estado de Bell. Digamos $|\Phi_-\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle_{AB} - |10\rangle_{AB})$. [3.25]

2) Andrea aplica sobre su qubit uno de los cuatro **operadores de Pauli**, a saber: I, $X = \sigma_x$,

[3.25] Se puede usar cualquiera de los estados de Bell, con las modificaciones correspondientes.

$Y = \sigma_y, Z = \sigma_z$,

$$\begin{aligned} I_A|\Phi_-\rangle_{AB} &= \frac{1}{\sqrt{2}}(|01\rangle_{AB} - |10\rangle_{AB}) = |\Phi_-\rangle_{AB} \\ X_A|\Phi_-\rangle_{AB} &= \frac{1}{\sqrt{2}}(|11\rangle_{AB} - |00\rangle_{AB}) = -|\Psi_-\rangle_{AB} \\ Y_A|\Phi_-\rangle_{AB} &= \frac{1}{\sqrt{2}}(i|11\rangle_{AB} + i|00\rangle_{AB}) = i|\Psi_+\rangle_{AB} \\ Z_A|\Phi_-\rangle_{AB} &= \frac{1}{\sqrt{2}}(|01\rangle_{AB} + |10\rangle_{AB}) = |\Phi_+\rangle_{AB} \end{aligned} \tag{3.46}$$

Cada operación produce un estado de Bell distinto.

3) Andrea envía físicamente el qubit A a Benito.

4) Benito mide el estado de Bell. (Se puede hacer usando el circuito inverso al de la Fig. 3.1.) Puesto que esos estados son ortogonales y Benito mide en esa misma base, el resultado es completamente determinista y sin ambigüedad, y de hecho la medida no cambia el estado. De ese modo Benito determina cuál de los cuatro operadores de Pauli se ha aplicado. Por tanto Andrea ha enviado *dos bits de información* con un solo qubit físico.

Aquí se ve que un *par de qubits máximamente entrelazados pero separados* (a veces denominado **ebit**) proporciona un recurso que puede ser utilizado con fines prácticos. En la codificación densa se tiene la relación 1ebit = 2cbits. Después usar el recurso (enviar los dos bits) los dos qubits siguen entrelazados pero no están separados; no se pueden usar tal cual para comunicación cuántica. Benito puede aplicar la operación de Pauli adecuada para restaurar el estado de Bell $|\Phi_-\rangle$, devolver el qubit A a Andrea y el ebit se puede reutilizar.

Es importante notar que la *formación de estados entrelazados requiere interacción*. La interacción siempre es local, por tanto no se puede generar entrelazamiento a distancia (no hay interacción a distancia). [3.26] Los dos qubits A y B deben estar juntos para ser entrelazados, por ejemplo Andrea los entrelaza en su estación y luego envía el qubit B a Benito. *Los dos qubits siguen entrelazados aunque estén separados.* Además permanecerán entrelazados indefinidamente si no se tocan y no se deterioran (pérdida de coherencia por interacción con el entorno). No hace falta que la formación del ebit se haga cada vez que se quiere enviar información. De hecho Andrea y Benito pueden tener un acopio de ebits para ir usando conforme hagan falta. También hay que notar que para medir el estado

[3.26] En cambio sí se puede desentrelazar a distancia: si AB comparten un estado de Bell, basta que A mida en su base computacional para que los dos qubits queden en un estado separable.

de Bell de nuevo hace falta interacción y por tanto la medida sólo se puede hacer si los dos qubit físicos están juntos. No se puede medir un estado de Bell con qubits separados.

3.3.2. Teleportación

En este protocolo Andrea tiene un qubit en un estado $|\psi\rangle$ y quiere enviar ese estado a Benito, pero no físicamente. Sólo quiere transferir el estado. Si se sabe de antemano que el $|\psi\rangle$ es o bien $|0\rangle$ o bien $|1\rangle$ (o cualquier otra base ortonormal conocida) se puede medir y comunicar el bit clásico. Pero suponemos que $|\psi\rangle$ ha sido recibido por Andrea e ignora qué estado es. Es posible **teleportar** $|\psi\rangle$ de Andrea a Benito, es decir, que pase de estar en A a estar en B, usando entrelazamiento. El protocolo es como sigue:

1) Sea A_1 el qubit que contiene al estado $|\psi\rangle$ en posesión de Andrea. Suponemos que Andrea y Benito comparten un par de qubits A_2 y B entrelazados en estado de Bell $|\Phi_-\rangle_{A_2B}$. Por tanto el estado completo es

$$\begin{aligned} |\chi\rangle &= |\psi\rangle_{A_1} \otimes |\Phi_-\rangle_{A_2B} = (\alpha|0\rangle + \beta|1\rangle)_{A_1} \otimes \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)_{A_2B} \\ &= \frac{1}{\sqrt{2}}(\alpha|00\rangle \otimes |1\rangle - \alpha|01\rangle \otimes |0\rangle + \beta|10\rangle \otimes |1\rangle - \beta|11\rangle \otimes |0\rangle) \end{aligned} \tag{3.47}$$

2) A continuación Andrea mide los qubits A_1A_2 en la base de Bell. Las amplitudes son

$$\begin{aligned} \langle\Psi_+|\chi\rangle &= \frac{1}{2}(\alpha|1\rangle - \beta|0\rangle) = \frac{1}{2}iY|\psi\rangle_B \\ \langle\Psi_-|\chi\rangle &= \frac{1}{2}(\alpha|1\rangle + \beta|0\rangle) = \frac{1}{2}X|\psi\rangle_B \\ \langle\Phi_+|\chi\rangle &= \frac{1}{2}(-\alpha|0\rangle + \beta|1\rangle) = -\frac{1}{2}Z|\psi\rangle_B \\ \langle\Phi_-|\chi\rangle &= \frac{1}{2}(-\alpha|0\rangle - \beta|1\rangle) = -\frac{1}{2}|\psi\rangle_B \end{aligned} \tag{3.48}$$

   Tras la medida, Andrea obtiene exactamente uno de los estados de Bell, cada uno con probabilidad $1/4$.

3) Andrea comunica a Benito cuál ha sido el resultado de la medida. Lo hace enviando dos bits de información por el canal clásico.

4) Con esa información Benito sabe qué operador de Pauli tiene que aplicar para reconstruir el estado $|\psi\rangle$ en el qubit B.

Benito se queda con $|\psi\rangle$, el entrelazamiento entre A_2 y B se pierde, y los qubits A_1A_2 se quedan en un estado de Bell conocido pero Andrea no conserva $|\psi\rangle$ (y en ningún momento necesita saber qué estado es). Ha sido necesaria una comunicación clásica para enviar dos bits. El balance sería $1\,\text{ebit} + 2\,\text{cbits} = 1\,\text{qubit}$.

3.4. Condiciones de separabilidad

Como se ha ilustrado con codificación densa y teleportación, el entrelazamiento es un recurso. Determinar el estado de separabilidad o entrelazamiento para estados puros es simple, como ya hemos visto: se puede usar la forma de Schmidt o la matriz densidad reducida. Sin embargo no hay un criterio similar para decir si un estado mezcla es separable o entrelazado. Hay algunas condiciones suficientes que garantizan la existencia de entrelazamiento.

3.4.1. Criterio basado en desigualdades de Bell

Un criterio para determinar que un estado mezcla es entrelazado y no separable se basa en las desigualdades de Bell. Si el estado viola alguna desigualdad de Bell, eligiendo los observables A_1, A_2, B_1, B_2 apropiados, es seguro que estado es entrelazado. Desafortunadamente este criterio no es siempre concluyente porque (a diferencia de lo que ocurre para estados puros) hay no pocos estados mezcla entrelazados que no violan ninguna desigualdad de Bell.

3.4.2. Criterio de traspuesto parcial positivo

Un criterio más fuerte es el **traspuesto parcial positivo**. [3.27] Si tenemos una matriz densidad ρ, la matriz traspuesta ρ^T también define un operador matriz densidad ya que es positivo y con traza 1.

3.4.2.1. Apartado matemático: Operadores traspuesto y conjugado

Si $\{|n\rangle\}$ es una cierta base ortonormal, un operador $\hat{A}$ queda caracterizado por sus elementos de matriz

[3.27] A. Peres, *Separability criterion for density matrices,* Phys. Rev. Lett. **77** (1996), 1413-1415. [2]

$\langle n|\hat{A}|m\rangle$. Entonces se pueden definir los nuevos operadores $\hat{A}^T$, $\hat{A}^*$ y $\hat{A}^\dagger$ por sus elementos de matriz:

$$\langle n|\hat{A}^T|m\rangle = \langle m|\hat{A}|n\rangle, \qquad \langle n|\hat{A}^*|m\rangle = \langle n|\hat{A}|m\rangle^*, \qquad \langle n|\hat{A}^\dagger|m\rangle = \langle m|\hat{A}|n\rangle^*. \tag{3.49}$$

Sin embargo los operadores $\hat{A}^T$ y $\hat{A}^*$ dependen de la base ortonormal usada para construirlos: Si A es la matriz de $\hat{A}$ en una base, en otra base ortonormal la matriz será $UAU^\dagger$. Si se traspone en la nueva base y se lleva el resultado a la base original se tendrá

$$U^\dagger(UAU^\dagger)^T U = U^\dagger U^* A^T U^T U = VA^T V^\dagger, \qquad V \equiv U^\dagger U^*. \tag{3.50}$$

Transponer en dos bases distintas produces operadores distintos, aunque relacionados unitariamente. Igual pasa con $\hat{A}^*$,

$$U^\dagger(UAU^\dagger)^* U = U^\dagger U^* A^* U^T U = VA^* V^\dagger. \tag{3.51}$$

En cambio $\hat{A}^\dagger$ no depende de la base (ortonormal).

En todo caso $\hat{A}^T$ tiene el mismo espectro que $\hat{A}$. El espectro son las raíces de la ecuación secular $\det(A - \lambda I) = 0$ y son las mismas para A y A^T por la propiedad del determinante $\det(A) = \det(A^T)$. En particular $\mathrm{Tr}(\hat{A}) = \mathrm{Tr}(\hat{A}^T)$. Igualmente se concluye que el espectro de $\hat{A}^*$ es el conjugado del espectro de $\hat{A}$.

Cuando $\hat{A}$ es hermítico $\hat{A}^T = \hat{A}^*$ y también es hermítico. Se concluye que si $\hat{A}$ es un operador positivo $\hat{A}^T$ también, y si ρ es una matriz densidad ρ^T también.

Si tenemos un sistema bipartito $\mathscr{H}_A \otimes \mathscr{H}_B$ con matriz densidad ρ_{AB}, se define su **traspuesto parcial** respecto de B, $\rho_{AB}^{T_B}$, mediante

$$\langle n,m|\rho_{AB}^{T_B}|n',m'\rangle \equiv \langle n,m'|\rho_{AB}|n',m\rangle. \tag{3.52}$$

Esta definición depende de la base $\{|m\rangle_B\}$ que se utilice, aunque las distintas definiciones son todas de la forma $I_A \otimes V_B \rho^{T_B} I_A \otimes V_B^\dagger$, donde V_B es un operador unitario en $\mathscr{H}_B$ que sólo depende de la base en $\mathscr{H}_B$ (es independiente de ρ_{AB}). En particular esto implica que todas las versiones tienen el mismo espectro (y por tanto la misma traza). Además, como $\rho_{AB}^{T_A} = (\rho_{AB}^{T_B})^T$ y trasponer no cambia el espectro, se deduce que $\rho_{AB}^{T_A}$ y $\rho_{AB}^{T_B}$ tienen el mismo espectro.

$\rho_{AB}^{T_B}$ es un operador hermítico pero en general ρ_{AB} y $\rho_{AB}^{T_B}$ no tienen el mismo espectro.

Definición Se dice que un operador matriz densidad ρ_{AB} tiene **traspuesto parcial positivo** (TPP) cuando $\rho_{AB}^{T_B} \geq 0$.

Un estado separable siempre es TPP. En efecto, supongamos que ρ_{AB} es separable

$$\rho_{AB} = \sum_j p_j \rho_{Aj} \otimes \rho_{Bj}, \qquad \rho_{AB}^{T_B} = \sum_j p_j \rho_{Aj} \otimes \rho_{Bj}^T, \tag{3.53}$$

como ρ_{Bj}^T es una matriz densidad se deduce que $\rho_{AB}^{T_B}$ también es una matriz densidad (y separable) por tanto es un operador positivo.

Se tiene entonces un criterio *suficiente* para saber que el estado ρ_{AB} está entrelazado: si no es TPP es necesariamente entrelazado. Por otro lado el criterio no es también necesario ya que hay estados entrelazados pero con traspuesto parcial positivo.

Lo dicho es cierto para espacios genéricos. Para los casos particulares de $\mathbb{C}^2 \otimes \mathbb{C}^2$ (dos qubits) y $\mathbb{C}^2 \otimes \mathbb{C}^3$ (un qubit y un qutrit) la propiedad TPP es necesaria y suficiente para que el estado sea separable.

Por ejemplo, consideremos para dos qubits el estado

$$\rho_{AB} = p|\Phi_-\rangle\langle\Phi_-| + (1-p)|00\rangle\langle 00|, \qquad 0 \leq p \leq 1. \tag{3.54}$$

Puede probarse que este estado no viola ninguna desigualdad de Bell cuando $p \leq \frac{1}{\sqrt{2}} \approx 0.71$. Pero en realidad es entrelazado si $p > 0$, como se comprueba aplicando el criterio TPP. En la base computacional su matriz es

$$\rho_{AB} = \begin{array}{c} \begin{pmatrix} 1-p & 0 & 0 & 0 \\ 0 & p/2 & -p/2 & 0 \\ 0 & -p/2 & p/2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} 00 \\ 01 \\ 10 \\ 11 \end{matrix} \\ \begin{matrix} 00 & 01 & 10 & 11. \end{matrix} \end{array} \tag{3.55}$$

Entonces, para su traspuesto parcial [3.28]

$$\rho_{AB}^{T_B} = \begin{pmatrix} 1-p & 0 & 0 & -p/2 \\ 0 & p/2 & 0 & 0 \\ 0 & 0 & p/2 & 0 \\ -p/2 & 0 & 0 & 0 \end{pmatrix}. \tag{3.56}$$

[3.28] Si $\rho = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, $\rho^{T_A} = \begin{pmatrix} A & C \\ B & D \end{pmatrix}$, $\rho^{T_B} = \begin{pmatrix} A^T & B^T \\ C^T & D^T \end{pmatrix}$.

Es una matriz hermítica y con traza 1. Con respecto a sus autovalores, la ecuación secular de $\rho_{AB}^{T_B}$ es

$$(\frac{p}{2}-\lambda)^2\left(\lambda^2-(1-p)\lambda-\frac{p^2}{4}\right)=0. \tag{3.57}$$

Tiene tres raíces no negativas pero la cuarta es $\lambda_4=\frac{1}{2}\left((1-p)-\sqrt{(1-p)^2+p^2}\right)<0$ si $p>0$. Por tanto el estado no tiene la propiedad TPP y es entrelazado excepto cuando $p=0$.

3.4.3. Testigo de entrelazamiento

Un operador hermítico W es un **testigo de entrelazamiento** para un estado entrelazado ρ_e cuando satisface dos propiedades:
1) $\mathrm{Tr}(\rho_s W)\geq 0$ para cualquier estado separable ρ_s.
2) $\mathrm{Tr}(\rho_e W)<0$.

W es un observable y por tanto puede medirse. Proporciona un método empírico para concluir que un estado es entrelazado.

Se puede construir un testigo de entrelazamiento para estados cuyo traspuesto parcial es negativo. Supongamos que ρ^{T_B} tiene un autovalor negativo λ_- con autovector $|\eta\rangle$. Usando la propiedad

$$\mathrm{Tr}(XY^{T_B})=\mathrm{Tr}_A(\mathrm{Tr}_B(XY^{T_B}))=\mathrm{Tr}_A(\mathrm{Tr}_B(X^{T_B}Y))=\mathrm{Tr}(X^{T_B}Y) \tag{3.58}$$

se tiene

$$\mathrm{Tr}(\rho(|\eta\rangle\langle\eta|)^{T_B})=\mathrm{Tr}(\rho^{T_B}|\eta\rangle\langle\eta|)=\lambda_-<0. \tag{3.59}$$

Mientras que para un estado ρ_s es separable

$$\mathrm{Tr}(\rho_s(|\eta\rangle\langle\eta|)^{T_B})=\mathrm{Tr}(\rho_s^{T_B}|\eta\rangle\langle\eta|)=\langle\eta|\rho_s^{T_B}|\eta\rangle\geq 0 \tag{3.60}$$

ya que $\rho_s^{T_B}\geq 0$. Por tanto $W=(|\eta\rangle\langle\eta|)^{T_B}$ es un testigo de entrelazamiento para el estado ρ.

3.4.4. Criterio DGCZ de separabilidad

Hay muchos criterios relativos a entrelazamiento de estados mezcla, a menudo basados en TPP. En el caso de variables continuas TPP es difícil de comprobar y vemos aquí una versión más simple.

El criterio DGCZ (Duan, Giedke, Cirac, Zoller) es para espacios de Hilbert de dimensión infinita, concretamente dos partículas moviéndose en $\mathbb{R}$. Los grados de libertad son x_1, p_1 (posición y momento

de la partícula 1) y x_2, p_2 (para la partícula 2). El mismo sistema se puede aplicar a fotones (o en general partículas cuánticas bosónicas) que puedan estar en dos orbitales 1 y 2. En efecto, como se ve al estudiar el oscilador armónico en el formalismo de operadores de creación y destrucción, se puede pasar de x_j, p_j a $a_j, a_j^\dagger$ mediante (unidades $m = \omega = \hbar = 1$)

$$x_j = \frac{a_j + a_j^\dagger}{\sqrt{2}}, \qquad p_j = \frac{a_j - a_j^\dagger}{i\sqrt{2}}, \qquad j = 1,2. \tag{3.61}$$

Las relaciones de conmutación en las dos versiones son

$$[x_j, x_k] = [p_j, p_k] = 0, \qquad [x_j, p_k] = i\delta_{jk}, \qquad [a_j, a_k] = [a_j^\dagger, a_k^\dagger] = 0, \qquad [a_j, a_k^\dagger] = \delta_{jk} \tag{3.62}$$

Aquí usaremos la versión x, p.

Definimos los observables

$$u = \alpha_1 x_1 + \alpha_2 x_2, \qquad v = \beta_1 p_1 + \beta_2 p_2, \qquad \alpha_j, \beta_j \in \mathbb{R} \tag{3.63}$$

Teorema Si ρ es separable entonces

$$(\Delta u)^2 + (\Delta v)^2 \geq |\alpha_1 \beta_1| + |\alpha_2 \beta_2| \tag{3.64}$$

(condición necesaria de separabilidad). Aquí $(\Delta A)^2$ es la varianza de A, es decir $(\Delta A)^2 = \langle A^2 \rangle - \langle A \rangle^2$.

Antes de ver la demostración veamos otra afirmación:

Proposición Para un estado ρ cualquiera (separable o entrelazado)

$$(\Delta u)^2 + (\Delta v)^2 \geq |\alpha_1 \beta_1 + \alpha_2 \beta_2| \tag{3.65}$$

Esta cota inferior es menos restrictiva que la anterior, ya que es menor cuando $\alpha_1\beta_1$ y $\alpha_2\beta_2$ tienen signos opuestos.

Para demostrar proposición y teorema, necesitamos primero

Lema 1 Para X operador cualquiera y ρ estado cualquiera [3.29]

$$|\langle X \rangle|^2 \leq \langle X^\dagger X \rangle \tag{3.66}$$

[3.29] Cualesquiera pero siempre se supone suficientemente regulares. Aquí X cualquiera quiere decir no necesariamente hermítico.

(Igualmente $|\langle X\rangle|^2 \leq \langle XX^\dagger\rangle$.)

Demostración: Para $\rho = |\psi\rangle\langle\psi|$ (estado puro)

$$\langle X^\dagger X\rangle = \langle X\psi|X\psi\rangle \geq \langle X\psi|\rho|X\psi\rangle = \langle X\psi|\psi\rangle\langle\psi|X\psi\rangle = |\langle X\rangle|^2 \tag{3.67}$$

(se ha usado que $\rho \leq I$).

Ahora, en el caso general, $\rho = \sum_k p_k\rho_k$, con $p_k \geq 0$, $\sum_k p_k = 1$ y ρ_k son estados puros.

Para aligerar las fórmulas voy a usar la siguiente **notación**

$$\overline{\xi_k} \equiv \sum_k p_k\xi_k \qquad (\xi_k \text{ es una magnitud arbitraria}) \tag{3.68}$$

Así por ejemplo $\rho = \overline{\rho_k}$. Con esta notación

$$\langle X\rangle = \overline{\langle X\rangle_k}, \qquad \langle X^\dagger X\rangle = \overline{\langle X^\dagger X\rangle_k}, \tag{3.69}$$

Aplicamos la desigualdad de Cauchy-Schwarz: [3.30]

$$x_k, y_k \in \mathbb{C} \qquad |\overline{x_k y_k}|^2 \leq \overline{|x_k|^2}\,\overline{|y_k|^2} \tag{3.70}$$

En particular, $|\overline{x_k}|^2 \leq \overline{|x_k|^2}$ (tomando $y_k \equiv 1$) entonces

$$|\langle X\rangle|^2 = |\overline{\langle X\rangle_k}|^2 \leq \overline{|\langle X\rangle_k|^2} \leq \overline{\langle X^\dagger X\rangle_k} = \langle X^\dagger X\rangle \tag{3.71}$$

□

También necesitamos el siguiente resultado

Lema 2 Para A, B operadores hermíticos y ρ arbitrario [3.31]

$$|\langle i[A,B]\rangle| \leq (\Delta A)^2 + (\Delta B)^2 \tag{3.72}$$

Demostración:

$$\begin{aligned} |\langle A+iB\rangle|^2 \leq \langle (A-iB)(A+iB)\rangle &\implies \langle A\rangle^2 + \langle B\rangle^2 \leq \langle A^2\rangle + \langle B^2\rangle + \langle i[A,B]\rangle \\ &\implies -\langle i[A,B]\rangle \leq (\Delta A)^2 + (\Delta B)^2 \end{aligned} \tag{3.73}$$

[3.30] Equivale a decir que para dos vectores $|\langle\psi|\phi\rangle|^2 \leq \|\psi\|^2\,\|\phi\|^2$.

[3.31] Si se reescala $A \to \lambda A$, $B \to \lambda^{-1}B$, la mínima cota superior se obtiene con $\lambda^2 = \Delta B/\Delta A$ e implica $|\langle i[A,B]\rangle| \leq 2\Delta A\Delta B \leq (\Delta A)^2 + (\Delta B)^2$.

Intercambiando A y B se deduce igualmente $\langle i[A,B]\rangle \leq (\Delta A)^2 + (\Delta B)^2$, lo que demuestra el Lema. □

Si se aplica el Lema 2 a u y v, teniendo en cuenta que

$$[u,v] = i\alpha_1\beta_1 + i\alpha_2\beta_2 \tag{3.74}$$

se deduce, para ρ cualquiera

$$(\Delta u)^2 + (\Delta v)^2 \geq |\alpha_1\beta_1 + \alpha_2\beta_2| \tag{3.75}$$

que demuestra la **proposición**.

Para demostrar el **teorema** (ρ no entrelazado) procedemos usando la identidad

$$(\Delta(A+B))^2 = (\Delta A)^2 + (\Delta B)^2 + 2\mathrm{Cov}(A,B) \tag{3.76}$$

siendo la covarianza $\mathrm{Cov}(A,B) = \langle \frac{1}{2}\{A,B\}\rangle - \langle A\rangle\langle B\rangle$, y si A y B conmutan $\mathrm{Cov}(A,B) = \langle AB\rangle - \langle A\rangle\langle B\rangle$.

La aplicamos a u y v en los estados puros separables ρ_k

$$\begin{aligned}(\Delta u)_k^2 &= \alpha_1^2(\Delta x_1)_k^2 + \alpha_2^2(\Delta x_2)_k^2 + 2\alpha_1\alpha_2\mathrm{Cov}(x_1,x_2)_k \\ (\Delta v)_k^2 &= \beta_1^2(\Delta p_1)_k^2 + \beta_2^2(\Delta p_2)_k^2 + 2\beta_1\beta_2\mathrm{Cov}(p_1,p_2)_k\end{aligned} \tag{3.77}$$

Pero por hipótesis los ρ_k son separables, eso implica

$$\langle x_1x_2\rangle_k = \langle x_1\rangle_k\langle x_2\rangle_k \implies \mathrm{Cov}(x_1,x_2)_k = 0 \qquad \text{ídem} \quad \mathrm{Cov}(p_1,p_2)_k = 0 \tag{3.78}$$

Por tanto

$$\begin{aligned}(\Delta u)_k^2 &= \alpha_1^2(\Delta x_1)_k^2 + \alpha_2^2(\Delta x_2)_k^2 \\ (\Delta v)_k^2 &= \beta_1^2(\Delta p_1)_k^2 + \beta_2^2(\Delta p_2)_k^2\end{aligned} \tag{3.79}$$

Si se usa ahora el Lema 2

$$\begin{aligned}[\alpha_1x_1, \beta_1p_1] = i\alpha_1\beta_1 \implies \quad & \alpha_1^2(\Delta x_1)_k^2 + \beta_1^2(\Delta p_1)_k^2 \geq |\alpha_1\beta_1| \\ \text{ídem} \quad & \alpha_2^2(\Delta x_2)_k^2 + \beta_2^2(\Delta p_2)_k^2 \geq |\alpha_2\beta_2|\end{aligned} \tag{3.80}$$

se deduce (sumando las igualdades de x y p)

$$(\Delta u)_k^2 + (\Delta v)_k^2 \geq |\alpha_1\beta_1| + |\alpha_2\beta_2| \tag{3.81}$$

y en consecuencia

$$\overline{(\Delta u)_k^2 + (\Delta v)_k^2} \geq |\alpha_1\beta_1| + |\alpha_2\beta_2| \tag{3.82}$$

Para completar la demostración del teorema necesitamos otro resultado:

Lema 3 Para A hermítico y ρ arbitrario

$$(\Delta A)^2 = \overline{(\Delta A)_k^2} + \mathrm{Var}(\langle A\rangle_k)\,. \tag{3.83}$$

Por tanto $(\Delta A)^2 \geq \overline{(\Delta A)_k^2}$.

Demostración: En efecto,

$$\begin{aligned}(\Delta A)^2 &= \langle A^2\rangle - \langle A\rangle^2 = \overline{\langle A^2\rangle_k} - \left(\overline{\langle A\rangle_k}\right)^2 \\ &= \overline{(\Delta A)_k^2 + \langle A\rangle_k^2} - \left(\overline{\langle A\rangle_k}\right)^2 = \overline{(\Delta A)_k^2} + \mathrm{Var}(\langle A\rangle_k) \qquad \mathrm{Var}(\langle A\rangle_k) \geq 0\end{aligned} \tag{3.84}$$

□

Aplicando el Lema a u y v, finalmente se concluye que para ρ separable $(\Delta u)^2 + (\Delta v)^2 \geq |\alpha_1\beta_1| + |\alpha_2\beta_2|$.

Obsérvese que partir directamente de

$$\begin{aligned}(\Delta u)^2 &= \alpha_1^2(\Delta x_1)^2 + \alpha_2^2(\Delta x_2)^2 + 2\alpha_1\alpha_2\,\mathrm{Cov}(x_1,x_2) \\ (\Delta v)^2 &= \beta_1^2(\Delta p_1)^2 + \beta_2^2(\Delta p_2)^2 + 2\beta_1\beta_2\,\mathrm{Cov}(p_1,p_2)\end{aligned} \tag{3.85}$$

sumando las dos expresiones para aplicar

$$\alpha_1^2(\Delta x_1)^2 + \beta_1^2(\Delta p_1)^2 \geq |\alpha_1\beta_1|, \qquad \alpha_2^2(\Delta x_2)^2 + \beta_2^2(\Delta p_2)^2 \geq |\alpha_2\beta_2| \tag{3.86}$$

no serviría porque aunque ρ no sea entrelazado eso no implica $\mathrm{Cov}(x_1,x_2) = 0$, $\mathrm{Cov}(p_1,p_2) = 0$. Un ρ no entrelazado todavía puede tener correlaciones clásicas (estadísticas):

$$\mathrm{Cov}(x_1,x_2) = \overline{\langle x_1x_2\rangle_k} - \overline{\langle x_1\rangle_k}\,\overline{\langle x_2\rangle_k} \underset{\rho_k \text{ sep.}}{=} \overline{\langle x_1\rangle_k\langle x_2\rangle_k} - \overline{\langle x_1\rangle_k}\,\overline{\langle x_2\rangle_k} = \mathrm{Cov}(\langle x_1\rangle_k, \langle x_2\rangle_k) \tag{3.87}$$

que no se anula en general.

Para un ρ dado se puede jugar con las constantes α_j, β_j para intentar encontrar valores para los que la desigualdad se viole, lo que garantizaría que el estado es entrelazado.

La condición de entrelazamiento del teorema es de hecho necesaria y suficiente para estados gaussianos (es decir, su transformada de Wigner es una gaussiana en el espacio de las fases) [3.32]

Como caso particular

$$\alpha_1 = \alpha_2 = \beta_1 = 1 = -\beta_2, \qquad u = x_1 + x_2, \qquad v = p_1 - p_2 \tag{3.88}$$

Como en este caso $[u,v] = 0$ el principio de incertidumbre (la Proposición) no dice nada no trivial: $(\Delta u)^2 + (\Delta v)^2 \geq 0$. Δu y Δv se pueden hacer arbitrariamente pequeños, pero el teorema dice que $(\Delta u)^2 + (\Delta v)^2 < 2$ sólo puede producirse si ρ es un estado entrelazado.

3.4.5. Criterio basado en reordenación de operadores

Cada criterio de entrelazamiento sólo cubre una parte de los casos. Por ejemplo dos fotones en el estado $|\Phi_+\rangle = \frac{1}{\sqrt{2}}(|0\rangle_1|1\rangle_2 + |1\rangle_1|0\rangle_2)$ están entrelazados pero el criterio DGCZ no lo detecta. Aquí $|n_1\rangle_1|n_2\rangle_2$ representa el estado de n_1 fotones en el orbital 1 y n_2 fotones en el orbital 2, [3.33]

$$|n_1\rangle_1|n_2\rangle_2 = \frac{(a_1^\dagger)^{n_1}}{\sqrt{n_1!}}\frac{(a_2^\dagger)^{n_2}}{\sqrt{n_2!}}|0\rangle \tag{3.89}$$

(Para ver que el entrelazamiento de $|\Phi_+\rangle$ no es detectado por el criterio DGCZ en (3.64) hay que usar las relaciones (3.61) para llevar u, v a forma $a_j, a_j^\dagger$, o bien llevar $|\Phi_+\rangle$ a forma x_j, p_j; $|0\rangle_j$ es la gaussiana del estado fundamental del oscilador $\propto e^{-x_j^2/2}$ y $|1\rangle_j$ el primer estado excitado $\propto x_j e^{-x_j^2/2}$.)

Vemos aquí otro criterio (de los muchos que hay) que sí es capaz de ver que ese estado es entrelazado.

Teorema Sea ρ separable y X, Y operadores (no necesariamente hermíticos) en $\mathscr{H}_A$ y $\mathscr{H}_B$ respectivamente, Entonces

$$|\langle XY^\dagger\rangle|^2 \leq \langle X^\dagger X Y^\dagger Y\rangle \tag{3.90}$$

Nótese que:
1) La desigualdad $|\langle XY^\dagger\rangle|^2 \leq \langle XY^\dagger YX^\dagger\rangle$ se cumple siempre, sea ρ entrelazado o separable (por el Lema 1).

[3.32] L-M. Duan, G. Giedke, J. I. Cirac and P. Zoller, *Inseparability criterion for continuous variable systems,* Phys. Rev. Lett. **84**(2000)2722 [3].

[3.33] Un orbital es cualquier estado monoparticular, por ejemplo, con momento y polarización definidas.

2) En general $A, B \geq 0$ no implica $AB \geq 0$, pero sí si A e B conmutan (basta escribirlos en una base diagonal común) por tanto $\langle X^\dagger X Y^\dagger Y \rangle \geq 0$ aunque el estado no sea separable.

Demostración:

$$\begin{aligned} |\langle XY^\dagger \rangle|^2 = \left|\overline{\langle XY^\dagger \rangle_k}\right|^2 &\underset{\text{Cauchy-Schwarz}}{\leq} \overline{|\langle XY^\dagger \rangle_k|^2} \\ &\underset{\rho_k \text{ sep. y } \langle Y^\dagger \rangle = \langle Y \rangle^*}{=} \overline{|\langle X \rangle_k|^2 |\langle Y \rangle_k|^2} \underset{\text{Lema 1}}{\leq} \overline{\langle X^\dagger X \rangle_k \langle Y^\dagger Y \rangle_k} \\ &\underset{\rho_k \text{ sep.}}{=} \overline{\langle X^\dagger X Y^\dagger Y \rangle_k} = \langle X^\dagger X Y^\dagger Y \rangle \end{aligned} \tag{3.91}$$

□

Podemos aplicar este criterio al estado $|\Phi_+\rangle$, con $X = a_1, Y = a_2$. Recordando que para operadores de creación y destrucción bosónicos

$$a|n\rangle = \sqrt{n}|n-1\rangle, \qquad a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle, \qquad \langle n|m\rangle = \delta_{nm}, \qquad a^\dagger a|n\rangle = n|n\rangle \tag{3.92}$$

($a^\dagger a = \hat{N}$ es el operador número)

Entonces

$$\langle a_1 a_2^\dagger \rangle = \langle \Phi_+ | a_1 a_2^\dagger | \Phi_+ \rangle = \frac{1}{2}\langle 0,1|a_1 a_2^\dagger|1,0\rangle = \frac{1}{2}\langle 0|a|1\rangle\langle 1|a^\dagger|0\rangle = \frac{1}{2} \tag{3.93}$$

Mientras que

$$\langle a_1^\dagger a_1 a_2^\dagger a_2 \rangle = \langle \hat{N}_1 \hat{N}_2 \rangle = 0 \tag{3.94}$$

Se concluye que el estado está entrelazado.

3.5. Destilación y formación de entrelazamiento

El entrelazamiento es esencial en computación cuántica (la puerta CNOT entrelaza y es esencial en cualquier circuito, por ejemplo en el problema de Deutsch). Igualmente es un recurso en comunicación cuántica. Los **ebits**, entendidos como un par de qubits máximamente entrelazados compartidos (generalmente entre sistemas separados) ya han mostrado su utilidad en el caso de codificación densa y teleportación. Veamos dos aplicaciones más. Pero antes necesitamos el concepto de **operaciones locales**.

3.5.1. Operaciones locales y comunicación clásica (OLCC)

En comunicación cuántica tenemos dos (o más) estaciones, $\mathscr{H} = \mathscr{H}_A \otimes \mathscr{H}_B$ que están espacialmente separadas. Como ya se ha explicado no es posible medir observables genéricos ni producir estados genéricos (entrelazados) de $\mathscr{H}$ porque las interacciones físicas son locales, no hay acción a distancia, cualquier efecto se produce en un sitio y luego se tiene que propagar para llegar a otro lado y actuar. La propagación siempre es a velocidad sublumínica o lumínica, pero no supralumínica.

Las denominadas operaciones locales y comunicación clásica (OLCC) son realizables. **Comunicación clásica** se refiere al envío o intercambio de bits (equivalentemente enviar qubits físicos siempre en la base computacional). Las **operaciones locales** son acciones que se hacen en $\mathscr{H}_A$ o en $\mathscr{H}_B$ o ambos, pero siempre por separado. Son las siguientes:

i) Operaciones unitarias. Estos son operadores unitarios del tipo $U_A \otimes U_B$. Es decir, actúan en A y/o en B por separado.

ii) Medidas de observables de A y/o de B (observables del tipo $O_A \otimes O_B$).

iii) Añadir estados auxiliares no entrelazados con la otra parte. Por ejemplo aunque inicialmente nos concentrábamos en $\mathscr{H}_A$, en realidad siempre hay más grados de libertad disponibles que simplemente antes no se usaban: $\mathscr{H}_A \otimes \mathscr{H}_{A_1}$ (A_1 está en un entorno local de A). Se puede entonces usar A_1 entrelazando con A, etc. (Como en el protocolo de teleportación).

iv) Descartar parte del sistema A y/o B. Se refiere a tomar traza sobre espacios auxiliares del tipo $\mathscr{H}_{A_1}$, es decir, no utilizar información de esos grados de libertad.

Aparte de esto se pueden intercambiar físicamente estados cuánticos (por ejemplo qubits). Esta operación no está entre las OLCC.

3.5.2. Destilación de ebits (concertación de entrelazamiento)

Andrea y Benito comparten n pares de qubits que no están máximamente entrelazados. La **destilación** consiste en concentrar el entrelazamiento, formar a partir de esos qubits el mayor número posible de ebits (es decir, pares máximamente entrelazados) usando OLCC.

Discutimos aquí un método que no es óptimo, pero sí simple. Inicialmente Andrea y Benito com-

parten un par de qubits en el estado entrelazado [3.34]

$$|\Psi\rangle_{AB} = \cos(\theta)|00\rangle_{AB} + \text{sen}(\theta)|11\rangle_{AB}, \qquad 0 < \theta < \frac{\pi}{4} \tag{3.95}$$

y quieren extraer el estado de Bell $|\Psi_+\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB})$.

El protocolo consta de tres pasos:

1) Andrea añade un qubit en el estado $|0\rangle_{A'}$. El nuevo estado es

$$|\Psi\rangle_{AB} \otimes |0\rangle_{A'} = \cos(\theta)|00\rangle_{AA'} \otimes |0\rangle_B + \text{sen}(\theta)|10\rangle_{AA'} \otimes |1\rangle_B. \tag{3.96}$$

2) Andrea aplica una transformación unitaria U_A (que actúa en AA') tal que

$$\begin{aligned} U_A|00\rangle_{AA'} &= \tan(\theta)|00\rangle_{AA'} + \sqrt{1-\tan^2(\theta)}|01\rangle_{AA'} \\ U_A|10\rangle_{AA'} &= |10\rangle_{AA'} \end{aligned} \tag{3.97}$$

Tal y como está U_A conserva norma y producto escalar, por tanto se puede extender su definición sobre $|01\rangle$ y $|11\rangle$ (de muchas maneras) para que sea un operador unitario en el sistema de dos qubits.

El estado pasa a ser

$$\begin{aligned} &U_A \otimes I_B|\Psi\rangle_{AB} \otimes |0\rangle_{A'} \\ &\quad = \left(\text{sen}(\theta)|00\rangle_{AA'} + \sqrt{1-2\,\text{sen}^2(\theta)}|01\rangle_{AA'}\right) \otimes |0\rangle_B + \text{sen}(\theta)|10\rangle_{AA'} \otimes |1\rangle_B \\ &\quad = \sqrt{2}\,\text{sen}(\theta)|\Psi_+\rangle_{AB} \otimes |0\rangle_{A'} + \sqrt{1-2\,\text{sen}^2(\theta)}|00\rangle_{AB} \otimes |1\rangle_{A'}. \end{aligned} \tag{3.98}$$

3) Andrea mide el qubit A' en la base computacional. Si encuentra $|0\rangle_{A'}$ entonces sabe que tiene el estado $|\Psi_+\rangle$ en AB. Ha destilado un ebit y el proceso ha tenido éxito. Si encuentra $|1\rangle_{A'}$, el estado AB es $|00\rangle_{AB}$, se ha perdido el entrelazamiento y el par de qubits se descarta.

La probabilidad de éxito es $p = 2\,\text{sen}^2(\theta)$. Si AB comparten un número n de estados $|\Psi\rangle_{AB}$ después de aplicar el protocolo acaban con np ebits (y desperdician $n(1-p)$ estados). Como se verá en el teorema BBPS (pág. 37) este valor de p no es el máximo posible.

[3.34] Por la descomposición de Schmidt, cualquier estado bipartito se puede llevar a la forma $|\Psi\rangle_{AB} = \cos(\theta)|00\rangle_{AB} + \text{sen}(\theta)|11\rangle_{AB}$, con $0 \le \theta \le \frac{\pi}{4}$, eligiendo las bases $\{|0\rangle, |1\rangle\}$ adecuadas en A y en B.

3.5.3. Dilución de ebits

En este caso lo que se quiere es producir un estado compartido entre A y B

$$|\psi\rangle_{AB} = \alpha|00\rangle + \beta|11\rangle \tag{3.99}$$

usando OLCC y se dispone del estado compartido $|\Phi_-\rangle_{AB}$ como recurso. No funcionaría que A y B actúen cada uno sobre su qubit, esto es aplicar un operador unitario local $U_A \otimes U_B$

$$\begin{aligned} \rho_B &\longrightarrow \mathrm{Tr}_A(U_A \otimes U_B|\Phi_-\rangle\langle\Phi_-|U_A^\dagger \otimes U_B^\dagger) = U_B\,\mathrm{Tr}_A(|\Phi_-\rangle\langle\Phi_-|)U_B^\dagger = \frac{1}{2}I_B \\ &\neq \mathrm{Tr}_A(|\psi\rangle\langle\psi|_{AB}) = |\alpha|^2|0\rangle\langle 0| + |\beta|^2|1\rangle\langle 1|. \end{aligned} \tag{3.100}$$

Un método (no óptimo) que sí funciona es el siguiente:

1) Andrea añade $\alpha|00\rangle_{A_1A_2} + \beta|11\rangle_{A_1A_2}$ a $|\Phi_-\rangle_{AB}$ y forma un estado $|\chi\rangle_{A_1A_2AB}$.

$$|\chi\rangle_{A_1A_2AB} = (\alpha|00\rangle_{A_1A_2} + \beta|11\rangle_{A_1A_2}) \otimes \frac{1}{\sqrt{2}}(|01\rangle_{AB} - |10\rangle_{AB}) \tag{3.101}$$

2) Andrea teleporta A_2 (o A_1) a B usando el ebit disponible. Para ello mide A_1A en la base de Bell y comunica el resultado a Benito para que aplique el operador de Pauli adecuado:

$$\begin{aligned} {}_{A_1A}\langle\Psi_+|\chi\rangle &= \frac{1}{2}(\alpha|01\rangle - \beta|10\rangle)_{A_2B} = \frac{1}{2}I \otimes (-iY)|\psi\rangle_{A_2B} \\ {}_{A_1A}\langle\Psi_-|\chi\rangle &= \frac{1}{2}(\alpha|01\rangle + \beta|10\rangle)_{A_2B} = \frac{1}{2}I \otimes X|\psi\rangle_{A_2B} \\ {}_{A_1A}\langle\Phi_+|\chi\rangle &= \frac{1}{2}(-\alpha|00\rangle + \beta|11\rangle)_{A_2B} = \frac{1}{2}I \otimes (-Z)|\psi\rangle_{A_2B} \\ {}_{A_1A}\langle\Phi_-|\chi\rangle &= \frac{1}{2}(-\alpha|00\rangle - \beta|11\rangle)_{A_2B} = \frac{1}{2}I \otimes (-I)|\psi\rangle_{A_2B} \end{aligned} \tag{3.102}$$

Al final del proceso AB comparten el estado $|\psi\rangle_{AB}$. Aquí el problema no era transmitir la información $(\alpha\beta)$, que podía ser conocida, como construir localmente un estado entrelazado concreto no local. El método no es óptimo porque han gastado totalmente su ebit para producir un estado de menor entrelazamiento. Se puede hacer con menos gasto de ebits por par producido (Teorema BBPS).

3.6. Medidas de entrelazamiento

Necesitamos cuantificar el grado de entrelazamiento. Hay varias formas de hacerlo y usamos una que es útil en teoría de la información. Empezamos por estados puros. (Los estados mezcla son más complicados, recuérdese que mezclar estados entrelazados puede resultar en una mezcla separable.)

3.6.1. Entropía de von Neumann como medida de entrelazamiento

Como se dijo, si un estado bipartito puro es separable, las matrices densidad reducidas también lo son, y en cambio si son múltiplos de la identidad, el entrelazamiento es máximo. [3.35] Es decir, una medida de entrelazamiento sería cómo de concentrado está el espectro de las matrices densidad reducidas, y eso lo mide la entropía de Shannon del espectro, o equivalentemente la entropía de von Neumann.

Entropía de Shannon

Tenemos un sistema (clásico o cuántico) que puede estar en N estados con etiqueta $k = 1, \ldots, N$ cada uno con probabilidad p_k, equivalentemente un objeto (por ejemplo unas llaves) que estoy buscando y que puede estar en N cajas. Quiero medir cómo de localizado está el objeto. Si la distribución de p_k está muy concentrada en pocos k entonces tengo mucha información sobre dónde está el objeto y si está muy repartida tengo poca información. En un sentido vago **entropía**: lo que **ignoro** cuándo aún no he encontrado el objeto o equivalentemente lo que **aprendo** cuando lo encuentro. (Quien dice un objeto, dice el valor de una variable aleatoria antes y después de hacer el experimento aleatorio y descubrir su valor). La entropía de Shannon cuantifica esa ignorancia (esa cantidad de información):

$$p_k \geq 0, \qquad \sum_k p_k = 1, \qquad H(\{p_k\}) \equiv -\sum_k p_k \log(p_k) = \overline{-\log(p_k)} \qquad \text{(Entropía de Shannon)} \tag{3.103}$$

(Los p_k nulos no contribuyen, $x\log(x) \underset{x\to 0}{\to} 0^+$). Aquí, y en lo que sigue, $\log(x) = \log_2(x) = \ln(x)/\ln(2)$. Ésas son las unidades convencionales en teoría de la información, basada en bits.

Así por ejemplo, si $p_1 = 1$ y $p_k = 0$ para $k > 1$, $H = 0$. Ya se sabe todo y no se gana ninguna información al descubrir que el valor de k correcto es $k = 1$. En el extremo contrario, si $p_k = 1/N$ para todo k, $H = \log(N)$. Éste es el caso de máxima entropía, en el que se tiene menor información inicial. En particular si $N = 2^n$, $H = n$, se dice que se tienen n bits de información (cada k codifica n bits; se necesitan n bits para especificar un valor de k).

[3.35] Aquí estamos implícitamente suponiendo que los dos espacios son de igual dimensión.

Se pueden definir otras entropías cambiando la función log por otras funciones, pero la entropía de Shannon tiene la propiedad de **aditividad**: si lo tengo organizado en contenedores $a = 1,\ldots,N$, con probabilidades q_a, cada uno con cajas $k = 1,\ldots,N_a$, con probabilidades $p_{ak} = \text{Prob}(k|a)$ ($\sum_a q_a = 1$, $\sum_k p_{ak} = 1$) entonces

$$H(\{q_a p_{ak} \mid \forall a,k\}) = H(\{q_a\}) + \sum_a q_a H(\{p_{ak} \mid \forall k\}) \tag{3.104}$$

Es decir, la información que obtengo secuencialmente (primero descubro el valor de a y luego el de k –parte derecha de la ecuación–) es la misma que se obtiene al descubrir ak directamente –parte izquierda de la ecuación–.

Entropía de von Neumann

$$\text{S}(\rho) \equiv -\text{Tr}(\rho \log(\rho)) = \langle -\log(\rho)\rangle_\rho \qquad (\textbf{Entropía de von Neumann}) \tag{3.105}$$

Satisface (siendo $\{|\phi_i\rangle\}$ la base ortonormal propia de ρ y λ_i los valores propios)

$$\rho = \sum_i \lambda_i |\phi_i\rangle\langle\phi_i|, \qquad \text{S}(\rho) = H(\{\lambda_i\}) \tag{3.106}$$

Nótese que la definición de $S(\rho)$ no es $H(\{p_k\})$, siendo $\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k|$ una mezcla de estados normalizados pero no ortogonales (de hecho $H(\{p_k\}) \geq S(\rho)$).

Con esta definición $S(\rho)$ se anula para un estado puro (el espectro no nulo de ρ es 1 –no contribuye a S– y el espectro nulo tampoco contribuye a S). En el otro extremo, para $\rho = \frac{1}{d} I$, se obtiene $S = \log(d)$.

Es interesante notar que (salvo un factor trivial de normalización, $\langle -k_B \ln(\rho)\rangle_\rho$) la entropía de Shannon es la entropía de Gibbs de mecánica estadística clásica. La entropía de von Neumann de un sistema físico no es más que su entropía termodinámica. Para un sistema con estados propios de la energía $\hat{H}|n\rangle = E_n|n\rangle$, poblados con probabilidad p_n, la entropía es $S(\{p_n\})$. Si se busca la distribución que maximice la entropía con $\langle\hat{H}\rangle = \sum_n p_n E_n$ fijo, se obtiene la distribución de Boltzmann $p_n = e^{-\beta E_n}/Z$ ($\beta = 1/T$, unidades $k_B = 1$). En este caso

$$S = -\sum_n p_n \ln(p_n) = \frac{1}{Z}\sum_n e^{-\beta E_n}(\beta E_n + \ln(Z)) \implies U - TS = -T\ln(Z) \tag{3.107}$$

$U = \langle\hat{H}\rangle$ es la energía interna, $-T\ln(Z) = A$ es la energía libre de Helmholtz.

Volvemos al tema de cuantificar el entrelazamiento.

Para un estado bipartito puro $|\psi\rangle_{AB}$ denotamos $E(|\psi\rangle_{AB})$ la cantidad de entrelazamiento del estado, y la definimos como la entropía de von Neumann de las matrices densidad reducidas,

$$E(|\psi\rangle_{AB}) \equiv \mathrm{S}(\rho_A) = \mathrm{S}(\rho_B) \tag{3.108}$$

Esta definición se basa en que para estados puros, ρ_A y ρ_B tienen el mismo espectro no nulo.

Con esta definición $E(|\psi\rangle_{AB})$ se anula para un estado separable ya que en ese caso ρ_A es un estado puro y S se anula. Para un estado máximamente entrelazado $\rho_A = \frac{1}{d_A} I_A$, se obtiene $E = \log(d_A)$. [3.36] Para qubits $0 \leq E \leq 1$.

Veamos otras propiedades que sugieren que $E(\psi)$ es una buena medida del entrelazamiento:

1) El entrelazamiento de sistemas independientes es aditivo.

   Demostración: Supongamos que $\mathscr{H}_A = \mathscr{H}_{A_1} \otimes \mathscr{H}_{A_2}$ y $\mathscr{H}_B = \mathscr{H}_{B_1} \otimes \mathscr{H}_{B_2}$, y además el estado es de la forma

   $$|\psi\rangle = |\phi_1\rangle_{A_1B_1} \otimes |\phi_2\rangle_{A_2B_2} \tag{3.109}$$

   entonces

   $$\rho_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|) = \mathrm{Tr}_{B_1}(|\phi_1\rangle\langle\phi_1|_{A_1B_1}) \otimes \mathrm{Tr}_{B_2}(|\phi_2\rangle\langle\phi_2|_{A_2B_2}) = \rho_{A_1} \otimes \rho_{A_2} \tag{3.110}$$

   ρ_{A_1} y ρ_{A_2} actúan en espacios distintos y por ello conmutan. Entonces [3.37]

   $$\log(\rho_{A_1} \otimes \rho_{A_2}) = \log(\rho_{A_1}) \otimes I_{A_2} + I_{A_1} \otimes \log(\rho_{A_2}) \tag{3.111}$$

   Entonces

   $$\begin{aligned} S(\rho_A) &= -\mathrm{Tr}_A(\rho_A \log(\rho_A)) \\ &= -\mathrm{Tr}_{A_1A_2}\Big(\rho_{A_1} \otimes \rho_{A_2}\big(\log(\rho_{A_1}) \otimes I_{A_2} + I_{A_1} \otimes \log(\rho_{A_2})\big)\Big) \\ &= S(\rho_{A_1}) + S(\rho_{A_2}) \\ E(|\phi_1\rangle \otimes |\phi_2\rangle) &= E(|\phi_1\rangle) + E(|\phi_2\rangle) \end{aligned} \tag{3.112}$$

   □

[3.36] Que este es realmente el máximo se comprueba más abajo.

[3.37] Si X, Y conmutan $e^{X+Y} = e^X e^Y$, si $A, B \geq 0$ y conmutan $\log(AB) = \log(A) + \log(B)$. En general $e^X e^Y = e^{X+Y+\frac{1}{2}[X,Y]+\cdots}$ es la fórmula de Campbell-Hausdorff.

2) $E(|\psi\rangle)$ es conservado bajo operaciones locales unitarias.

Demostración: Una operación unitaria local significa

$$|\psi'\rangle_{AB} = U_A \otimes U_B |\psi\rangle_{AB} \tag{3.113}$$

En este caso, para la matriz densidad reducida (usando la propiedad cíclica de la traza)

$$\rho'_A = \mathrm{Tr}_B \left(U_A \otimes U_B |\psi\rangle\langle\psi|_{AB} U_A^\dagger \otimes U_B^\dagger\right) = U_A \,\mathrm{Tr}_B \left(U_B |\psi\rangle\langle\psi|_{AB} U_B^\dagger\right) U_A^\dagger = U_A \rho_A U_A^\dagger \tag{3.114}$$

Puesto que ρ_A y ρ'_A tienen el mismo espectro y la entropía sólo depende del espectro $S(\rho'_A) = S(\rho_A)$. O también, usando la propiedad $f(TAT^{-1}) = Tf(A)T^{-1}$ para $f(x) = -x\log(x)$, y de nuevo la propiedad cíclica,

$$S(\rho'_A) = -\,\mathrm{Tr}(U_A \rho_A \log(\rho_A) U_A^\dagger) = S(\rho_A) \tag{3.115}$$

Es decir $E(\psi') = E(\psi)$. □

3) En promedio, el valor de E (la cuantificación del entrelazamiento basada en la entropía definida en (3.108)) no puede aumentarse mediante OLCC. La demostración se presenta más abajo (Temas 3.6.4–3.6.6)

4) (Teorema BBPS) La magnitud E es una medida de la eficiencia asintótica óptima con la que el entrelazamiento puede ser concentrado o diluido usando sólo OLCC. [3.38] El teorema es para estados puros de qubits y por tanto $0 \leq E \leq 1$.

Esto significa lo siguiente:

i) Si Andrea y Benito comparten n copias de un estado (parcialmente) entrelazado ψ, con entrelazamiento $f = E(\psi)$, usando un método óptimo (pero que involucre sólo OLCC) pueden llegar a destilar un número $n_e = fn \leq n$ de estados de Bell, cuando $n \to \infty$.

ii) Viceversa, si Andrea y Benito comparten n_e ebits, usando un método óptimo basado en OLCC, pueden llegar a producir hasta $n = n_e/f \geq n_e$ copias de un estado ψ con entrelazamiento $f = E(\psi)$, cuando $n_e \to \infty$.

Obsérvese que en el método de destilación del Tema 3.5.2 a partir de $|\psi\rangle = c|00\rangle + s|11\rangle$ $(0 < s^2 < 1/2)$, la probabilidad de éxito era $p = 2s^2$. Este valor es menor que el óptimo $f = E(\psi) = -c^2\log(c^2) - s^2\log(s^2)$, como exige el teorema. ($f = p$ para en los casos extremos $s = 0$ ó $s = 1/\sqrt{2}$.)

Análogamente el método de dilución descrito en el Tema 3.5.2, de estados de Bell en estados $|\psi\rangle = c|00\rangle + s|11\rangle$ consume un ebit por cada estado. De acuerdo con el teorema BBPS se puede hacer consumiendo sólo una fracción f de ebits.

[3.38] C.H. Bennett, H.J. Bernstein, S. Popescu, y B. Schumacher, *Concentrating partial entanglement by local operations*, Phys. Rev. A**53**(1996)2046 [4].

3.6.2. Entropía relativa y desigualdad de Klein

Necesitamos algunos resultados auxiliares para obtener conclusiones sobre la entropía de von Neumann como medida de entrelazamiento. Para ello definimos la **entropía relativa** de un estado ρ con respecto a otro estado σ,

$$S(\rho\|\sigma) \equiv \mathrm{Tr}\left(\rho(\log(\rho)-\log(\sigma))\right). \tag{3.116}$$

También se denomina "divergencia de Kullback-Leibler", [3.39] principalmente en el contexto clásico. Nótese que puede hacerse $+\infty$ si σ se anula en un subespacio y ρ no.

Una propiedad importante de la entropía relativa es que es **semidefinida positiva**:

$$S(\rho\|\sigma) \geq 0, \qquad S(\rho\|\sigma) = 0 \quad \text{sii} \quad \rho = \sigma. \tag{3.117}$$

Examinamos sólo la positividad.

Definición (Función convexa) Se dice que una función real de variable real $f(x)$ es **convexa** cuando el grafo *por encima* de la curva $\{(x, y \geq f(x))\}$ es un conjunto convexo. Equivale a decir que, para cualquier promedio $\overline{f(x_k)} \geq f(\overline{x_k})$, y también a decir que $f''(x) \geq 0$ (si $f(x)$ es diferenciable dos veces). Se dice que $f(x)$ es **cóncava** cuando $-f(x)$ es convexa.

Lema si $X, Y \geq 0$ también $\mathrm{Tr}(XY) \geq 0$, por $\mathrm{Tr}(XY) = \sum_n X_n \langle n|Y|n\rangle \geq 0$.

Teorema (Desigualdad de Klein) Si A y B son dos matrices hermíticas y $f(x)$ es convexa

$$\mathrm{Tr}(f(A)) - \mathrm{Tr}(f(B)) \geq \mathrm{Tr}\left((A-B)f'(B)\right) \quad \text{(desigualdad de Klein)} \tag{3.118}$$

Demostración: La función

$$F(t) = \mathrm{Tr}\left(f(tA+(1-t)B)\right) \qquad t \in \mathbb{R} \tag{3.119}$$

también es convexa: Suponiendo que f sea dos veces diferenciable, [3.40] una forma de verlo es derivando:

$$\begin{aligned} F'(t) &= \mathrm{Tr}\left((A-B)f'(tA+(1-t)B)\right), \\ F''(t) &= \mathrm{Tr}\left((A-B)^2 f''(tA+(1-t)B)\right) \geq 0. \end{aligned} \tag{3.120}$$

[3.39] En teoría de la información una "divergencia" es una especie de distancia pero asimétrica.

[3.40] Suponemos esto para simplificar la demostración. La desigualdad vale en general.

La última desigualdad se sigue del Lema anterior. Se deduce que $F'(t)$ es creciente y por tanto

$$F(1)-F(0)=\int_0^1 F'(t)\,dt \geq F'(0). \tag{3.121}$$

(Usando $F'(t)\geq F'(0)$.) Sustituyendo en la definición de $F(t)$ se obtiene la desigualdad de Klein. □

Demostración: (de la positividad de la entropía relativa)

Como la función $f(x)=x\ln(x)$ es convexa en $x\geq 0$, la desigualdad de Klein se aplica a matrices densidad. Usando $f'(x)=\ln(x)+1$ y $\mathrm{Tr}(\rho)=\mathrm{Tr}(\sigma)=1$

$$\begin{aligned}\mathrm{Tr}\left(\rho\ln(\rho)-\sigma\ln(\sigma)\right) &\geq \mathrm{Tr}\left((\rho-\sigma)(\ln(\sigma)+I)\right)=\mathrm{Tr}\left(\rho\ln(\sigma)\right)-\mathrm{Tr}\left(\sigma\ln(\sigma)\right)\\ &\Longrightarrow \mathrm{Tr}\left(\rho\ln(\rho)-\rho\ln(\sigma)\right)\geq 0\,.\end{aligned} \tag{3.122}$$

□

3.6.3. Más propiedades de la entropía de von Neumann

Explotando que la entropía relativa es positiva podemos obtener más información sobre la entropía y el entrelazamiento:

i) Para un sistema d-dimensional

$$0\leq S(\rho)\leq \log(d)\,. \tag{3.123}$$

Demostración: Puesto que $-x\log(x)\geq 0$ para $0\leq x\leq 1$, la cota $0\leq S(\rho)$ es obvia. Por otro lado, si consideramos el estado $\sigma=\dfrac{1}{d}I$ y la entropía relativa de ρ y σ

$$0\leq S(\rho\|\sigma)=-S(\rho)-\langle\log(\sigma)\rangle_\rho=-S(\rho)+\log(d)\,. \qquad \square \tag{3.124}$$

Para un sistema de n qubits $d=2^n$, $0\leq S\leq n$. En general $S_{\text{máx}}$ cuenta el número de grados de libertad de un sistema.

ii) Si tenemos un conjunto de N de estados mezcla $\{\rho_k\}$ en $\mathscr{H}_A$ y una base ortonormal $|k\rangle$ de un espacio $\mathscr{H}_B$ (con dimensión al menos N) y unos pesos p_k, se puede formar una **mezcla disjunta** en el espacio producto

$$\tilde{\rho}=\sum_k p_k\,\rho_k\otimes|k\rangle\langle k|=\overline{\rho_k\otimes|k\rangle\langle k|}\,. \tag{3.125}$$

Es disjunta porque todavía se puede distinguir cada ρ_k, no están realmente mezclados, como sería en $\rho_A = \sum_k p_k \rho_k = \mathrm{Tr}_B(\tilde{\rho})$.

Entonces

$$S(\tilde{\rho}) = H(\{p_k\}) + \overline{S(\rho_k)} \tag{3.126}$$

Demostración: $\tilde{\rho}$ es diagonal en bloques, en cada bloque $|k\rangle\langle k|$ es la identidad en B, por tanto

$$\begin{aligned} \tilde{\rho}\log(\tilde{\rho}) &= \sum_k p_k \rho_k \big(\log(p_k) I_A + \log(\rho_k)\big) \otimes |k\rangle\langle k| \\ S(\tilde{\rho}) &= -\mathrm{Tr}\left(\tilde{\rho}\log(\tilde{\rho})\right) = -\sum_k p_k \mathrm{Tr}_A\left(\rho_k\big(\log(p_k) I_A + \log(\rho_k)\big)\right) \\ &= H(\{p_k\}) + \sum_k p_k S(\rho_k) \qquad \square \end{aligned} \tag{3.127}$$

iii) La entropía es **subaditiva**. Es decir, para un sistema bipartito con estado mezcla ρ_{AB}

$$S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B) \qquad \text{(Subaditividad de la entropía)} \tag{3.128}$$

Aquí $\rho_A = \mathrm{Tr}_B(\rho_{AB})$ y $\rho_B = \mathrm{Tr}_A(\rho_{AB})$. De hecho se satisface la propiedad más general

$$|S(\rho_A) - S(\rho_B)| \leq S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B) \qquad \text{(Desigualdad triangular de la entropía)} \tag{3.129}$$

Nótese que para estados puros ρ_A y ρ_B tenían el mismo espectro no nulo y en consecuencia la misma entropía, pero eso ya no es así para estados mezcla AB. Por ejemplo

$$\rho_{AB} = \rho_A \otimes \rho_B, \qquad \rho_A = |0\rangle\langle 0|, \quad \rho_B = \frac{1}{2} I_B. \tag{3.130}$$

Demostración: Aplicamos la positividad de la entropía relativa para $\rho = \rho_{AB}$ y $\sigma = \rho_A \otimes \rho_B$,

$$\begin{aligned} 0 \leq S(\rho_{AB} \| \rho_A \otimes \rho_B) &= \mathrm{Tr}\left(\rho_{AB}\log(\rho_{AB})\right) - \mathrm{Tr}\left(\rho_{AB}\log(\rho_A \otimes \rho_B)\right) \\ &= -S(\rho_{AB}) - \mathrm{Tr}\left(\rho_{AB}\log(\rho_A) \otimes I_B\right) - \mathrm{Tr}\left(\rho_{AB} I_A \otimes \log(\rho_B)\right) \\ &= -S(\rho_{AB}) + S(\rho_A) + S(\rho_B) \end{aligned} \tag{3.131}$$

$\square$

iv) Finalmente, una propiedad de utilidad inmediata, a saber la *concavidad de la entropía*:

Para una mezcla de estados mezcla ρ_k con pesos p_k $(p_k \geq 0, \sum_k p_k = 1)$

$$S(\overline{\rho_k}) \geq \overline{S(\rho_k)} \qquad (S(\rho) \text{ es una función cóncava}) \tag{3.132}$$

Es decir, $S(\sum_k p_k\rho_k) \geq \sum_k p_k S(\rho_k)$.

Demostración: Empleando un espacio auxiliar $\mathscr{H}_B$ podemos formar la mezcla disjunta $\tilde{\rho} = \overline{\rho_k \otimes |k\rangle\langle k|}$, que como hemos visto en (3.126), satisface

$$S(\tilde{\rho}) = \overline{S(\rho_k)} + H(\{p_k\})\,. \tag{3.133}$$

Por otro lado

$$\rho_A = \overline{\rho_k}, \qquad \rho_B = \overline{|k\rangle\langle k|} \tag{3.134}$$

como ρ_B ya está en forma diagonal

$$S(\rho_B) = H(\{p_k\}). \tag{3.135}$$

Entonces, por la desigualdad triangular,

$$S(\tilde{\rho}) \leq S(\rho_A) + S(\rho_B) = S(\overline{\rho_k}) + H(\{p_k\}) \tag{3.136}$$

y se concluye que $\overline{S(\rho_k)} \leq S(\overline{\rho_k})$. □

La concavidad de la entropía implica que mezclar más nunca disminuye la entropía.

3.6.4. Efecto de medidas locales

Aparte de comunicación clásica, había cuatro acciones locales que cualifican como OLCC, a saber, i) operaciones unitarias, ii) medidas proyectivas, iii) añadir estados y iv) descartar parte del sistema.

Ya se ha visto (ec. (3.115)) que las operaciones unitarias locales no cambian el entrelazamiento compartido entre A y B. Igual pasa al añadir estados localmente: si se tiene $|\psi\rangle_{AB}$, el estado $|\phi\rangle_{A'} \otimes |\psi\rangle_{AB}$ tiene el mismo entrelazamiento entre A y B (o entre AA' y B) que antes ya que ρ_B es el mismo.

Veamos qué ocurre con las dos operaciones locales que faltan. Empezamos con medidas locales.

Tenemos $|\psi\rangle_{AB} \in \mathscr{H} = \mathscr{H}_A \otimes \mathscr{H}_B$. Se quiere investigar qué efecto tiene realizar medidas (por supuesto locales, digamos en A) sobre el entrelazamiento compartido.

Un observable k que puede tomar n valores distintos en A equivale a decir hay una base ortonormal asociada $\{|k,\alpha\rangle_A, \alpha = 1,\ldots,r_k,\ k = 1,\ldots,n\}$, aquí r_k es la degeneración del valor k. Sea $\{|j\rangle_B\}$ una base ortonormal de $\mathscr{H}_B$.

Un estado cualquiera

$$|\psi\rangle_{AB} = \sum_{k,j} \sum_{\alpha=1}^{r_k} \psi_{k\alpha j} |k,\alpha\rangle_A \otimes |j\rangle_B \tag{3.137}$$

puede expresarse como una suma normalizada de estados ortogonales

$$|\psi\rangle_{AB} = \sum_k |\tilde{\psi}_k\rangle_{AB}, \qquad |\tilde{\psi}_k\rangle_{AB} \equiv \sum_j \sum_{\alpha=1}^{r_k} \psi_{k\alpha j} |k,\alpha\rangle_A \otimes |j\rangle_B, \qquad \sum_k \| |\tilde{\psi}_k\rangle_{AB} \|^2 = 1 \tag{3.138}$$

Hacer una **medida proyectiva selectiva** (sinónimo de **filtrante**) de k consiste en quedarse con un sumando

$$|\psi\rangle_{AB} \to |\tilde{\psi}_k\rangle_{AB} \tag{3.139}$$

y eso ocurre con probabilidad

$$p_k = \| |\tilde{\psi}_k\rangle_{AB} \|^2. \tag{3.140}$$

Después de la medida selectiva con resultado k el estado pasa a ser $|\tilde{\psi}_k\rangle_{AB}$ antes de normalizar, o

$$|\psi_k\rangle_{AB} = \frac{1}{\sqrt{p_k}} |\tilde{\psi}_k\rangle_{AB} \tag{3.141}$$

después de normalizar. Nótese que

$$|\tilde{\psi}_k\rangle_{AB} = P_k |\psi\rangle_{AB} \tag{3.142}$$

siendo P_k el proyector ortogonal sobre el subespacio (con valor propio) k de $\mathscr{H}_A$, de dimensión r_k, $\sum_k P_k = I_A$ (según el contexto identificamos P_k con $P_k \otimes I_B$).

Efecto de medidas selectivas

Veamos qué ocurre con el entrelazamiento después de la medida selectiva del número cuántico k.

Si k *no está degenerado* ($r_k = 1$) claramente $|\tilde{\psi}_k\rangle_{AB}$ es separable, el entrelazamiento pasa a ser 0.

Si k está degenerado el entrelazamiento puede incluso *aumentar*. Veamos un ejemplo: A tiene dos qubits (k y α) y B uno, y el estado es

$$\begin{aligned} |\psi\rangle_{AB} &= \sqrt{\frac{p}{2}} \left(\overset{k\alpha}{|00\rangle}_A \otimes |0\rangle_B + |01\rangle_A \otimes |1\rangle_B \right) + \sqrt{1-p} |10\rangle_A \otimes |0\rangle_B, \\ &= \sqrt{p} |0\rangle_k \otimes |\Psi_+\rangle_{\alpha B} + \sqrt{1-p} |1\rangle_k \otimes |00\rangle_{\alpha B}, \qquad 0 \le p \le 1\,. \end{aligned} \tag{3.143}$$

Para este estado

$$\rho_B = \left(1 - \frac{p}{2}\right)|0\rangle\langle 0| + \frac{p}{2}|1\rangle\langle 1| \tag{3.144}$$

y recuérdese que $E(\psi_{AB}) = S(\rho_B)$. Se ve que cuando $p \ll 1$, ρ_B es casi un estado puro y por tanto $E \ll 1$.

Si ahora mido k y resulta $k = 0$ (lo cual ocurre con probabilidad p) el estado pasa a ser

$$|\psi_0\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|00\rangle_A \otimes |0\rangle_B + |01\rangle_A \otimes |1\rangle_B\right) = |0\rangle_k \otimes |\Psi_+\rangle_{\alpha B} \tag{3.145}$$

que está máximamente entrelazado (la matriz reducida B pasa a ser $\frac{1}{2}I_B$).

Pero esto sólo ocurre con probabilidad $p \ll 1$. De hecho, no se puede aumentar el entrelazamiento *sistemáticamente* haciendo medidas selectivas, al revés, el entrelazamiento tiende a decrecer:

$$E(|\psi\rangle_{AB}) \geq \overline{E(|\psi_k\rangle_{AB})} \tag{3.146}$$

En el ejemplo anterior $E(\psi_{AB}) = S(\rho_B) = H(p/2)$ siendo

$$H(x) \equiv -x\log(x) - (1-x)\log(1-x), \tag{3.147}$$

la entropía de Shannon de un bit. Mientras que $\overline{E(|\psi_k\rangle_{AB})} = p \times 1 + (1-p) \times 0 = p$, y en efecto $p \leq H(p/2)$.

Demostración: (de la ec. (3.146)). La demostración se basa en la concavidad de la entropía.

$$\overline{E(|\psi_k\rangle_{AB})} = \overline{S(\rho_{B,k})} \leq S(\overline{\rho_{B,k}}) = S(\rho_B) = E(\psi_{AB}) \tag{3.148}$$

Se ha usado la propiedad

$$\overline{\rho_{B,k}} = \rho_B, \tag{3.149}$$

En efecto

$$\begin{aligned} \overline{\rho_{B,k}} &= \sum_k p_k \rho_{B,k} = \mathrm{Tr}_A\left(\sum_k p_k |\psi_k\rangle\langle\psi_k|\right) = \mathrm{Tr}_A\left(\sum_k |\tilde{\psi}_k\rangle\langle\tilde{\psi}_k|\right) \\ &= \mathrm{Tr}_A\left(\sum_k P_k|\psi\rangle\langle\psi|P_k\right) = \mathrm{Tr}_A\left(\sum_k P_k|\psi\rangle\langle\psi|\right) = \mathrm{Tr}_A\left(|\psi\rangle\langle\psi|\right) = \rho_B \end{aligned} \tag{3.150}$$

Se ha usado que P_k actúa sólo en $\mathscr{H}_A$ para aplicar la propiedad cíclica, así como $\sum_k P_K = I_A$. □

La relación $\overline{\rho_{B,k}} = \rho_B$ equivale a decir que una medida no filtrante no tiene efecto para B.

Efecto de medidas no selectivas

Las medidas no filtrantes no tienen efecto alguno a nivel clásico, pero sí cuánticamente para el sistema AB completo. Después de una medida no filtrante el estado puro puede pasar a estado mezcla

$$\rho'_{AB} = \overline{\rho_{AB,k}} = \sum_k P_k |\psi\rangle\langle\psi| P_k = \sum_k |\tilde{\psi}_k\rangle\langle\tilde{\psi}_k| \tag{3.151}$$

Sin embargo la medida no tiene efecto sobre ρ_B: como se acaba de ver $\rho'_B = \overline{\rho_{B,k}} = \rho_B$. Por lo tanto el entrelazamiento no cambia. [3.41]

Si hay una colección formada por muchas copias de estados $|\psi\rangle_{AB}$, una medida selectiva equivale a decir que tras la medida los estados se clasifican por el resultado de k en cada caso. Eso permite tratar de forma distinta cada clase k. No filtrante es que la colección no se clasifica, se usan todos los estados postmedida por igual, por ejemplo en promedios.

Desde el punto de vista de B (que está separado de A) "filtrante" quiere decir que A comunica el resultado de sus medidas (su clasificación en clases etiquetadas por k) a B por el canal clásico, lo cual requiere conexión causal.

En otro caso, es decir si no hay comunicación clásica, para B la medida es *no filtrante*, y ρ_B es exactamente el mismo antes y después de la medida, ya que va a usar toda la colección de estados por igual. Es decir, para cualquier observable $\mathscr{O}_B$ de $\mathscr{H}_B$, no cambia ni su valor esperado ni la distribución de probabilidad de los resultados de medirlo (la probabilidad de cada valor del espectro de $\mathscr{O}_B$ en el estado ρ_B).

Esto es consistente con que el solo hecho de hacer la medida en A no permite enviar información supralumínica de A a B.

3.6.5. Medida de entrelazamiento en estados mezcla

Desechar estados es tomar traza sobre el espacio descartado. Eso puede producir estados mezcla. [3.42] Se concluye que es necesario definir una medida de entrelazamiento para estados mezcla.

[3.41] En realidad aún no hemos definido el entrelazamiento de mezclas, en este caso ρ'_{AB}, pero se puede anticipar que las medidas no filtrantes no tienen efecto sobre el entrelazamiento.

[3.42] El argumento cualitativo "tomar traza es mezclar" no es válido siempre, por ejemplo, $\mathrm{Tr}(\rho) = 1$, 1 es un estado puro del espacio remanente $\mathbb{C}$. De hecho, dada un matriz densidad ρ_{AB}, $S(\rho_A)$ puede ser mayor o menor que $S(\rho_{AB})$.

El entrelazamiento de un ebit es 1. Entonces la idea es definir E como el número medio de ebits requeridos para formar un estado $|\psi\rangle_{AB}$ o un estado ρ_{AB} (se entiende, usando un método óptimo)

Por el Th. BBPS (pág. 37) para estados puros $S(\rho_B)$ es justamente el número medio de pares de Bell necesarios para formar $|\psi\rangle_{AB}$, de ahí la definición $E(\psi) = S(\rho_B)$.

Entonces para definir $E(\rho_{AB})$ se sigue la misma idea: si ρ_{AB} se obtiene como la mezcla $\sum_k p_k |\psi_k\rangle\langle\psi_k|$, formar $|\psi_k\rangle$ requiere $E(\psi_k)$ ebits, y formar esa mezcla concreta requiere $\overline{E(\psi_k)}$. Eso nos da una primera estimación de $E(\rho_{AB})$. Sin embargo no basta. Por ejemplo, para dos qubits con $\rho_{AB} = \frac{1}{4} I_{AB}$, se puede descomponer en la base computacional o en estados de Bell y las estimaciones son distintas (0 y 1, respectivamente).

Como hay muchas descomposiciones de un mismo ρ_{AB}, el *mínimo* número de ebits necesarios es el que se adopta como medida de entrelazamiento:

$$E(\rho_{AB}) = \text{ínf}\left\{ \overline{E(\psi_k)} \right\} \tag{3.152}$$

El ínfimo (mayor de las cotas inferiores) es sobre todas las mezclas $(p_k, |\psi_k\rangle_{AB})$ que produzcan el mismo ρ_{AB}. Es una magnitud bien definida pero no fácil de determinar en general.

Hay otras definiciones y ésta es el que se denomina **entrelazamiento de formación** del estado ρ_{AB}.

Obviamente cuando el estado es puro $\rho_{AB} = |\psi\rangle\langle\psi|$, (y por tanto la descomposición es única por ser ρ extremal) $E(\rho_{AB}) = E(\psi)$.

También, si ρ_{AB} es separable se verifica que $E(\rho_{AB}) = 0$.

Demostración: si ρ_{AB} es separable entonces se puede escribir como una mezcla de estados puros separables y para una de esas descomposiciones $E(\psi_k) = 0$ y $\overline{E(\psi_k)} = 0$, que es el valor mínimo posible para E, eso implica que es el ínfimo. □

Otra propiedad importante:

Teorema $E(\rho)$ es una función convexa,

$$E(\overline{\rho_k}) \leq \overline{E(\rho_k)} \qquad (E(\rho) \text{ es una función convexa }) \tag{3.153}$$

Demostración: En efecto, las mejores descomposiciones en estados puros para cada ρ_k (las que dan el ínfimo de E) proporcionan una descomposición para $\rho := \sum_k p_k \rho_k$. Esa descomposición produce una cota superior del ínfimo sobre todas las descomposiciones posibles de ρ. □

La convexidad de $E(\rho)$ implica que mezclar no puede aumentar el entrelazamiento.

3.6.6. Efecto de desechar información local

Descartar o desechar información técnicamente es tomar traza sobre ciertos grados de libertad. Es el análogo cuántico de marginalizar una distribución de probabilidad. Equivale a ignorar esa información. No se puede aumentar el entrelazamiento desechando información.

Teorema Sea $\mathscr{H} = \mathscr{H}_A \otimes \mathscr{H}_B$ y $\mathscr{H}_A = \mathscr{H}_{A_1} \otimes \mathscr{H}_{A_2}$. Dados ρ_{AB} y $\rho_{A_1B} = \mathrm{Tr}_{A_2}(\rho_{AB})$, entonces

$$E(\rho_{AB}) \geq E(\rho_{A_1B}) \tag{3.154}$$

Demostración:

Consideramos primero el caso en que ρ_{AB} es un estado puro $|\psi\rangle_{AB}$. En este caso $E(\rho_{AB}) = E(\psi_{AB}) = S(\rho_B)$ y $\rho_B = \mathrm{Tr}_A(|\psi\rangle\langle\psi|_{AB}) = \mathrm{Tr}_{A_1}(\rho_{A_1B})$.

Como toda matriz densidad, ρ_{A_1B} admite una descomposición (no única) en estados puros de $\mathscr{H}_{A_1} \otimes \mathscr{H}_B$

$$\rho_{A_1B} = \sum_k p_k \rho_{A_1B,k}, \qquad \rho_{A_1B,k} \text{ estado puro} \tag{3.155}$$

Para esos estados $E(\rho_{A_1B,k}) = S(\rho_{B,k})$ siendo $\rho_{B,k} = \mathrm{Tr}_{A_1}(\rho_{A_1B,k})$. Entonces

$$E(\rho_{AB}) = S(\rho_B) = S(\overline{\rho_{B,k}}) \geq \overline{S(\rho_{B,k})} = \overline{E(\rho_{A_1B,k})} \geq E(\rho_{A_1B})\,. \tag{3.156}$$

Se ha usado la concavidad de la entropía, así como la relación $\overline{\rho_{B,k}} = \rho_B$.

Ahora supongamos que ρ_{AB} es una mezcla, y sea

$$\rho_{AB} = \sum_k p_k \rho_{AB,k}, \qquad (\rho_{AB,k} \text{ estados puros}) \tag{3.157}$$

concretamente la descomposición óptima en estado estados puros $\rho_{AB,k}$ tal que

$$E(\rho_{AB}) = \overline{E(\rho_{AB,k})}. \tag{3.158}$$

Por el resultado anterior, al descartar A_2 se tendrá

$$E(\rho_{AB}) \geq \overline{E(\rho_{A_1B,k})}, \qquad \rho_{A_1B,k} \equiv \mathrm{Tr}_{A_2}(\rho_{AB,k}). \tag{3.159}$$

Finalmente por convexidad de E

$$E(\rho_{AB}) \geq E(\overline{\rho_{A_1B,k}}) = E(\rho_{A_1B}). \tag{3.160}$$

□

En definitiva, el entrelazamiento es un recurso (permite o facilita hacer cosas). Para entrelazar hace falta una interacción y la interacción (en particular medidas) es siempre local (no hay acción a distancia). El entrelazamiento se crea localmente y luego se pueden separar los subsistemas entrelazados.

OLCC no puede producir entrelazamiento (al menos de manera sistemática). Sí es posible destilar entrelazamiento (gastando entrelazamiento parcial ya existente).

3.6.7. Concurrencia

Evaluar el entrelazamiento de formación de un estado mezcla es difícil en general, sin embargo para el caso especial de dos qubits sí hay una fórmula que lo proporciona.

Para un estado puro de dos qubits en bases estándar [3.43] y de Schmidt

$$|\psi\rangle = \sum_{j,k=0}^{1} \psi_{jk}|j\rangle \otimes |k\rangle = \sum_{\alpha=1}^{2} \sqrt{\lambda_\alpha}|u_\alpha\rangle \otimes |v_\alpha\rangle, \qquad \lambda_1 \geq \lambda_2 \geq 0, \qquad \lambda_1 + \lambda_2 = 1\,. \tag{3.161}$$

El estado es separable si y sólo si $\lambda_2 = 0$. $\lambda_1 = \lambda_2 = 1/2$ corresponde a máximo entrelazamiento. En general

$$\rho_A = \sum_{\alpha=1}^{2} \lambda_\alpha |u_\alpha\rangle\langle u_\alpha|, \qquad \mathrm{E}(\psi) = S(\rho_A) = H(\lambda_1) \tag{3.162}$$

siendo $H(x)$ la entropía de un bit, (3.147).

$\sqrt{\lambda_{1,2}}$ son los módulos de los valores propios de la matriz ψ_{jk}. Dado que sólo hay un parámetro independiente (digamos λ_1) el entrelazamiento se puede expresar como una función del determinante de esa matriz

$$\frac{1}{2}C(\psi) \equiv \sqrt{\lambda_1\lambda_2} = |\det(\psi_{jk})|, \qquad 0 \leq C \leq 1 \tag{3.163}$$

C es la denominada **concurrencia** del estado ψ. En función de la concurrencia

$$\lambda_{1,2} = \frac{1}{2}(1 \pm \sqrt{1 - C^2}) \tag{3.164}$$

[3.43] En todo caso, una base ortonormal separable del sistema bipartito.

y por tanto

$$E(\psi) = H\left(\frac{1+\sqrt{1-C^2}}{2}\right) \equiv \mathscr{E}(C) \tag{3.165}$$

Una forma de obtener el determinante de una matriz $n \times n$ A es mediante la expresión

$$\det(A) = \frac{1}{n!}\varepsilon_{i_1\ldots i_n}\varepsilon_{j_1\ldots j_n}A_{i_1 j_1}\cdots A_{i_n j_n} \tag{3.166}$$

siendo $\varepsilon_{i_1\ldots i_n}$ el tensor completamente antisimétrico con $\varepsilon_{1\ldots n} = 1$. Aplicado a qubits, teniendo en cuenta que $\varepsilon_{jk} = (i\sigma_y)_{jk}$, se tiene

$$C(\psi) = |\langle\psi|\tilde{\psi}\rangle|, \qquad |\tilde{\psi}\rangle \equiv \sigma_y \otimes \sigma_y |\psi^*\rangle \tag{3.167}$$

y $|\psi^*\rangle$ se obtiene conjugando las componentes ψ_{jk}. (Nótese que para un qubit $|\widetilde{\hat{n}}\rangle = |-\hat{n}\rangle$.)

Esto es innecesario para estados puros, pero su utilidad es que se extiende a estados mezcla. Para un estado mezcla ρ se define la matriz

$$\tilde{\rho} \equiv \sigma_y \otimes \sigma_y \rho^* \sigma_y \otimes \sigma_y \tag{3.168}$$

y se define la concurrencia $C(\rho)$ como [3.44]

$$C(\rho) \equiv \text{máx}(0, \sqrt{\mu_1} - \sqrt{\mu_2} - \sqrt{\mu_3} - \sqrt{\mu_4}) \tag{3.169}$$

siendo $\mu_1 \geq \mu_2 \geq \mu_3 \geq \mu_4$ los autovalores de $\rho\tilde{\rho}$. En este caso, puede demostrarse [3.45] que el entrelazamiento de formación de ρ satisface

$$E(\rho) = \mathscr{E}(C(\rho)). \tag{3.170}$$

3.6.8. Apéndice: Entrelazamiento bloqueado

Debe notarse que no todos los estados mezcla pueden ser destilados para concentrar el entrelazamiento mediante OLCC. Esto no contradice el teorema BBPS porque ese resultado es para estados puros. Los estados que no pueden usarse para destilación de ebits se denominan **bloqueados** (para

[3.44] Propiamente la definición es $C(\rho) = \text{ínf}(\overline{C(\psi_k)})$ sobre todas las posibles descomposiciones de ρ, pero equivale a la dada.

[3.45] W.K. Wootters, *Entanglement of formation of an arbitrary state of two qubits*, Phys. Rev. Lett. 80(1998)2245 [5].

entrelazamiento). Se puede probar que si un estado bipartito entrelazado pasa el test TPP (su traspuesto parcial es positivo) entonces su entrelazamiento está bloqueado y no es destilable.

Para mostrar un ejemplo de entrelazamiento bloqueado, introducimos primero otro criterio de entrelazamiento de mezclas bipartitas:

Criterio del recorrido. Si un estado mezcla ρ_{AB} es separable entonces existe un conjunto de estados separables $|\psi_{Ak}\rangle \otimes |\psi_{Bk}\rangle$ que subtiende el recorrido de ρ_{AB}, y además los estados $|\psi_{Ak}\rangle \otimes |\psi_{Bk}^*\rangle$ subtienden el recorrido de $\rho_{AB}^{T_B}$ (ambos, traspuesto y conjugado tomados en la misma base ortonormal de B, por lo demás arbitraria).

Queremos un estado entrelazado que pase el criterio TPP, eso no es posible en $\mathbb{C}^2 \otimes \mathbb{C}^2$ o $\mathbb{C}^2 \otimes \mathbb{C}^3$. Consideramos $\mathbb{C}^3 \otimes \mathbb{C}^3$ (dos qutrits, dimensión 9). Definimos el siguiente conjunto de estados ortonormales separables

$$\begin{aligned} |v_0\rangle &= \frac{1}{\sqrt{2}}|0\rangle(|0\rangle - |1\rangle), \qquad |v_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)|2\rangle, \qquad |v_2\rangle = \frac{1}{\sqrt{2}}|2\rangle(|1\rangle - |2\rangle), \\ |v_3\rangle &= \frac{1}{\sqrt{2}}(|1\rangle - |2\rangle)|0\rangle, \qquad |v_4\rangle = \frac{1}{3}(|0\rangle + |1\rangle + |2\rangle)(|0\rangle + |1\rangle + |2\rangle) \end{aligned} \tag{3.171}$$

Este es un **conjunto inextensible**: aunque no son suficientes para formar una base ortonormal, no hay más estados *separables* que sean ortogonales a los ya presentes.

Entonces, el estado mezcla

$$\rho_{AB} = \frac{1}{4}(I - P), \qquad P \equiv \sum_{j=0}^{4} |v_j\rangle\langle v_j|, \tag{3.172}$$

es entrelazado pero pasa el test TPP. En efecto, $I - P$ es el proyector sobre el espacio ortogonal a los $|v_j\rangle$, ese espacio no tiene estados separables, por tanto el recorrido de ρ_{AB} no admite una base separable y por el criterio del recorrido, debe ser entrelazado. [3.46] Por otro lado, en nuestro caso $\rho_{AB}^{T_B} = \rho_{AB}$ que es positivo. Esto es inmediato porque tanto I como P son invariantes bajo T_B, ya que los $|v_j\rangle$ son separables y las funciones de onda son reales:

$$|v_j\rangle\langle v_j| = \rho_{A,j} \otimes \rho_{B,j} \xrightarrow[T_B]{} \rho_{A,j} \otimes \rho_{B,j}^T = \rho_{A,j} \otimes \rho_{B,j}^* = \rho_{A,j} \otimes \rho_{B,j} \tag{3.173}$$

Luego ρ_{AB} pasa el test TPP aunque es entrelazado. Eso implica que no se puede destilar.

[3.46] Obsérvese que tanto I como P son operadores positivos separables, pero ρ_{AB} no es una combinación convexa de I y P ya que hay pesos negativos (se trata por tanto de una extrapolación en vez de una interpolación entre los dos estados) y el estado resulta ser entrelazada.

4. Dinámica cuántica generalizada

Un sistema A en interacción con otro E (el entorno o ambiente) se denomina un **sistema abierto**. A y E pueden intercambiar información, así como energía, partículas, etc. La evolución del sistema completo es unitaria, pero la evolución de la matriz densidad reducida del subsistema, ρ_A, no es unitaria en general. La evolución de ρ_A es descrita por un **canal cuántico**.

4.1. Canales cuánticos

4.1.1. Canales cuánticos y operadores de Kraus

Tenemos un sistema $\mathscr{H}_A \otimes \mathscr{H}_E$ (A va a ser el sistema abierto y E el ambiente) con operador de evolución U_{AE}. Para un estado separable

$$|\Psi\rangle_{AE} = |\psi\rangle_A \otimes |\phi\rangle_E \to |\Psi'\rangle_{AE} = U_{AE}|\Psi\rangle_{AE}\,. \tag{4.1}$$

Nos interesa la evolución en el sector A. $|\phi\rangle_E$ es un estado dado de E y queremos estudiar la dependencia lineal del resultado con respecto del estado $|\psi\rangle_A$. Sea $\{|m\rangle_E\}_{m=1}^M$ una base ortonormal de $\mathscr{H}_E$,

$$|\Psi'\rangle_{AE} = \sum_m (A_m|\psi\rangle_A) \otimes |m\rangle_E \qquad A_m|\psi\rangle_A \equiv {}_E\langle m|U_{AE}(|\psi\rangle_A \otimes |\phi\rangle_E) = {}_E\langle m|\Psi'\rangle_{AE}\,. \tag{4.2}$$

Esta relación define los operadores lineales A_m, que actúan en $\mathscr{H}_A$. Son los denominados **operadores de Kraus**; además de U_{AE}, dependen de $|\phi\rangle_E$ y de la elección de base $\{|m\rangle_E\}$.

Por construcción los operadores de Kraus satisfacen la relación de normalización

$$\sum_{m=1}^{M} A_m^\dagger A_m = I_A \tag{4.3}$$

Demostración:

$$\langle\psi_1|\sum_m A_m^\dagger A_m|\psi_2\rangle = \sum_m {}_A\langle\psi_1| \otimes {}_E\langle\phi|U_{AE}^\dagger|m\rangle_E\,{}_E\langle m|U_{AE}|\psi_2\rangle_A \otimes |\phi\rangle_E = \langle\psi_1|\psi_2\rangle \qquad \square \tag{4.4}$$

Veamos cómo evoluciona ρ_A

$$\begin{aligned} |\psi\rangle\langle\psi|_A &= \mathrm{Tr}_E\left(|\Psi\rangle\langle\Psi|\right) \to \mathrm{Tr}_E\left(|\Psi'\rangle\langle\Psi'|\right) \\ &= \sum_m {}_E\langle m|\Psi'\rangle\langle\Psi'|m\rangle_E = \sum_m A_m|\psi\rangle\langle\psi|A_m^\dagger \end{aligned} \tag{4.5}$$

es decir

$$\rho \to \rho' = T(\rho) \equiv \sum_{m=1}^{M} A_m \rho A_m^\dagger, \qquad \sum_{m=1}^{M} A_m^\dagger A_m = I. \tag{4.6}$$

4.1.1.1. Apartado matemático: superoperadores

Dado un espacio de Hilbert $\mathscr{H}$, $\mathscr{B}(\mathscr{H})$ denota el conjunto de **operadores acotados** definidos en $\mathscr{H}$. [4.1] Una aplicación lineal $T : \mathscr{B}(\mathscr{H}) \to \mathscr{B}(\mathscr{H}')$ se suele denominar **superoperador**. T está definido por su acción sobre una base $\{|i\rangle\langle j|\}$ de $\mathscr{B}(\mathscr{H})$,

$$T(|i\rangle\langle j|) = \sum_{k,\ell} T_{k\ell,ij} |k\rangle\langle\ell| \quad \Longrightarrow \quad \forall \mathscr{O} \in \mathscr{B}(\mathscr{H}) \quad T(\mathscr{O}) = \sum_{i,j,k,\ell} T_{k\ell,ij} |k\rangle\langle i|\mathscr{O}|j\rangle\langle\ell| \tag{4.7}$$

Aquí $\{|k\rangle\langle\ell|\}$ es una base $\mathscr{B}(\mathscr{H}')$. Por tanto T siempre puede expresarse en la forma $T(\mathscr{O}) = \sum_\alpha A_\alpha \mathscr{O} B_\alpha^\dagger$, donde $A_\alpha : \mathscr{H} \to \mathscr{H}'$ y lo mismo B_α.

Cuando $T(\mathscr{O}) = \sum_\alpha A_\alpha \mathscr{O} A_\alpha^\dagger$, se dice que el superoperador admite una **representación de Kraus**. La representación de Kraus está **normalizada** cuando $\sum_\alpha A_\alpha^\dagger A_\alpha = I$.

Definición Un **canal cuántico** es un superoperador T que admite una **representación de Kraus normalizada**, es decir, como en (4.6). Una generalización de los canales cuánticos son las **operaciones cuánticas**, son superoperadores que admiten una representación de Kraus subnormalizada, es decir, $\sum_m A_m^\dagger A_m \leq I$.

La representación de Kraus (4.6) se ha obtenido a partir de una evolución unitaria en un espacio mayor. El siguiente teorema establece que siempre es así.

4.1.1.2. Apartado matemático: operadores isométricos

Sean $\mathscr{H}_1$ y $\mathscr{H}_2$ dos espacios de Hilbert (complejos) de igual dimensión, y $\mathscr{V}$ un subespacio de $\mathscr{H}_1$, y sea U un **operador isométrico** definido en $\mathscr{V}$, es decir, $\forall|\psi\rangle \in \mathscr{V} \;\; \||\psi\rangle\| = \|U|\psi\rangle\|$. Equivale a decir que conserva el producto escalar. Entonces el núcleo es trivial y U es un operador inyectivo. Por tanto $\mathscr{V}$ y $\mathscr{W} \equiv U\mathscr{V}$

[4.1] En espacios de Hilbert finitos, todos los operadores son acotados.

tienen la misma dimensión. $U : \mathscr{V} \to \mathscr{W}$ es unitario, y se puede extender (de muchas formas) a un operador unitario $\mathscr{H}_1 \to \mathscr{H}_2$. En efecto, $\mathscr{V}^\perp$ y $\mathscr{W}^\perp$ también tienen la misma dimensión, basta elegir sendas bases ortonormales $\{|a_i\rangle\}$ y $\{|b_i\rangle\}$ en esos subespacios y completar la definición de U mediante $U|a_i\rangle = |b_i\rangle$.

Teorema Un canal cuántico T en $\mathscr{H}_A$ con M operadores de Kraus se puede hacer derivar de una evolución unitaria U_{AE} en $\mathscr{H}_A \otimes \mathscr{H}_E$ con $\dim \mathscr{H}_E \geq M$, y $|\phi\rangle_E \in \mathscr{H}_E$ normalizado cualquiera, de modo que

$$T(\rho) = \mathrm{Tr}_E \left(U_{AE}\, \rho \otimes |\phi\rangle\langle\phi|_E\, U_{AE}^\dagger \right). \tag{4.8}$$

Demostración: Elegimos una base ortonormal $\{|m\rangle_E\}$ de $\mathscr{H}_E$, y definimos el operador U_{AE} mediante

$$U_{AE}|\psi\rangle_A \otimes |\phi\rangle_E = \sum_{m=1}^{M} A_m|\psi\rangle_A \otimes |m\rangle_E\,. \tag{4.9}$$

Ahora mismo U_{AE} está definido en el subespacio $\mathscr{V} = \mathscr{H}_A \otimes |\phi\rangle_E$. Para ver que se puede extender a un operador unitario basta comprobar que es isométrico (conserva el producto escalar):

$$\begin{aligned} \langle\psi'|\psi\rangle = (\langle\psi'| \otimes \langle\phi|)(|\psi\rangle \otimes |\phi\rangle) &\underset{U_{AE}}{\to} \left(\sum_{m'} \langle\psi'|A_{m'}^\dagger \otimes \langle m'| \right) \left(\sum_m A_m|\psi\rangle \otimes |m\rangle \right) \\ &= \sum_m \langle\psi'|A_m^\dagger A_m|\psi\rangle = \langle\psi'|\psi\rangle. \end{aligned} \tag{4.10}$$

□

4.1.2. Propiedades de los canales cuánticos

Observaciones:

i) El canal cuántico T está definido como una aplicación lineal. El mismo canal puede admitir representaciones con distintos operadores de Kraus y distintos valores de M.

ii) Un sistema cerrado o aislado corresponde a $M = 1$, y en este caso la evolución es unitaria, por $A_1^\dagger A_1 = I$. Los canales cuánticos generalizan la evolución unitaria.

iii) Los canales cuánticos se pueden componer y forman un semigrupo. T sólo es invertible cuando es unitario, es decir, cuando $M = 1$. [4.2]

[4.2] Más exactamente: Si T es una aplicación invertible, T^{-1} es también un canal cuántico si y sólo si T es unitario.

iv) Una combinación convexa (un promedio) de canales cuánticos es también un canal cuántico. Los canales cuánticos forman un conjunto convexo en el espacio de superoperadores.

v) $\rho \to T(\rho) = \rho'$ y en general ρ' puede ser una matriz densidad en otro espacio $\mathscr{H}'_A \neq \mathscr{H}_A$. Es decir, $T : \mathscr{B}(\mathscr{H}_A) \to \mathscr{B}(\mathscr{H}'_A)$. En este caso $A_\mu : \mathscr{H}_A \to \mathscr{H}'_A$ y $A^\dagger_\mu : \mathscr{H}'_A \to \mathscr{H}_A$.

Por ejemplo si $\mathscr{H}_A = \mathscr{H}_{A_1} \otimes \mathscr{H}_{A_2}$, $T(\mathscr{O}) = \mathrm{Tr}_{A_1}(\mathscr{O})$ es un canal cuántico de $\mathscr{H}_A$ en $\mathscr{H}_{A_2}$.

vi) Partiendo de $\rho = |\psi\rangle\langle\psi|$,

$$T(\rho) = \sum_m A_m |\psi\rangle\langle\psi| A^\dagger_m = \sum_m |\tilde{\psi}_m\rangle\langle\tilde{\psi}_m|, \qquad |\tilde{\psi}_m\rangle \equiv A_m|\psi\rangle. \tag{4.11}$$

Se puede interpretar como $|\psi\rangle \to |\tilde{\psi}_m\rangle$ con probabilidad $\||\tilde{\psi}_m\rangle\|^2$. Las medidas no selectivas son ejemplos de canales cuánticos. Las medidas selectivas son operaciones cuánticas.

Para caracterizar de forma más completa los canales cuánticos, necesitamos introducir algunas propiedades de los superoperadores.

Proposición Si un superoperador $T : \mathscr{B}(\mathscr{H}_A) \to \mathscr{B}(\mathscr{H}'_A)$ admite una representación de Kraus, $T(\mathscr{O}) = \sum_m A_m \mathscr{O} A^\dagger_m$, la representación está normalizada si y sólo si T conserva la traza:

$$\sum_m A^\dagger_m A_m = I \quad \Longleftrightarrow \quad \forall \mathscr{O} \in \mathscr{B}(\mathscr{H}_A) \quad \mathrm{Tr}(\mathscr{O}) = \mathrm{Tr}(T(\mathscr{O})) \tag{4.12}$$

Demostración: En efecto, por la identidad

$$\mathrm{Tr}_{A'}(T(\mathscr{O})) = \mathrm{Tr}_{A'}(\sum_m A_m \mathscr{O} A^\dagger_m) = \mathrm{Tr}_A(\sum_m A^\dagger_m A_m \mathscr{O}), \tag{4.13}$$

$\sum_m A^\dagger_m A_m = I \implies \mathrm{Tr}_A(\mathscr{O}) = \mathrm{Tr}_{A'}(T(\mathscr{O}))$. Y viceversa: Si $\forall \mathscr{O} \;\; \mathrm{Tr}_A(\mathscr{O}) = \mathrm{Tr}_{A'}(T(\mathscr{O}))$

$$\delta_{jk} = \mathrm{Tr}_A(|j\rangle\langle k|) = \mathrm{Tr}_{A'}(T(|j\rangle\langle k|)) = \mathrm{Tr}_A(\sum_m A^\dagger_m A_m |j\rangle\langle k| = \langle j| \sum_m A^\dagger_m A_m |k\rangle_A, \tag{4.14}$$

y por tanto $\sum_m A^\dagger_m A_m = I_A$. □

La conservación de la traza y la normalización de los operadores de Kraus son condiciones equivalentes. En particular, *los canales cuánticos conservan la traza*.

Definición Un superoperador es **positivo** cuando $\mathscr{O} \geq 0 \implies T(\mathscr{O}) \geq 0$. En este caso T también conserva hermiticidad.

Es claro que todo superoperador que admite una representación de Kraus (y en particular un canal cuántico) es positivo.

Dado un superoperador T en un espacio $\mathscr{H}_A$, se puede considerar su **extensión por la identidad** a un espacio $\mathscr{H}_A \otimes \mathscr{H}_B$, es decir, $T \otimes \hat{I}_B$. Esta extensión se define mediante

$$T \otimes \hat{I}_B(A \otimes B) \equiv T(A) \otimes B \qquad A \in \mathscr{B}(\mathscr{H}_A), \quad B \in \mathscr{B}(\mathscr{H}_B), \tag{4.15}$$

y la definición se extiende por linealidad a operadores arbitrarios de $\mathscr{H}_A \otimes \mathscr{H}_B$.

La idea es que se han añadido grados de libertad auxiliares y T no actúa sobre ellos. Nótese que si T conserva la traza, $T \otimes \hat{I}_B$ también.

Definición Un superoperador T es **completamente positivo** cuando todas sus extensiones por la identidad son positivas.

Obviamente todo superoperador completamente positivo es también positivo. Pero no al revés: La condición de ser completamente positivo es estrictamente más fuerte.

Un ejemplo de superoperador positivo pero no completamente positivo es $T(\mathscr{O}) = \mathscr{O}^T$ (trasponer en una cierta base es positivo ya que no cambia el espectro ni la hermiticidad) sin embargo si se extiende a $\mathscr{H}_A \otimes \mathscr{H}_B$, $T \otimes \hat{I}_B(\rho) = \rho^{T_A}$ es el traspuesto parcial. Ya vimos ejemplos de matrices densidad ρ tales que ρ^{T_A} no era positivo (requería que ρ fuera no separable, Tema 3.4.2).

Definición *(Dualidad canal-estado)* A todo superoperador $T : \mathscr{B}(\mathscr{H}_A) \to \mathscr{B}(\mathscr{H}_A')$ se le puede asociar un operador $\tilde{\rho} \in \mathscr{B}(\mathscr{H}_A' \otimes \mathscr{H}_B)$ de manera unívoca por la relación

$$\tilde{\rho} = \sum_{j,k=1}^{N} T(|j\rangle\langle k|_A) \otimes |j\rangle\langle k|_B, \qquad T(|j\rangle\langle k|) = \langle j|\tilde{\rho}|k\rangle_B \,. \tag{4.16}$$

Aquí $N = \dim \mathscr{H}_A = \dim \mathscr{H}_B$ y $\{|j\rangle_A\}$ y $\{|j\rangle_B\}$ son bases ortonormales cualesquiera e independientes. $\mathscr{H}_B$ es una copia de $\mathscr{H}_A$. [4.3]

$\tilde{\rho}$ es la denominada **matriz de Choi** de T. La biyección $\tilde{\rho} \leftrightarrow T$ depende de la elección de bases ortonormales.

La matriz de Choi también puede expresarse como

$$\tilde{\rho} = T \otimes \hat{I}_B(|\Xi\rangle\langle\Xi|_{AB}), \qquad |\Xi\rangle_{AB} \equiv \sum_{j=1}^{N} |j\rangle_A \otimes |j\rangle_B \,. \tag{4.17}$$

[4.3] En la literatura la construcción se hace directamente con $\mathscr{H}_A' \otimes \mathscr{H}_A$.

En efecto:

$$\begin{aligned} T\otimes\hat{I}_B(|\Xi\rangle\langle\Xi|_{AB}) &= T\otimes\hat{I}_B\left(\sum_{j=1}^{N}|j\rangle_A\otimes|j\rangle_B\sum_{k=1}^{N}\langle k|_A\otimes\langle k|_B\right) \\ &= \sum_{j,k=1}^{N}T(|j\rangle\langle k|_A)\otimes|j\rangle\langle k|_B=\tilde{\rho}\,. \quad \square \end{aligned} \tag{4.18}$$

Teorema Sea $T:\mathscr{B}(\mathscr{H}_A)\to\mathscr{B}(\mathscr{H}'_A)$ y $\tilde{\rho}$ su matriz de Choi. Las siguientes afirmaciones son equivalentes:

i) T admite una representación de Kraus.

ii) T es completamente positivo.

iii) $\tilde{\rho}$ es un operador positivo.

Demostración:

[i) $\Longrightarrow$ ii)] Si $T(\mathscr{O})=\sum_m A_m\mathscr{O}A_m^\dagger$, claramente T es positivo. Sean A y B operadores en $\mathscr{H}_A$ y $\mathscr{H}_B$,

$$T\otimes\hat{I}_B(A\otimes B)=\left(\sum_m A_m A A_m^\dagger\right)\otimes B=\sum_m A_m\otimes I_B A\otimes B A_m^\dagger\otimes I_B \tag{4.19}$$

Es decir

$$T\otimes\hat{I}_B(\mathscr{O})=\sum_m A_m\otimes I_B\,\mathscr{O}A_m^\dagger\otimes I_B\,. \tag{4.20}$$

$T\otimes\hat{I}_B$ también admite una representación de Kraus y por tanto es positivo. Luego T es completamente positivo.

[ii) $\Longrightarrow$ iii)] Para la matriz de Choi ya se han elegido $\mathscr{H}_B\approx\mathscr{H}_A$ y sendas bases ortonormales. Teniendo en cuenta (4.17), $\tilde{\rho}=T\otimes\hat{I}_B(|\Xi\rangle\langle\Xi|_{AB})$.

$T\otimes\hat{I}_B$ es positivo por ser T completamente positivo, y también $|\Xi\rangle\langle\Xi|_{AB}\geq 0$, entonces $\tilde{\rho}\geq 0$.

[iii) $\Longrightarrow$ i)] Si $\tilde{\rho}\geq 0$, admite una descomposición

$$\tilde{\rho}=\sum_\mu|\tilde{\phi}_\mu\rangle\langle\tilde{\phi}_\mu|_{A'B}\,. \tag{4.21}$$

Comparando con (4.16) esto implica

$$T(|j\rangle\langle k|) = {}_B\langle j|\tilde{\rho}|k\rangle_B = \sum_\mu {}_B\langle j|\tilde{\phi}_\mu\rangle\langle\tilde{\phi}_\mu|k\rangle_B\,. \tag{4.22}$$

Si ahora se definen los operadores de Kraus $A_\mu : \mathscr{B}(\mathscr{H}_A) \to \mathscr{B}(\mathscr{H}'_A)$ mediante

$$A_\mu|j\rangle_A \equiv {}_B\langle j|\tilde{\phi}_\mu\rangle_{A'B}\,, \tag{4.23}$$

se tiene

$$T(|j\rangle\langle k|) = \sum_\mu A_\mu|j\rangle\langle k|A_\mu^\dagger\,. \tag{4.24}$$

Por linealidad de T se deduce

$$\forall\mathscr{O}\in\mathscr{B}(\mathscr{H}_A) \qquad T(\mathscr{O}) = \sum_\mu A_\mu\mathscr{O}A_\mu^\dagger. \tag{4.25}$$

Es decir, T admite una representación de Kraus. □

Corolario Un superoperador T es un canal cuántico (es decir, admite una representación de Kraus normalizada) si y sólo si T es completamente positivo y conserva la traza.

El Corolario equivale a decir que un canal cuántico es cualquier aplicación lineal que transforma matrices densidad en matrices densidad, también cuando se extiende trivialmente por la identidad.

Hemos visto tres formas de representar un canal cuántico T:
i) mediante operadores de Kraus (eq. (4.6)),
ii) mediante un operador unitario en un espacio extendido (eq. (4.8)).
iii) mediante la matriz de Choi $\tilde{\rho}$, un operador en $\mathscr{H}'_A\otimes\mathscr{H}_B$, con la construcción en (4.16). Aquí, $\mathscr{H}_B$ es una copia de $\mathscr{H}_A$.

Como se ha visto, T es complemente positivo si y sólo si $\tilde{\rho}\geq 0$. Por otro lado, la condición necesaria y suficiente para que $\tilde{\rho}$ corresponda a un superoperador T que conserve la traza es

$$\mathrm{Tr}_{A'}(\tilde{\rho}) = I_B\,. \tag{4.26}$$

Demostración: Sea $\mathscr{O} = \sum_{jk}\mathscr{O}_{jk}|j\rangle\langle k|_A$. Entonces

$$\mathrm{Tr}_{A'}(T(\mathscr{O})) = \mathrm{Tr}_{A'}(\sum_{jk}\mathscr{O}_{jk}T(|j\rangle\langle k|)) = \mathrm{Tr}_{A'}(\sum_{jk}\mathscr{O}_{jk\,B}\langle j|\tilde{\rho}|k\rangle_B) = \sum_{jk}\mathscr{O}_{jk\,B}\langle j|\,\mathrm{Tr}_{A'}(\tilde{\rho})|k\rangle_B. \tag{4.27}$$

Dado que $\mathrm{Tr}_A(\mathscr{O}) = \sum_i\mathscr{O}_{ii}$, a partir de esta expresión ya es inmediato que
$\forall\mathscr{O}\in\mathscr{B}(\mathscr{H}_A)\;\; \mathrm{Tr}_A(\mathscr{O}) = \mathrm{Tr}_{A'}(T(\mathscr{O})) \iff {}_B\langle j|\,\mathrm{Tr}_{A'}(\tilde{\rho})|k\rangle_B = \delta_{jk} \iff \mathrm{Tr}_{A'}(\tilde{\rho}) = I_B.$ □

En el caso de un canal cuántico, por ser un operador hermítico en $\mathscr{B}(\mathscr{H}_A' \otimes \mathscr{H}_A)$, $\tilde{\rho}$ tiene a priori $(NN')^2$ parámetros reales libres (siendo $N = \dim \mathscr{H}_A$ y $N' = \dim \mathscr{H}_A'$), pero la condición (4.26) fija N^2 parámetros. En total el número de parámetros reales libres de un canal cuántico es $N^2(N'^2 - 1)$. Así, un canal cuántico de un qubit tiene 12 parámetros, frente a los 3 del caso unitario.

4.1.3. Propiedades de la representación de Kraus

Todo canal cuántico admite una representación de Kraus

$$T(\mathscr{O}) = \sum_\mu A_\mu \mathscr{O} A_\mu^\dagger, \qquad \sum_\mu A_\mu^\dagger A_\mu = I, \qquad \mathscr{O} \in \mathscr{B}(\mathscr{H}) \qquad T(\mathscr{O}) \in \mathscr{B}(\mathscr{H}') \tag{4.28}$$

Pero ésta no es única. Como se ha visto en la demostración del teorema, los operadores de Kraus se pueden asociar descomposiciones de $\tilde{\rho}$ como una mezcla. El número M de operadores de Kraus en la representación de un mismo canal cuántico puede ser arbitrariamente grande. Sin embargo diagonalizar $\tilde{\rho}$ requiere a lo sumo $M \leq NN'$ siendo $N = \dim \mathscr{H}_A$ y $N' = \dim \mathscr{H}_A'$. Por tanto, la representación de Kraus se puede hacer con no más de NN' operadores A_μ, por muy grande que sea la dimensión del espacio ambiente. Un canal cuántico concreto se puede representar con un número $M = \mathrm{rank}(\tilde{\rho})$ de operadores de Kraus, pero no menos.

Todo esto es análogo a lo que ocurre con matrices densidad. Basándose en la identificación $T \leftrightarrow \tilde{\rho}$ se tiene también:

Teorema Para operadores A_μ y B_ν de $\mathscr{B}(\mathscr{H}_A) \to \mathscr{B}(\mathscr{H}_A')$, la igualdad

$$\forall \mathscr{O} \in \mathscr{B}(\mathscr{H}_A) \qquad \sum_\mu A_\mu \mathscr{O} A_\mu^\dagger = \sum_\nu B_\nu \mathscr{O} B_\nu^\dagger \tag{4.29}$$

se satisface si y sólo si hay una matriz unitaria $U_{\mu\nu}$ tal que

$$A_\mu = \sum_\nu U_{\mu\nu} B_\nu \qquad U_{\mu\nu} \text{ unitaria.} \tag{4.30}$$

Nótese que se supone una representación de Kraus normalizada, pero la igualdad automáticamente garantiza $\sum_\mu A_\mu^\dagger A_\mu = \sum_\nu B_\nu^\dagger B_\nu$.

Este teorema nos indica la relación más general entre distintas representaciones de Kraus de un canal cuántico, o más generalmente, de una operación cuántica.

Demostración: Sea $T(\mathscr{O}) \equiv \sum_\mu A_\mu \mathscr{O} A_\mu^\dagger$ y $\tilde{\rho}$ su matriz de Choi:

$$\tilde{\rho} = \sum_{j,k} T(|j\rangle\langle k|_A) \otimes |j\rangle\langle k|_B \,. \tag{4.31}$$

Por una lado

$$\tilde{\rho} = \sum_{j,k}\sum_\mu A_\mu |j\rangle\langle k|_A A_\mu^\dagger \otimes |j\rangle\langle k|_B = \sum_\mu |\tilde{\psi}_\mu\rangle\langle\tilde{\psi}_\mu|, \qquad |\tilde{\psi}_\mu\rangle \equiv \sum_j A_\mu |j\rangle_A \otimes |j\rangle_B \,, \tag{4.32}$$

y por otro

$$\tilde{\rho} = \sum_\nu |\tilde{\phi}_\nu\rangle\langle\tilde{\phi}_\nu|, \qquad |\tilde{\phi}_\nu\rangle \equiv \sum_j B_\nu |j\rangle_A \otimes |j\rangle_B \,. \tag{4.33}$$

Por el primer teorema del Tema 2.6 se sigue que existe una matriz unitaria $U_{\mu\nu}$ tal que

$$|\tilde{\psi}_\mu\rangle = \sum_\nu U_{\mu\nu} |\tilde{\phi}_\nu\rangle \tag{4.34}$$

Proyectando sobre los ${}_B\langle j|$ se deduce [4.4]

$$\forall j \quad A_\mu |j\rangle_A = \sum_\nu U_{\mu\nu} B_\nu |j\rangle_A \implies A_\mu = \sum_\nu U_{\mu\nu} B_\nu \,. \tag{4.35}$$

□

4.2. Ejemplos de canales cuánticos

4.2.1. Despolarización

Como ejemplo consideremos un qubit $|\psi\rangle$ en un espacio ambiente de dimensión al menos 4, con una evolución isótropa del tipo

$$U|\psi\rangle \otimes |0\rangle = \sqrt{1-p}\,|\psi\rangle \otimes |0\rangle + \sqrt{\frac{p}{3}} \sum_{i=1}^{3} \sigma_i |\psi\rangle \otimes |i\rangle, \qquad 0 \le p \le 1 \,. \tag{4.36}$$

Usando $\sigma_i^2 = I$ es fácil comprobar que conserva el producto escalar.

[4.4] Nótese que partir de $\sum_\mu A_\mu |\psi\rangle\langle\psi| A_\mu^\dagger = \sum_\nu B_\nu |\psi\rangle\langle\psi| B_\nu^\dagger$ ya implica $A_\mu|\psi\rangle = \sum_\nu U_{\mu\nu} B_\nu |\psi\rangle$, pero faltaría probar que $U_{\mu\nu}$ se puede además elegir común para todos los $|\psi\rangle$.

Su representación de Kraus se obtiene con

$$\rho' = T(\rho) = \mathrm{Tr}_E(U\rho|0\rangle\langle 0|U^\dagger) = (1-p)\rho + \frac{p}{3}\sum_{i=1}^{3}\sigma_i\rho\sigma_i \tag{4.37}$$

corresponde a operadores de Kraus A_μ, $\mu = 0,1,2,3$,

$$A_0 = \sqrt{1-p}I \qquad A_i = \sqrt{\frac{p}{3}}\sigma_i, \quad i = 1,2,3 \tag{4.38}$$

Si $\rho = \frac{1}{2}(I + \boldsymbol{u}\boldsymbol{\sigma})$, $\boldsymbol{u}$ es el vector en la esfera de Bloch ($\boldsymbol{u} \in \mathbb{R}^3$, $\|\boldsymbol{u}\| \leq 1$) un sencillo cálculo [4.5] produce

$$\rho' = \frac{1}{2}(I + \boldsymbol{u}'\boldsymbol{\sigma}), \qquad \boldsymbol{u}' = (1 - \frac{4}{3}p)\boldsymbol{u}\,. \tag{4.39}$$

El canal respeta isotropía, pero reduce la polarización, ya que $\|\boldsymbol{u}'\| \leq \|\boldsymbol{u}\|$. En el caso extremo $p = \dfrac{3}{4}$ el canal T reparte el estado en las cuatro direcciones del espacio ambiente por igual.

4.2.2. Pérdida de coherencia cuántica

La coherencia cuántica se refiere la información contenida en las fases relativas en una superposición de estados. Está contenida en los elementos de matriz no diagonales de la matriz densidad. Veamos un ejemplo esquemático de como la coherencia cuántica puede desaparecer por acoplamiento del sistema con el ambiente. El modelo se define por un operador unitario tal que

$$\begin{aligned} |0\rangle_A \otimes |0\rangle_E &\to \sqrt{1-p}|0\rangle_A \otimes |0\rangle_E + \sqrt{p}|0\rangle_A \otimes |1\rangle_E, \\ |1\rangle_A \otimes |0\rangle_E &\to \sqrt{1-p}|1\rangle_A \otimes |0\rangle_E + \sqrt{p}|1\rangle_A \otimes |2\rangle_E, \end{aligned} \tag{4.40}$$

siendo $0 \leq p \leq 1$. Equivalentemente

$$i = 0,1 \quad U|i\rangle_A \otimes |0\rangle_E = |i\rangle_A \otimes |\phi_i\rangle_E, \qquad \langle\phi_0|\phi_1\rangle = 1 - p. \tag{4.41}$$

Nótese que aquí el estado del sistema, sea $|0\rangle_A$ o $|1\rangle_A$, no cambia pero con probabilidad p el ambiente sí cambia de estado. Sin embargo la propiedad $|i\rangle_A \to |i\rangle_A$ sólo es valida en una base privilegiada, para otros estados $|\psi\rangle_A$ ya no se cumple, de hecho se obtiene un estado entrelazado en general.

[4.5] Usando la identidad $\sum_i \sigma_i\sigma_k\sigma_i = -\sigma_k$.

U define un canal cuántico $\rho' = T(\rho)$. Mediante cálculo directo se obtiene

$$\rho' = \mathrm{Tr}_E(U\rho\otimes|0\rangle\langle 0|_E U^\dagger) = \mathrm{Tr}_E\Big(\sum_{ij}\rho_{ij}|i\rangle\langle j|\otimes|\phi_i\rangle\langle\phi_j|\Big) = \sum_{ij}\rho_{ij}|i\rangle\langle j|\langle\phi_j|\phi_i\rangle \tag{4.42}$$

es decir, $\langle i|\rho'|j\rangle = \langle\phi_j|\phi_i\rangle\langle i|\rho|j\rangle$,

$$\begin{pmatrix}\rho'_{00} & \rho'_{01}\\ \rho'_{10} & \rho'_{11}\end{pmatrix} = \begin{pmatrix}\rho_{00} & (1-p)\rho_{01}\\ (1-p)\rho_{10} & \rho_{11}\end{pmatrix} \tag{4.43}$$

o en términos de operadores de Kraus

$$\rho' = (1-p)\rho + p\sum_{i=0,1}E_i\rho E_i, \qquad E_i = |i\rangle\langle i| \tag{4.44}$$

es decir, $A_2 = \sqrt{1-p}\,I, \quad A_i = \sqrt{p}\,E_i, \;\; i = 0,1.$

El efecto de $\rho \to \sum_i E_i\rho E_i$ (que ocurre con probabilidad p) es una pérdida de coherencia

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \to |\alpha|^2|0\rangle\langle 0| + |\beta|^2|1\rangle\langle 1|. \tag{4.45}$$

Se pierden las fases relativas en la base privilegiada.

Si el proceso ocurre repetidamente con probabilidad $p = \Gamma dt$ en cada dt, $1-p = 1-\Gamma dt = e^{-\Gamma dt}$ y el efecto acumulado es $1-p = e^{-\Gamma t}$,

$$\begin{pmatrix}|\alpha|^2 & \alpha^*\beta\\ \beta^*\alpha & |\beta|^2\end{pmatrix} \longrightarrow \begin{pmatrix}|\alpha|^2 & e^{-\Gamma t}\alpha^*\beta\\ e^{-\Gamma t}\beta^*\alpha & |\beta|^2\end{pmatrix} \underset{t\gg 1/\Gamma}{\longrightarrow} \begin{pmatrix}|\alpha|^2 & 0\\ 0 & |\beta|^2\end{pmatrix} \tag{4.46}$$

Aunque la unitaridad y la coherencia cuántica se mantiene en el sistema completo $\mathscr{H}_A\otimes\mathscr{H}_E$, desde el punto de vista del sistema, la coherencia, la propiedad genuinamente cuántica, se pierde para tiempos $t \gg \tau = \frac{1}{\Gamma}$. Para un sistema macroscópico esto ocurre con mucha rapidez (el famoso gato de Schrödinger).

Un ejemplo idealizado de aplicación de este canal es un rayo cósmico energético, digamos un protón, que es el sistema A, que se desplaza en presencia del fondo cósmico de radiación, fotones muy blandos, que forman el ambiente. La interacción ocurre repetidamente, el protón prácticamente no se ve afectado por los fotones (no pierde energía), pero sí coherencia cuántica. Si por ejemplo el protón se creó en una superposición coherente de dos estados de momento bien definido con distintas direcciones, eventualmente se tendrá un estado mezcla, es decir, con cierta probabilidad acabará en un estado de momento o en el otro.

Lo mismo ocurre cuando se mide un observable de un sistema cuántico. El aparato de medida (el ambiente) es macroscópico y rápidamente toda la coherencia cuántica del sistema medido se evapora al ambiente para no volver.[4.6] El sistema cuántico queda en un estado mezcla, justamente en la base del observable que se quiere medir (ya que el dispositivo experimental se ha diseñado para seleccionar precisamente esa base).

Es importante notar que la pérdida de coherencia ocurre en una base privilegiada (no es una propiedad covariante bajo cambios de base). Bases privilegiadas son típicamente las asociadas a la posición o al momento. Son los observables que aparecen en un hamiltoniano típico $H = \frac{\boldsymbol{p}^2}{2m} + V(\boldsymbol{x})$. La posición porque la interacción siempre es local, la posición se conserva inmediatamente después de una interacción o las partículas se crean y destruyen localmente, no a distancia. La de momento porque una vez creada una partícula ésta se propaga conservando su momento mientras no interaccione.

4.2.3. Ecuación de Lindblad

Para un sistema cuántico cerrado (es decir, aislado) la evolución es unitaria: [4.7]

$$\rho(t) = e^{-i(t-t_0)H}\rho(t_0)e^{i(t-t_0)H}\,. \tag{4.47}$$

Para un tiempo infinitesimal

$$\rho(t+dt) = (1-iHdt)\rho(t)(1+iHdt) = \rho(t) - i[H,\rho(t)]dt. \tag{4.48}$$

Es decir,

$$\frac{d\rho(t)}{dt} = \mathscr{L}_H(\rho(t)), \qquad \mathscr{L}_H(\mathscr{O}) := -i[H,\mathscr{O}]\,. \tag{4.49}$$

$\mathscr{L}_H$ es un superoperador que controla la ecuación maestra que describe la evolución de $\rho(t)$ en el caso unitario. La solución de la ecuación puede expresarse como

$$\rho(t) = e^{(t-t_0)\mathscr{L}_H}\rho(t_0) \tag{4.50}$$

que es equivalente a eq. (4.47).

[4.6] El sistema completo evoluciona unitariamente, la información de las fases relativas no desaparece, pero en la práctica el proceso es irreversible. Es análogo al funcionamiento del segundo principio de la termodinámica, que actúa aunque la evolución hamiltoniana subyacente sea reversible.

[4.7] Por simplicidad suponemos un sistema conservativo (esto es, H independiente de t). La extensión caso no conservativo es inmediata. Aquí usamos unidades $\hbar = 1$.

Se quiere obtener una ecuación maestra análoga para el caso de un sistema abierto,

$$\frac{d\rho(t)}{dt} = \mathscr{L}(\rho(t)), \tag{4.51}$$

de modo que $T_{dt} := \hat{I} + dt\mathscr{L}$ sea un *canal cuántico infinitesimal*, y la evolución a tiempos finitos sea

$$\rho(t) = e^{(t-t_0)\mathscr{L}}\rho(t_0). \tag{4.52}$$

Toda la evolución ocurre en un mismo espacio $\mathscr{H}_A$. En el apéndice 4.2.4 se discute cuándo puede aplicarse este enfoque.

El canal infinitesimal T_{dt} describe la evolución del sistema abierto de t a $t+dt$. El canal cuántico actúa según

$$\rho(t+dt) = T_{dt}(\rho(t)) = A_0\rho(t)A_0^\dagger + \sum_{\mu>0} A_\mu\rho(t)A_\mu^\dagger. \tag{4.53}$$

Debe cumplir que $T_0(\rho) = \rho$. Esto requiere [4.8]

$$A_0 \underset{dt\to 0}{\to} I, \qquad A_\mu \underset{dt\to 0}{\to} 0 \quad \mu > 0. \tag{4.54}$$

Además el comportamiento cuando $dt \to 0$ debe ser tal que $\rho(t)$ sea derivable. Eso va a requerir

$$A_\mu = \sqrt{dt}L_\mu + O(dt) \quad \mu > 0. \tag{4.55}$$

Por otro lado A_0 debe ser de la forma

$$A_0 = I - (K + iH)dt + O(dt^2), \qquad H, K \text{ hermíticos.} \tag{4.56}$$

Posibles términos con $\sqrt{dt}$ en A_0 deben anularse ya que su contribución produciría términos $\sqrt{dt}$ en $T_{dt}(\rho)$ y $\rho(t)$ no sería derivable.

La condición de normalización requiere

$$I - \sum_{\mu>0} A_\mu^\dagger A_\mu = A_0^\dagger A_0 = I - 2Kdt + O(dt^2) \tag{4.57}$$

que implica

$$K = \frac{1}{2}\sum_{\mu>0} L_\mu^\dagger L_\mu. \tag{4.58}$$

[4.8] Esta implicación se justifica en el apéndice 4.2.4.

Teniendo en cuenta que

$$A_0 \rho A_0^\dagger = \rho - \{K,\rho\}dt - i[H,\rho]dt + O(dt^2), \tag{4.59}$$

se obtiene la **ecuación de Lindblad**

$$\frac{d\rho}{dt} = -i[H,\rho] + \sum_\mu \left(L_\mu \rho L_\mu^\dagger - \frac{1}{2}\{L_\mu^\dagger L_\mu, \rho\} \right) \equiv \mathscr{L}(\rho(t)) \,. \tag{4.60}$$

$\mathscr{L}$ es el superoperador de Lindblad, y la solución de la ecuación es $\rho(t) = e^{(t-t_0)\mathscr{L}}\rho(t_0)$. Por construcción la traza de ρ se conserva, $\mathrm{Tr}(\mathscr{L}(\rho)) = 0$.

La ecuación de Lindblad es la ecuación maestra más general que es lineal en ρ, conserva la traza y mantiene positividad de ρ en la evolución.

La ecuación tiene una parte de evolución unitaria y otra de interacción con el ambiente, que rompe unitaridad. Los operadores L_μ representan interacciones de A con el entorno, que se ven como saltos cuánticos no hamiltonianos en A. [4.9]

4.2.4. Apéndice: Aplicación de la ecuación de Lindblad

Justificación de ec. (4.54)

Si T_{dt} tiene operadores de Kraus $B_\nu(dt)$, la condición $T_0 = \hat{I}$ equivale a

$$\sum_{\nu=0}^{M} B_\nu(0) \mathscr{O} B_\nu^\dagger(0) = \mathscr{O}. \tag{4.61}$$

Como el mismo canal cuántico $\hat{I}$ se puede representar con operadores de Kraus $A_0 = I$ y $A_\mu = 0$ para $\mu = 1,\ldots,M$, se deduce que los $B_\nu(0)$ son una rotación unitaria de los A_μ. Aplicando la rotación inversa a $B_\nu(dt)$, se obtiene una nueva representación de Kraus de T_{dt} con operadores $A_\mu(dt)$ que ahora satisfacen (4.54).

Aplicabilidad de la ecuación de Lindblad

Denotemos como T_{t_2,t_1} una familia de canales cuánticos que realicen la evolución del sistema abierto entre dos tiempos cualesquiera $t_2 \geq t_1$,

$$\rho_{t_2} = T_{t_2,t_1}(\rho_{t_1}). \tag{4.62}$$

[4.9] De hecho, a menudo la ecuación se resuelve mediante un proceso estocástico: $\rho(t)$ se hace evolucionar unitariamente con H, pero con cierta probabilidad por unidad de tiempo de saltar a otro estado, controlado por los operadores L_μ.

Debe cumplirse entonces que

$$T_{t_3,t_1} = T_{t_3,t_2} \circ T_{t_2,t_1} \qquad t_3 \geq t_2 \geq t_1. \tag{4.63}$$

Esta condición de tipo markoviano se cumple para la evolución unitaria y para la ecuación de Lindblad. Pero no para canales cuánticos generales.

Los canales cuánticos proceden de una ecuación del tipo

$$T(\rho) = \mathrm{Tr}_E(U_{AE}\,\rho \otimes \rho_E\, U_{AE}^{\dagger}). \tag{4.64}$$

Para que la aplicación $\rho \to T(\rho)$ esté bien definida, el estado inicial AE debe ser factorizable. Se supone que el entorno E es fijo mientras ρ varía, para definir la aplicación T. En un estado general ρ_{AE}, incluso aunque $\rho_E = \mathrm{Tr}_A(\rho_{AE})$ se mantenga fijo, la información contenida en $\rho = \mathrm{Tr}_E(\rho_{AE})$ no va a determinar unívocamente la evolución posterior en general. Entonces, para poder aplicar canales cuánticos sucesivos, como en (4.63), tiene que ocurrir que el estado final $U_{AE}\,\rho \otimes \rho_E\, U_{AE}^{\dagger}$ sea de nuevo de tipo separable, $\rho' \otimes \rho'_E$, al menos de manera aproximada.

En el ejemplo del rayo cósmico energético propagándose por el fondo de cósmico de radiación discutido anteriormente (Tema 4.2.2) la partícula se encuentra continuamente con un estado ρ_E similar (a saber, nuevos fotones en el estado térmico del fondo cósmico) y la ecuación maestra va a ser aplicable. Otro ejemplo es un sistema A que interacciona con un átomo del medio. El átomo está en el estado fundamental en ausencia de interacción. Eventualmente la interacción con A puede producir un estado excitado que se mantiene durante un tiempo de latencia τ, luego se desexcita y queda otra vez en el estado fundamental. Para valores de τ mucho menores que los tiempos típicos de evolución (u observación) la hipótesis markoviana puede ser correcta.

4.2.5. Imagen de Heisenberg

Si $T_t(\rho)$ es una canal cuántico, de modo que $\rho(t) = T_t(\rho(t_0))$ eso describe la evolución del sistema en imagen de Schrödinger. Se puede también describir en imagen de Heisenberg (referida a un t_0 dado) teniendo en cuenta la relación entre las dos imágenes de evolución:

$$\langle \mathscr{O}(t_0)\rangle_{\rho(t)} = \langle \mathscr{O}(t)\rangle_{\rho(t_0)}, \tag{4.65}$$

es decir

$$\mathrm{Tr}(\rho(t)\mathscr{O}(t_0)) = \mathrm{Tr}(\rho(t_0)\mathscr{O}(t)). \tag{4.66}$$

Dado que

$$\mathrm{Tr}(\rho(t)\mathscr{O}(t_0)) = \mathrm{Tr}(\sum_{\mu} A_\mu(t)\rho(t_0)A_\mu^{\dagger}(t)\mathscr{O}(t_0)) = \mathrm{Tr}(\rho(t_0)\sum_{\mu} A_\mu^{\dagger}(t)\mathscr{O}(t_0)A_\mu(t)) \tag{4.67}$$

se deduce [4.10]

$$\mathscr{O}(t) = T_t^\dagger(\mathscr{O}(t_0)) \equiv \sum_\mu A_\mu^\dagger(t)\mathscr{O}(t_0)A_\mu(t). \tag{4.68}$$

Para el caso markoviano

$$\frac{d\mathscr{O}(t)}{dt} = i[H,\mathscr{O}] + \sum_\mu \left(L_\mu^\dagger \mathscr{O} L_\mu - \frac{1}{2}\{L_\mu L_\mu^\dagger, \mathscr{O}\} \right) \equiv \mathscr{L}^\dagger(\mathscr{O}(t)). \tag{4.69}$$

La propiedad dual de conservar la traza es ahora

$$T^\dagger(I) = I, \qquad \mathscr{L}^\dagger(I) = 0\,. \tag{4.70}$$

Un superoperador que transforma I en I es **unital**.

4.3. Canales cuánticos prohibidos

4.3.1. Teorema de no clonación

Algunas operaciones cuánticas hipotéticas no son viables, entre ellas la **clonación de estados cuánticos**. Clonación quiere decir recibir como input un estado $|\psi\rangle$ arbitrario desconocido (conocemos el espacio al que pertenece pero no qué estado concreto es) y producir una copia exacta de ese estado, es decir, con la misma función de onda. Se distingue de la teleportación en que se mantiene el original, de modo que al final se tendría el original y su copia exacta.

Clonar estados ortogonales conocidos es posible, basta incluirlos en una base ortonormal y medir en esa base. Una vez conocido el estado es posible crear un número arbitrario de copias de él. Sin embargo hay limitaciones para clonar estados desconocidos o no ortogonales.

El **teorema de no clonación** afirma que no es posible clonar perfectamente estados cuánticos arbitrarios. Lo que no se puede clonar es un *estado de superposición* desconocido (aunque sea una superposición de estados conocidos).

Demostración: Supongamos que sí fuera posible. El esquema más general sería un operador unitario U que actúe del siguiente modo:

$$U|\psi\rangle_A \otimes |b\rangle_B \otimes |0\rangle_C = |\psi\rangle_A \otimes |\psi\rangle_B \otimes |Q_\psi\rangle_C \tag{4.71}$$

[4.10] $T^\dagger$ es el adjunto del superoperador T, para el producto escalar entre operadores $\langle A|B\rangle = \mathrm{Tr}(A^\dagger B)$.

Copia el registro A en el B, los espacios $\mathscr{H}_A$ y $\mathscr{H}_B$ son isomorfos de dimensión d. C es un registro auxiliar. El output debe ser separable para distinguir claramente cada registro (A y B se van procesar por separado, por ejemplo). Los estados $|\psi\rangle$, $|b\rangle$, $|Q_\psi\rangle$, están normalizados. $|b\rangle$ es el estado inicial, en blanco, de B. Se permite que el estado final del registro C, $|Q_\psi\rangle$, pueda depender de $|\psi\rangle$.

Por supuesto hay que suponer que $d > 1$ ya que en otro caso $|\psi\rangle$ sería conocido (cuando $d = 1$ sólo hay un estado).

Ya se ve que en la definición de U hay un problema de linealidad respecto de $|\psi\rangle$. Para hacerlo patente, sean $\{|i\rangle_A\}$ y $\{|i\rangle_B\}$ bases ortonormales, de modo que

$$U|i\rangle_A \otimes |b\rangle_B \otimes |0\rangle_C = |i\rangle_A \otimes |i\rangle_B \otimes |Q_i\rangle_C \tag{4.72}$$

Sea un estado arbitrario $|\psi\rangle = \sum_{i=1}^{d} \alpha_i |i\rangle$. Por un lado

$$|\psi\rangle \otimes |b\rangle \otimes |0\rangle \to |\psi\rangle \otimes |\psi\rangle \otimes |Q_\psi\rangle = \sum_{i,j=1}^{d} \alpha_i \alpha_j |i\rangle \otimes |j\rangle \otimes |Q_\psi\rangle \equiv |\chi\rangle \tag{4.73}$$

Por otro, por linealidad de U,

$$|\psi\rangle \otimes |b\rangle \otimes |0\rangle = \sum_{i=1}^{d} \alpha_i |i\rangle \otimes |b\rangle \otimes |0\rangle \longrightarrow \sum_{i=1}^{d} \alpha_i |i\rangle \otimes |i\rangle \otimes |Q_i\rangle = |\chi\rangle \tag{4.74}$$

Usando ambas expresiones, la proyección de $|\chi\rangle$ sobre el estado $|i,j\rangle_{AB}$ puede escribirse de dos formas

$$_{AB}\langle i,j|\chi\rangle = \alpha_i \alpha_j |Q_\psi\rangle = \delta_{ij} \alpha_i |Q_i\rangle \tag{4.75}$$

Como $d \geq 2$, podemos elegir $\alpha_i \alpha_j \neq 0$ para $i \neq j$. Eso requiere $|Q_\psi\rangle = 0$ y por tanto $|\chi\rangle = 0$ lo cual está en contradicción con que U sea unitario. □

El teorema no prohíbe clonación aproximada (es decir, producir una copia con fidelidad menor que 1) o clonar sólo con cierta probabilidad, y tampoco clonar estados con cierta información adicional. Por ejemplo, si se sabe que $|\psi\rangle$ es un estado de una base ortonormal conocida $\{|i\rangle\}$ (pero no cuál en concreto) se puede utilizar el operador

$$U|i\rangle_A \otimes |0\rangle_B = |i\rangle_A \otimes |i\rangle_B, \qquad i = 0, \ldots, d-1 \tag{4.76}$$

que es isométrico y en consecuencia se puede extender a unitario. La demostración anterior no se aplica porque entre los estados a clonar aquí no se incluye el caso $\alpha_i \alpha_j \neq 0$ para $i \neq j$. Otra forma

de ver que el conjunto $\{|i\rangle\}$ es clonable, es simplemente hacer una medida selectiva en esa base. Eso proporciona el valor de i sin cambiar el estado y una vez conocido no hay dificultad en producir copias del estado $|i\rangle$. (Este es el mismo mecanismo de una clonación clásica.)

Para matrices densidad también hay resultados. [4.11] Sea una colección de matrices densidad conocidas $\{\rho_i\}$. Una de ellas ρ_i (no sabemos cuál) es el input de nuestro canal cuántico, y queremos que el canal produzca ρ_{AB} tal que $\rho_A = \rho_B = \rho_i$. El teorema afirma que la condición necesaria y suficiente para que ese canal exista es

$$\forall i, j \;\; [\rho_i, \rho_j] = 0\,. \tag{4.77}$$

Para clonación estricta, $\rho_{AB} = \rho_A \otimes \rho_B$, las condiciones son más restrictivas, del tipo $\rho_i \rho_j = 0$ o $\rho_i = \rho_j$.

Este teorema implica que no es posible crear copias exactas de un ρ dado desconocido arbitrario (lo que se denomina "retransmitir" el estado).

Nótese que una clonación aproximada no está prohibida. Por ejemplo el mecanismo en (4.76) produce copias aproximadas que son más fieles cuanto más próximo sea $|\psi\rangle$ a alguno de los estados de la base, y empeoran drásticamente al alejarse de ese caso. Existen otros mecanismos que producen copias aproximadas de calidad más uniforme.

4.3.2. Comunicación supralumínica

Si la clonación perfecta fuera posible también sería posible la comunicación instantánea entre puntos separados espacialmente. En efecto, supongamos que Andrea y Benito comparten un ebit en estado singlete $|\Phi_-\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A|1\rangle_B - |1\rangle_A|0\rangle_B)$. Andrea puede elegir medir su qubit en la base $\{|0\rangle, |1\rangle\}$ o en la base $\{|+\rangle, |-\rangle\}$. "Después de la medida" [4.12] Benito podría clonar su qubit repetidamente hasta tener muchas copias. A continuación puede medir los estados de esas copias en la base $\{|0\rangle, |1\rangle\}$. Si siempre obtiene el mismo valor se deduce que Andrea midió en la base $\{|0\rangle, |1\rangle\}$, si el valor se reparte entre 0 y 1 al 50 % se deduce que Andrea midió en la base $\{|+\rangle, |-\rangle\}$. De esa forma se transmitiría un bit de información a velocidad supralumínica.

Puesto que la comunicación supralumínica es imposible (sería paradójica, podría llegar la res-

[4.11] H. Barnum, C.M. Caves, C.A. Fuchs, R. Jozsa y B. Schumacher, *Noncommuting mixed states cannot be broadcast*, Phys. Rev. Lett. 76(1996)2818-2821. [6]

[4.12] En realidad no está definido qué ocurre antes y qué después, siendo la separación tipo espacio. Después de la medida el estado es separable, pero eso no es esencial para que B aplique un operador U_B a su qubit ya que por linealidad siempre se puede usar una base separable, aunque el estado esté entrelazado.

puesta antes de enviar la pregunta) debe deducirse que la clonación cuántica es imposible (y en efecto lo es), pero además los métodos de clonación aproximados deben tener limitaciones. Si fueran demasiado eficientes se podrían adaptar, quizá con bases $\{|0\rangle, |1\rangle\}$ y $\{|0\rangle', |1\rangle'\}$ más parecidas, para que la clonación sea más eficiente, para enviar información instantánea con pequeña probabilidad de error. Sin embargo, si se eligen bases parecidas para mejorar la clonación se empeora la discriminación al hacer las medidas. Deber ser así para que no se pueda enviar una señal por encima del ruido sin involucrar un canal clásico.

5. Medidas cuánticas

5.1. Medidas proyectivas o estándar

5.1.1. Modelo de von Neumann

Veamos un modelo simplificado de medida debido a von Neumann.

S denota el sistema cuántico en el que se quiere hacer la medida, con espacio de Hilbert $\mathscr{H}_S$, y el observable que se quiere medir es A, un operador hermítico en $\mathscr{H}_S$. La evolución del sistema la describe el hamiltoniano H_S. Para simplificar suponemos que A es una constante de movimiento, $[A,H_S]=0$, pero con espectro posiblemente degenerado:

$$A|j\alpha\rangle = a_j|j\alpha\rangle, \qquad H_S|j\alpha\rangle = \varepsilon_{j\alpha}|j\alpha\rangle, \qquad \alpha = 1,\ldots,d_j\,, \tag{5.1}$$

y los a_j son todos distintos.

El aparato de medida o puntero p se introduce en el modelo como una partícula no relativista de masa M y espacio de Hilbert $\mathscr{H}_p$ (suponemos que se mueve en una dimensión para simplificar). La hipótesis es que esta partícula se comporta de manera suficientemente clásica de modo que su posición X es fácilmente observable. El hamiltoniano correspondiente es $H_p = P^2/(2M)$ siendo P el operador momento, es decir, $[X,P] = i\hbar I_S$.

La medida implica una interacción entre el sistema y el puntero. Se modela con un hamiltoniano de interacción del tipo $H_I = gPA$, siendo g una constante de acoplamiento. El sistema completo está en $\mathscr{H} = \mathscr{H}_S \otimes \mathscr{H}_p$ y el hamiltoniano completo en ese espacio es

$$H = \frac{P^2}{2M} + H_S + gPA \tag{5.2}$$

La idea es que el valor de X (posición del puntero) dependa del valor de A, que es lo que se quiere medir, de modo que observar X es medir A. El acoplamiento PA tiene la forma adecuada porque P es precisamente el generador de las traslaciones de X y gPA va a producir una traslación (por unidad de tiempo) proporcional a A. Otra ventaja adicional es que P conmuta con $P^2/(2M)$ lo que simplifica la discusión. Una traslación por unidad de tiempo es un boost, y se puede poner de manifiesto reescribiendo el hamiltoniano como

$$H = \frac{1}{2M}(P+gMA)^2 + H_S - \frac{M}{2}g^2A^2 \tag{5.3}$$

que corresponde a hacer un boost de velocidad gA en el puntero.

Se supone que p y S se acoplan en $t \geq 0$, y su estado en $t = 0$ es separable

$$|\Psi(0)\rangle_{pS} = |\psi\rangle_S \otimes |\phi\rangle_p \tag{5.4}$$

con $|\psi\rangle_S = \sum_{j,\alpha} \psi_{j\alpha} |j\alpha\rangle$.

El operador de evolución del sistema completo es

$$U(t) = e^{-iHt/\hbar} = U_p(t) U_S(t) U_I(t) \tag{5.5}$$

U factoriza dado que los tres términos del hamiltoniano H conmutan.

La evolución de $|\psi\rangle_S$ bajo $U_S(t)$ produce

$$|\psi\rangle_S \to \sum_{j,\alpha} e^{-i\varepsilon_{j\alpha}t/\hbar} \psi_{j\alpha} |j\alpha\rangle \tag{5.6}$$

Equivale a una redefinición de las fases de las amplitudes, $\psi_{j\alpha}(t) \equiv e^{-i\varepsilon_{j\alpha}t/\hbar}\psi_{j\alpha}$.

La evolución de $|\phi\rangle_p$ bajo $U_p(t)$ es el de una partícula libre. Cualitativamente corresponde a

$$\phi(x) \to \phi(x,t) = \phi(x - vt), \qquad v = \frac{1}{M}\langle P\rangle_\phi \tag{5.7}$$

acompañado de un ensanchamiento de la función de onda que puede estimarse como

$$\Delta x(t) = \Delta x + \frac{\Delta p}{M} t\,, \tag{5.8}$$

y a su vez, por el principio de incertidumbre $\Delta x \Delta p \simeq \hbar$, [5.1]

$$\Delta x(t) = \Delta x + \frac{\hbar t}{M \Delta x}. \tag{5.9}$$

La resolución en X depende del instante t en el que se va a hacer la medida y del estado inicial $|\phi\rangle$. Si se toma un estado con Δx demasiado grande se tendrá poca resolución, en cambio si es pequeño para mejorar la resolución, el ensanchamiento crecerá muy deprisa. Para un t dado el valor mínimo de $\Delta x(t)$ corresponde a

$$\Delta x_{\text{mín}}(t) = 2\sqrt{\hbar t / M}\,. \tag{5.10}$$

[5.1] Suponemos que $|\phi\rangle_p$ no tiene Δp mayor del imprescindible, para que Δx no crezca más de lo necesario ya que lo que se va a medir es X.

El efecto de v sobre $X(t)$ se debe tener en cuenta para descontarlo y no atribuirlo a la interacción con S. Para simplificar supondremos sencillamente que el estado $|\phi\rangle$ se ha elegido de modo que $\langle P\rangle = 0$. Es decir, suponemos que $U_p(t)$ no tiene efecto neto sobre $|\phi\rangle$ aparte del ensanchamiento.

Nos queda la acción de $U_I(t)$,

$$|\Psi(t)\rangle_{pS} = e^{-igPAt/\hbar}|\psi(t)\rangle_S \otimes |\phi\rangle_p = \sum_{j,\alpha} \psi_{j\alpha}(t)|j\alpha\rangle_S \otimes e^{-iga_jPt/\hbar}|\phi\rangle_p\,. \tag{5.11}$$

En general el estado queda entrelazado. Teniendo en cuenta que

$$e^{-iqP/\hbar}|\phi\rangle = \int dx\,|x\rangle e^{-iq/\hbar(-i\hbar\partial/\partial_x)}\phi(x) = \int dx\,\phi(x-q)|x\rangle \equiv |\phi(x-q)\rangle \tag{5.12}$$

$$|\Psi(t)\rangle_{pS} = \sum_{j,\alpha} \psi_{j\alpha}(t)|j\alpha\rangle_S \otimes |\phi(x-x_j(t))\rangle, \qquad x_j(t) \equiv ga_jt\,. \tag{5.13}$$

Puesto que x_j depende de a_j, observar la posición del puntero proporciona el valor de A. Para que no haya problemas de resolución entre estados $|j\rangle$ próximos debe cumplirse que

$$|x_{j+1} - x_j| > \Delta x_{\text{mín}}(t). \tag{5.14}$$

Equivalentemente

$$|a_{j+1} - a_j| > \frac{\Delta x_{\text{mín}}(t)}{gt} = \frac{2}{g}\sqrt{\frac{\hbar}{Mt}} \tag{5.15}$$

No hay problema de resolución para M, g ó t suficientemente grandes.

De acuerdo con el postulado de la medida, al medir X se obtiene uno de los valores x_j (dentro de la resolución) y por tanto se selecciona uno y sólo uno de los valores a_j. El estado final del sistema, salvo normalización, es $\sum_{\alpha=1}^{d_j} e^{-i\varepsilon_{j\alpha}t/\hbar}\psi_{j\alpha}|j\alpha\rangle$, que corresponde a $\sum_{\alpha=1}^{d_j}\psi_{j\alpha}|j\alpha\rangle$ en $t=0$ y posterior evolución durante un tiempo t. Cada valor a_j ocurre con probabilidad $\sum_{\alpha=1}^{d_j}|\psi_{j\alpha}|^2$.

En la teoría de los **universos paralelos** se tendría el estado $\sum_{j,\alpha}\psi_{j\alpha}|j\alpha\rangle \otimes |\phi(x)\rangle \otimes |O_0\rangle$ antes de la medida y $\sum_{j,\alpha}\psi_{j\alpha}|j\alpha\rangle \otimes |\phi(x_j)\rangle \otimes |O_j\rangle$ después. $|O\rangle \in \mathscr{H}_O$ es el estado del observador (o del registro del aparato de medida). Después de la medida los espacios $\mathscr{H}_O$ y $\mathscr{H}_S$ quedan entrelazados. No hay colapso: en cada *componente* de la función de onda (cada "universo") el observador ve que A toma el valor a_j. El papel del puntero es que es fácil de entrelazar con $\mathscr{H}_O$ mediante el aparato de medida. [5.2]

[5.2] En contra de lo a veces se afirma, la teoría de los universos paralelos no conlleva la creación de nuevas réplicas en cada medida; la evolución es unitaria y por tanto la dimensión del espacio total se conserva.

La conclusión $|\psi\rangle \to \sum_{\alpha=1}^{d_j} \psi_{j\alpha}|j\alpha\rangle$ es general y se obtiene igualmente aunque el observable A no sea una constante de movimiento.

5.1.2. Medida proyectiva sobre estados puros

Esto corresponde a una **medida ideal** (regla de Lüders). Una medida requiere siempre una interacción y según como se haga el estado final del sistema puede quedar más o menos afectado. El postulado de la medida lo que afirma es que las leyes de la mecánica cuántica permiten hacer una medida ideal. Pero no se puede ir más allá; por ejemplo, el resultado es intrínsecamente aleatorio.

Una forma de expresar el resultado de una medida ideal es decir que se ha aplicado a $|\psi\rangle$ el proyector ortogonal P_j que proyecta sobre el subespacio con $A = a_j$,

$$P_j = \sum_{\alpha=1}^{d_j} |j\alpha\rangle\langle j\alpha|. \tag{5.16}$$

En efecto

$$|\psi\rangle = \sum_j \sum_{\alpha=1}^{d_j} \psi_{j\alpha}|j\alpha\rangle, \qquad P_j|\psi\rangle = \sum_{\alpha=1}^{d_j} \psi_{j\alpha}|j\alpha\rangle \tag{5.17}$$

Dado que A admite la descomposición espectral

$$A = \sum_j a_j P_j \tag{5.18}$$

y los valores de a_j son distintos por definición, se sigue que medir a_j equivale decir cuál de los P_j se obtiene al hacer la medida. La colección de proyectores asociados a A satisface

$$P_j = P_j^\dagger, \qquad P_j P_k = \delta_{jk} I, \qquad \sum_j P_j = I\,. \tag{5.19}$$

El resultado de la medida es uno y sólo uno de los valores j. La colección de proyectores equivale a descomponer $\mathscr{H}$ como suma directa de subespacios ortogonales,

$$\mathscr{H} = \bigoplus_j \mathscr{H}_j, \qquad \mathscr{H}_j \equiv P_j\mathscr{H}, \qquad \dim \mathscr{H}_j = \operatorname{rank} P_j = d_j\,. \tag{5.20}$$

Medir equivale a descomponer $|\psi\rangle$ según esos subespacios

$$|\psi\rangle = \sum_j |\tilde{\psi}_j\rangle, \qquad |\tilde{\psi}_j\rangle \equiv P_j|\psi\rangle, \qquad 1 = \||\psi\rangle\|^2 = \sum_j \||\tilde{\psi}_j\rangle\|^2, \tag{5.21}$$

y quedarse con una de las componentes, con una probabilidad

$$p_j = \| |\tilde{\psi}_j\rangle \|^2 = \|P_j|\psi\rangle\|^2 = \langle\psi|P_j|\psi\rangle. \tag{5.22}$$

El estado normalizado tras la medida pasa a ser

$$|\psi_j\rangle = \frac{|\tilde{\psi}_j\rangle}{\| |\tilde{\psi}_j\rangle \|} = \frac{P_j|\psi\rangle}{\sqrt{\langle\psi|P_j|\psi\rangle}}. \tag{5.23}$$

Algunas observaciones:

i) Si se vuelve a medir el mismo observable A sobre el estado resultante $|\psi_j\rangle$ se obtendrá el mismo valor j (y el mismo estado) con probabilidad 1, por $P_jP_k = \delta_{jk}I$. La medida se puede refinar midiendo un observable que conmute con A pero rompa la degeneración. El máximo refinamiento es una medida asociada a una base ortonormal (módulo una fase), de modo que $\forall j\ d_j = 1$.

ii) Ni la probabilidad p_j ni el estado resultante $|\psi_j\rangle$ dependen de la base concreta $\{|j,\alpha\rangle\}_{\alpha=1}^{d_j}$ de $\mathscr{H}_j$. Una medida ideal no cambia las amplitudes relativas entre los distintos $\psi_{j\alpha}$. La medida implica una pérdida de coherencia cuántica entre j's distintas, pero no entre distintas α's.

iii) Una base ortonormal $|j\rangle$ define una medida maximal, pero todas las bases $e^{i\varphi_j}|j\rangle$ definen la misma descomposición en subespacios y por tanto la misma medida maximal.

iv) Si $\mathscr{H}_S = \mathscr{H}_A \otimes \mathscr{H}_B$ y P_j actúa sobre $\mathscr{H}_A$, su acción se extiende al espacio completo como $P_j \otimes I_B$ (y se denota P_j igualmente). La medida no afecta a $\mathscr{H}_B$. Así si $|\psi\rangle = \sum_{j,k}\psi_{jk}|j\rangle_A \otimes |k\rangle_B$, una medida con resultado j dará $P_j|\psi\rangle = |j\rangle_A \otimes \sum_k \psi_{jk}|k\rangle_B$ (salvo normalización) con probabilidad $\sum_k |\psi_{jk}|^2$. Si A es destruido después de la medida el estado resultante es $\sum_k \psi_{jk}|k\rangle_B$. [5.3]

v) Puesto que una medida requiere interacción y ésta es local, no se puede medir un observable definido sobre dos espacios A y B espacialmente separados. Sí se pueden medir observables separables del tipo $A \otimes B$. El resultado final no depende de en qué orden se apliquen los correspondientes proyectores (de quién produzca el "colapso de la función de onda").

[5.3] Por ejemplo un fotón puede ser absorbido y no existir después de la medida. Más exactamente, el estado del fotón pasa a ser de número de ocupación 0, pero el grado de libertad fotónico en sí sigue existiendo, puede volver a poblarse.

5.1.3. Medida proyectiva sobre estados mezcla

Si se tiene un estado mezcla ρ, se puede aplicar el resultado anterior a una de sus descomposiciones en estados puros

$$\rho = \sum_k q_k |\psi^{(k)}\rangle\langle\psi^{(k)}| \tag{5.24}$$

ρ se puede interpretar como una colección de estados con frecuencia relativa q_k para el estado $|\psi^{(k)}\rangle$. Si se hace una medida de A sobre $|\psi^{(k)}\rangle$ se obtendrá el resultado j con probabilidad $p_{j|k} = \langle\psi^{(k)}|P_j|\psi^{(k)}\rangle$, y quedará un estado normalizado $|\psi_j^{(k)}\rangle = P_j|\psi^{(k)}\rangle/\sqrt{p_{j|k}}$. La mezcla (no normalizada) de estados con resultado j será entonces

$$\tilde{\rho}_j = \sum_k q_k p_{j|k} |\psi_j^{(k)}\rangle\langle\psi_j^{(k)}| = \sum_k q_k P_j |\psi^{(k)}\rangle\langle\psi^{(k)}| P_j = P_j \rho P_j \tag{5.25}$$

La matriz completa, con todos los valores de j después de la medida, es

$$\rho' = \sum_j P_j \rho P_j \tag{5.26}$$

que está correctamente normalizada

$$\mathrm{Tr}(\rho') = \sum_j \mathrm{Tr}(P_j\rho) = \mathrm{Tr}(\rho) = 1 \tag{5.27}$$

ρ' representa el resultado de una **medida no selectiva**: se mide pero no se seleccionan los distintos casos de acuerdo con los valores de j.

Una **medida selectiva** los distintos estados de la colectividad se clasifican de acuerdo con el valor de j obtenido al medir. La matriz de los estados con valor j es $\tilde{\rho}_j$ pero normalizada

$$\rho_j = \frac{\tilde{\rho}_j}{\mathrm{Tr}(\tilde{\rho}_j)} = \frac{P_j \rho P_j}{\mathrm{Tr}(\rho P_j)}\,. \tag{5.28}$$

El denominador es la probabilidad p_j de obtener j en la medida

$$p_j = \mathrm{Tr}(\rho P_j) \tag{5.29}$$

entendida como la frecuencia relativa con la que se obtiene el resultado j al medir en la colectividad ρ. En efecto,

$$p_j = \sum_k p_{j|k} q_k = \sum_k q_k \langle\psi^{(k)}|P_j|\psi^{(k)}\rangle = \mathrm{Tr}(\rho P_j)\,. \tag{5.30}$$

Por supuesto si volvemos a mezclar los varios ρ_j con sus pesos p_j se recupera el resultado de la medida no filtrante:

$$\sum_j p_j \rho_j = \rho'. \tag{5.31}$$

Alternativamente se puede interpretar ρ como un único estado del cual sólo tenemos información parcial, [5.4] p_j es entonces la probabilidad de que en esa única medida el resultado sea j y ρ_j es la nueva mezcla que queda tras obtener esa información. Si el espectro de A está degenerado ρ_j seguirá siendo un estado mezcla en general.

En todo caso las fórmulas obtenidas muestran que no importa qué descomposición concreta de ρ se utilice, los resultados sólo dependen de la matriz densidad ρ. Y también es obvio que todas las fórmulas se reducen a las vistas anteriormente cuando ρ es un estado puro.

Nótese que la aplicación $\rho \to \tilde{\rho}_j = P_j \rho P_j$ (para un j dado) es lineal y de hecho es una operación cuántica (subnormalizada) con operadores de Kraus P_j. La aplicación $\rho \to \rho_j$ no es directamente lineal, [5.5] por eso a veces es más cómodo trabajar con estados o matrices densidad no normalizadas, y normalizar al final. (Todo esto es análogo a lo que ocurre en variables aleatorias y probabilidades en una teoría clásica.) La aplicación $\rho \to \rho' = \sum_j P_j \rho P_j$ es lineal y es un canal cuántico unital.

Otra observación es que una medida no selectiva $\rho \to \rho'$, que clásicamente no tendría ningún efecto, sí tiene un efecto en el caso cuántico. Así por ejemplo, para qubit en estado $|+\rangle$ una medida en la base $\{|+\rangle, |-\rangle\}$ producirá $|+\rangle$ con probabilidad 1. Sin embargo si en lugar de hacer esta medida directamente, se hace una medida no selectiva intermedia en la base $\{|0\rangle, |1\rangle\}$, el estado pasará ser $\rho' = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|$. Si ahora se mide en la base $\{|+\rangle, |-\rangle\}$ se obtendrá cualquiera de los dos valores con igual probabilidad.

Si sobre ρ_j se vuelve a medir el mismo observable se volverá a obtener el mismo estado post-medida ρ_j, y también se obtendrá ρ_j si se vuelve a medir el mismo observable sobre el resultado no selectivo ρ'. Si sobre ρ' se vuelve a hacer una medida no selectiva del mismo observable se volverá a obtener ρ', el canal cuántico asociado es idempotente.

5.2. Medidas generalizadas

Tenemos un sistema con espacio de Hilbert $\mathscr{H}$ y estado $|\psi\rangle$ o más generalmente ρ, y queremos

[5.4] Por ejemplo, se tiene una colectividad de estados y se toma uno al azar.

[5.5] Pero tampoco es una aplicación no lineal arbitraria, es del tipo lineal seguida de normalización. Realmente es una aplicación entre espacios proyectivos.

obtener información sobre el estado haciendo medidas (teniendo en cuenta que una medida afecta al estado). Las medidas cuánticas genuinas son las **medidas proyectivas** descritas anteriormente. Estas medidas están asociadas a una colección de proyectores ortogonales $\{P_j\}_{j=1}^N$

$$P_i P_j = \delta_{ij} P_j, \qquad P_j = P_j^\dagger, \qquad \sum_{j=1}^{N} P_j = I. \tag{5.32}$$

Al hacer la medida se obtiene uno y sólo uno de los valores j, con probabilidad igual a $\langle\psi|P_j|\psi\rangle$ o $\mathrm{Tr}(\rho P_j)$.

La propiedad de completitud $\sum_j P_j = I$ indica que los operadores definen una *medida que toma valores en proyectores*, o PVM (*projector-valued measure*). [5.6]

Sin embargo es posible relajar las condiciones sobre los operadores de medida y formular la medida mediante un conjunto de operadores $\{\Pi_m\}_{m=1}^M$ en $\mathscr{H}$ tales que

$$\Pi_m \geq 0, \qquad \sum_{m=1}^{M} \Pi_m = I, \tag{5.33}$$

de modo que al hacer la medida se obtiene uno y sólo uno de los valores m, con probabilidad $\langle\psi|\Pi_m|\psi\rangle$ o $\mathrm{Tr}(\rho\Pi_m)$. Obviamente se satisface $\sum_m \mathrm{Tr}(\rho\Pi_m) = 1$. Los Π_m son hermíticos al ser positivos pero no se requiere que sean idempotentes ni ortogonales entre sí.

Estas son las denominadas **medidas generalizadas**; las medidas proyectivas están incluidas como un caso particular. Los **operadores de medida** Π_m definen una *medida que toma valores en operadores positivos*, o POVM (*positive operator-valued measure*).

Como se verá, esta formulación generalizada es admisible porque dado una POVM, siempre se puede trabajar en un espacio extendido $\mathscr{H}_W$, en el que la medida generalizada es una medida proyectiva estándar y por tanto se puede realizar físicamente.

Al relajar la condición de ser proyectores, la elección de los operadores de medida es más flexible en una medida generalizada (por ejemplo $N \leq \dim\mathscr{H}$ pero M puede ser arbitrariamente grande). En todo caso, conocidos los Π_m, a través de la propiedad $\mathrm{Prob}(m) = \mathrm{Tr}(\rho\Pi_m)$ se obtiene información sobre el estado ρ al hacer una medida y obtener m como resultado. Otra cuestión es cómo queda el estado después de la medida (si es que sigue existiendo). La extensión $\mathscr{H} \to \mathscr{H}_W$ (que equivale a decir el montaje experimental asociado) no es única y puede consistir en producto tensorial, suma directa

[5.6] Aquí la palabra "medida" es en sentido matemático (como en "medida de Lebesgue") no el de medir un observable.

o ambos. Dependiendo de la extensión puede ocurrir que el estado final en $\mathscr{H}_W$ no corresponda de manera natural a ningún estado de $\mathscr{H}$. La formulación basada en POVM está pensada para obtener información sobre el estado que se mide más que para hacer una **preparación** de un estado postmedida. Además si el estado postmedida no existe en el espacio original $\mathscr{H}$, el concepto de "volver a medir" el mismo POVM sobre ese estado (para ver por ejemplo si se obtiene el mismo valor de m) ni siquiera está definido.

5.3. Reconstrucción de medidas generalizadas

Queremos ver que dado un POVM $\{\Pi_m\}$ en un espacio $\mathscr{H}$, siempre es posible construir un PVM $\{P_m\}$ en otro espacio $\mathscr{H}_W$, de modo que a cualquier estado $|\psi\rangle$ de $\mathscr{H}$ se le pueda asociar otro estado $|\psi_W\rangle$ de $\mathscr{H}_W$ tal que

$$\langle\psi|\Pi_m|\psi\rangle = \langle\psi_W|P_m|\psi_W\rangle. \tag{5.34}$$

Aquí $\mathscr{H}$ y POVM son conocidos, $\mathscr{H}_W$ y el PVM son a elegir. El estado $|\psi\rangle$ no es conocido y puede ser cualquiera. Para que la aplicación $|\psi\rangle \to |\psi_W\rangle$ sea realizable físicamente, debe ser lineal e isométrica, [5.7]

$$U : \mathscr{H} \to \mathscr{H}_W, \qquad U^\dagger U = I, \qquad |\psi_W\rangle = U|\psi\rangle, \tag{5.35}$$

entonces [5.8]

$$\forall|\psi\rangle \quad \langle\psi|\Pi_m|\psi\rangle = \langle\psi_W|P_m|\psi_W\rangle \iff \Pi_m = U^\dagger P_m U. \tag{5.36}$$

La condición de completitud $\sum_m \Pi_m = I$ en $\mathscr{H}$ se satisface por $\sum_m P_m = I_W$ en $\mathscr{H}_W$. En realidad esta última condición es innecesariamente restrictiva, basta que $\sum_m P_m$ actúe como la identidad en el espacio imagen de U, que denotamos $\mathscr{H}'$

$$\mathscr{H}' \equiv U\mathscr{H} = \mathrm{ran}(U) \subset \mathscr{H}_W \tag{5.37}$$

El operador U es isométrico como operador $\mathscr{H} \to \mathscr{H}_W$ y unitario como operador $\mathscr{H} \to \mathscr{H}'$, y cumple $UU^\dagger = P_{\mathscr{H}'}$, el proyector sobre $\mathscr{H}'$. El operador $U^\dagger : \mathscr{H}_W \to \mathscr{H}$ queda definido por la propiedad $\langle\psi|U^\dagger|\phi_W\rangle = \langle\phi_W|U|\psi\rangle^*$. $U^\dagger$ es U^{-1} cuando actúa en $\mathscr{H}'$ y se anula en $\mathscr{H}_W \ominus \mathscr{H}'$.

[5.7] Es importante notar que no se trata de un estado $|\psi\rangle$ conocido, digamos su función de onda, (o equivalentemente tener muchas copias del mismo) y aplicar unas fórmulas para construir un $|\psi_W\rangle$ asociado que cumpla (5.34), eso sería un canal clásico. Tenemos sólo un ejemplar de $|\psi\rangle$ que no podemos examinar sin alterarlo y debe procesarse sin conocerlo, de manera automática explotando las posibilidades que ofrece la amplitud de probabilidad y el canal genuinamente cuántico.

[5.8] Dados dos operadores A y B en un espacio de Hilbert complejo, se cumple:

$$A = B \iff \forall|\psi\rangle\ \langle\psi|A|\psi\rangle = \langle\psi|B|\psi\rangle.$$

Está claro que una reconstrucción U para estados puros se extiende a estados mezcla: $\mathrm{Tr}(\rho\Pi_m) = \mathrm{Tr}(\rho_W P_m)$ con $\rho_W = U\rho U^\dagger$.

5.3.1. Teorema de Naimark

A un operador de medida Π_m en $\mathscr{H}$ se le puede asociar un operador A_m tal que

$$\Pi_M = A_m^\dagger A_m \tag{5.38}$$

y elegir que A_m esté definido de $\mathscr{H} \to \mathscr{H}$ (no es la elección más general). En este caso la solución más general es

$$A_m = U_m \Pi_m^{1/2} \tag{5.39}$$

(se entiende la única raíz cuadrada positiva de Π_m) donde U_m es un operador unitario cualquiera. [5.9] La condición de normalización de los Π_m implica

$$\sum_m A_m^\dagger A_m = I \tag{5.40}$$

por tanto los operadores A_m son operadores de Kraus y definen un canal cuántico

$$\rho \to T(\rho) = \sum_{m=1}^{M} A_m \rho A_m^\dagger. \tag{5.41}$$

Como ya se vio (Tema 4.1.1) se puede reconstruir un canal cuántico como un operador unitario $\tilde{U}$ en un espacio extendido $\mathscr{H}_W = \mathscr{H} \otimes \mathscr{H}_B$ con $\dim \mathscr{H}_B \geq M$. Para $|\phi\rangle_B$ un estado normalizado cualquiera y $\{|m\rangle_B\}$ una base ortonormal, el operador es

$$\tilde{U}|\psi\rangle \otimes |\phi\rangle_B = \sum_{m=1}^{M} A_m|\psi\rangle \otimes |m\rangle_B \equiv |\psi_W\rangle \tag{5.42}$$

U corresponde a la aplicación lineal $|\psi\rangle \to |\psi_W\rangle$. U es isométrico y $\tilde{U}$ es unitario. Los proyectores ortogonales en $\mathscr{H}_W$ se definen como

$$P_m = I \otimes E_m, \qquad E_m = |m\rangle\langle m|_B \tag{5.43}$$

Por construcción

$$\langle\psi_W|P_m|\psi_W\rangle = \langle\psi|A_m^\dagger A_m|\psi\rangle = \langle\psi|\Pi_m|\psi\rangle \tag{5.44}$$

[5.9] Todo matriz cuadrada A se puede expresar como $A = UH$ siendo U unitaria y $H \geq 0$. $A^\dagger A = H^2$ implica $H = (A^\dagger A)^{1/2}$.

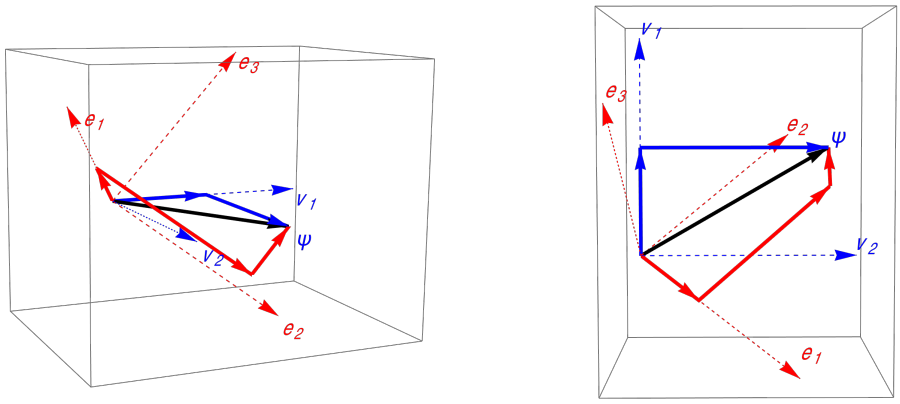


Figura 5.1: Los vectores v_1 y v_2 forman una base ortonormal del espacio $\mathscr{H}$ en el que está ψ y definen una medida proyectiva en ese espacio. Los vectores e_1, e_2 y e_3 forman una base ortonormal de $\mathscr{H}_W \supset \mathscr{H}$. Definen una medida proyectiva en ese espacio, que se ve como POVM en $\mathscr{H}$. Se muestran las dos descomposiciones, $\psi = \sum_{i=1}^{2} v_i \cdot \psi \, v_i = \sum_{\alpha=1}^{3} e_\alpha \cdot \psi \, e_\alpha$. La Fig. *b* muestra la proyección en $\mathscr{H}$.

Tras una medida proyectiva ideal del PVM $\{P_m\}$ con resultado m, el estado no normalizado es

$$P_m|\psi_W\rangle = A_m|\psi\rangle \otimes |m\rangle_B \tag{5.45}$$

Dado que es separable, en la construcción de Naimark se puede identificar el estado postmedida unívocamente como un estado en el espacio original $\mathscr{H}$, a saber, $A_m|\psi\rangle$ (salvo normalización). Sin embargo la medida no es proyectiva y eso implica que si se vuelve a medir el POVM sobre el estado $A_m|\psi\rangle$ puede obtenerse otro valor m', con probabilidad

$$\text{Prob}(m'|m,\psi) = \frac{\langle\psi|A_m^\dagger \Pi_{m'} A_m|\psi\rangle}{\langle\psi|\Pi_m|\psi\rangle} \qquad \text{(Naimark)} \tag{5.46}$$

En medidas proyectivas, el operador P_j desempeñaba el papel de operador de Kraus (el estado postmedida es $P_j|\psi\rangle$) y también de operador de medida (la probabilidad es $\langle\psi|P_j|\psi\rangle$). En medidas generalizadas A_m y Π_m son operadores distintos, relacionados mediante $\Pi_m = A_m^\dagger A_m$.

5.3.2. Construcción general

Tenemos un POVM $\{\Pi_m\}_{m=1}^{M}$ en un espacio $\mathscr{H}$. Cada operador de medida Π_m define un espacio

imagen $\mathscr{H}_m \equiv \mathrm{ran}(\Pi_m) = \Pi_m \mathscr{H}$ de dimensión d_m, con base ortonormal $\{|\Phi_{m\alpha}\rangle\}_{\alpha=1}^{d_m}$,

$$\Pi_m = \sum_{\alpha=1}^{d_m} h_{m\alpha} |\Phi_{m\alpha}\rangle\langle\Phi_{m\alpha}| \equiv \sum_{\alpha=1}^{d_m} |\widetilde{\Phi}_{m\alpha}\rangle\langle\widetilde{\Phi}_{m\alpha}| \tag{5.47}$$

Nótese que los $|\Phi_{m\alpha}\rangle$ son ortogonales para α's distintos pero no para m's distintos. Los pesos satisfacen $0 \leq h_{m\alpha} \leq 1$ pero no suman 1. Además $\mathscr{H}$ no es suma directa de los $\mathscr{H}_m$. Sí se cumple

$$\sum_{m\alpha} |\widetilde{\Phi}_{m\alpha}\rangle\langle\widetilde{\Phi}_{m\alpha}| = \sum_m \Pi_m = I. \tag{5.48}$$

Elegimos un espacio $\mathscr{H}_W$ tal que $\dim \mathscr{H}_W \geq \sum_{m=1}^{M} d_m$. Se puede entonces descomponer el espacio en suma directa de subespacios ortogonales $\mathscr{H}_{W,m}$ con dimensión d_m, más subespacios adicionales hasta completar la dimensión total. Por simplicidad supondremos no hacen falta tales subespacios.

En cada subespacio $\mathscr{H}_{W,m}$ elegimos una base ortonormal $\{|u_{m\alpha}\rangle\}_{\alpha=1}^{d_m}$, de modo que el proyector ortogonal P_m sobre $\mathscr{H}_{W,m}$ es

$$P_m = \sum_{\alpha=1}^{d_m} |u_{m\alpha}\rangle\langle u_{m\alpha}|, \qquad \sum_m P_m = I_W. \tag{5.49}$$

Podemos definir los operadores A_m y el operador U, de $\mathscr{H} \to \mathscr{H}_W$, como

$$A_m \equiv \sum_\alpha |u_{m\alpha}\rangle\langle\widetilde{\Phi}_{m\alpha}|, \qquad U \equiv \sum_m A_m \tag{5.50}$$

Nótese que $A_m^\dagger A_{m'} = 0$ si $m \neq m'$. Y también que $P_m U = A_m$. Se deduce que

$$\Pi_m = A_m^\dagger A_m = U^\dagger P_m U \tag{5.51}$$

y U es isométrico ya que

$$U^\dagger U = U^\dagger \sum_m P_m U = \sum_m \Pi_m = I. \tag{5.52}$$

Por construcción la probabilidad de obtener el valor m al medir el PVM en $\mathscr{H}_W$ es justamente $\langle\psi|\Pi_m|\psi\rangle$. El estado postmedida no normalizado es

$$|\widetilde{\psi}_{W,m}\rangle = P_m U|\psi\rangle = A_m|\psi\rangle \in \mathscr{H}_W. \tag{5.53}$$

En general este estado postmedida no se puede identificar de manera natural con ningún estado $|\psi'\rangle$ del espacio original $\mathscr{H}$ tal que $U|\psi'\rangle = A_m|\psi\rangle$. Aunque $U = \sum_m A_m$, en general $A_m\mathscr{H} \not\subset U\mathscr{H}$.

La construcción de Naimark es un caso particular. Sólo hay que tener en cuenta que los operadores A_m aquí no se corresponden exactamente con los de la construcción de Naimark. La relación es $A_m = A_m^{\text{Naimark}} \otimes |m\rangle_B$. En la construcción de Naimark $|u_{m\alpha}\rangle = |v_{m\alpha}\rangle \otimes |m\rangle_B$, siendo $\{|v_{m\alpha}\rangle\}_{\alpha=1}^{d_m}$ una base ortonormal cualquiera de $\mathscr{H}_m$, y $\mathscr{H}_{W,m} = \mathscr{H} \otimes |m\rangle_B$ de modo que $\mathscr{H}_W = \mathscr{H} \otimes \mathscr{H}_B$.

Otra observación importante es que la realización física de U (isométrico) sólo es posible de manera unitaria. Esto es, $\mathscr{H}$ debe ser un subespacio de un espacio $\mathscr{H}_P$ de modo que $\tilde{U} : \mathscr{H}_P \to \mathscr{H}_W$ sea unitario y U sea la restricción de $\tilde{U}$ a $\mathscr{H}$. Eso ocurre en la construcción de Naimark y también en los ejemplos que siguen.

5.3.3. Ejemplo de implementación de POVM y medida generalizada

En el espacio de un qubit consideremos el POVM $\{\Pi_a, \Pi_b, \Pi_c\}$, donde

$$\Pi_a = \begin{pmatrix} |t|^2 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \Pi_b = \frac{1}{2}\begin{pmatrix} |r|^2 & r^* \\ r & 1 \end{pmatrix}, \qquad \Pi_c = \frac{1}{2}\begin{pmatrix} |r|^2 & -r^* \\ -r & 1 \end{pmatrix}, \quad |t|^2 + |r|^2 = 1. \tag{5.54}$$

Los operadores están normalizados

$$\Pi_a + \Pi_b + \Pi_c = I, \tag{5.55}$$

y son positivos, ya que pueden expresarse como

$$\Pi_m = |\widetilde{\Phi}_m\rangle\langle\widetilde{\Phi}_m|, \qquad |\widetilde{\Phi}_a\rangle = \begin{pmatrix} t^* \\ 0 \end{pmatrix}, \qquad |\widetilde{\Phi}_b\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} r^* \\ 1 \end{pmatrix}, \qquad |\widetilde{\Phi}_c\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} -r^* \\ 1 \end{pmatrix}. \tag{5.56}$$

Nótese que los estados $|\widetilde{\Phi}_m\rangle$ no son ortogonales entre sí (no podrían, ya que un qubit no admite tres estados ortogonales no nulos). En este ejemplo los tres operadores de medida tienen rango 1, por tanto hace falta un espacio $\dim \mathscr{H}_W \geq 3$.

Al hacer una medida generalizada sobre un qubit $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, las probabilidades $\text{Prob}(m) = \langle\psi|\Pi_m|\psi\rangle$, son

$$\text{Prob}(a) = |t\alpha|^2, \qquad \text{Prob}(b) = |\frac{1}{\sqrt{2}}(\alpha r + \beta)|^2, \qquad \text{Prob}(c) = |\frac{1}{\sqrt{2}}(\alpha r - \beta)|^2. \tag{5.57}$$

A su vez, estas probabilidades fijan los operadores de medida Π_m unívocamente.

La reconstrucción como una medida proyectiva se puede hacer mediante fotones como se muestra en la Fig. 5.2a. $|\psi\rangle$ es un fotón y el qubit está codificado en su estado de polarización. Los detectores $D_{a,b,c}$ corresponden a los operadores de medida $\Pi_{a,b,c}$.

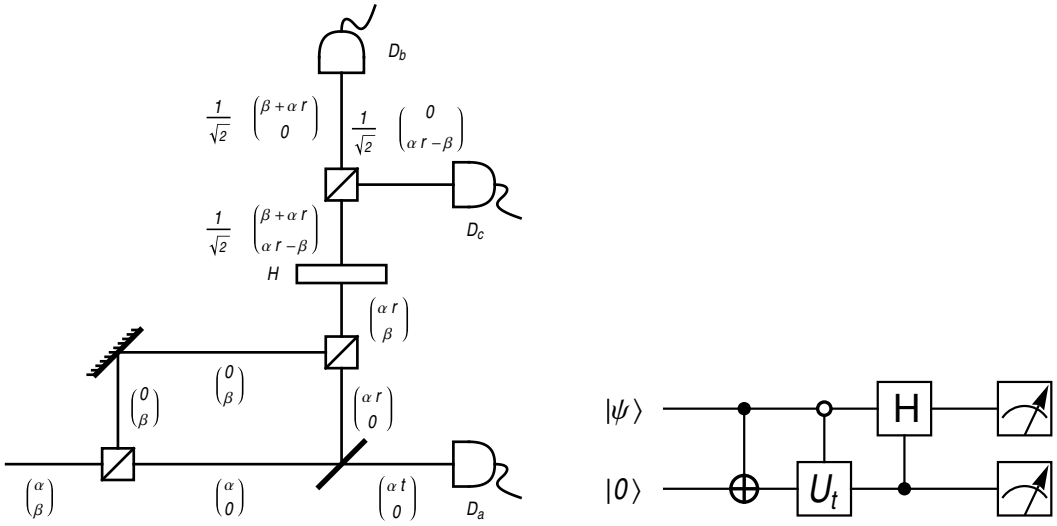


Figura 5.2: Circuito para $\Pi_a + \Pi_b + \Pi_c = I$ con fotones (con algunas licencias) y con puertas cuánticas.

Lo mismo se consigue con el circuito de la Fig. 5.2b. Ahí U_t es un operador unitario de un qubit

$$U_t = \begin{pmatrix} t & -r^* \\ r & t^* \end{pmatrix}, \qquad |t|^2 + |r|^2 = 1. \tag{5.58}$$

Las sucesivas puertas producen

$$\begin{aligned} |\psi\rangle \otimes |0\rangle &= (\alpha|0\rangle + \beta|1\rangle) \otimes |0\rangle = \alpha|0\rangle \otimes |0\rangle + \beta|1\rangle \otimes |0\rangle \\ &\to \alpha|0\rangle \otimes |0\rangle + \beta|1\rangle \otimes X|0\rangle = \alpha|0\rangle \otimes |0\rangle + \beta|1\rangle \otimes |1\rangle \\ &\to \alpha|0\rangle \otimes U_t|0\rangle + \beta|1\rangle \otimes |1\rangle = \alpha|0\rangle \otimes (t|0\rangle + r|1\rangle) + \beta|1\rangle \otimes |1\rangle \\ &\quad = \alpha t|0\rangle \otimes |0\rangle + (\alpha r|0\rangle + \beta|1\rangle) \otimes |1\rangle \\ &\to \alpha t|0\rangle \otimes |0\rangle + H(\alpha r|0\rangle + \beta|1\rangle) \otimes |1\rangle \\ &\quad = \alpha t|00\rangle + \frac{1}{\sqrt{2}}(\alpha r + \beta)|01\rangle + \frac{1}{\sqrt{2}}(\alpha r - \beta)|11\rangle \\ &\equiv |\widetilde{\psi}_{W,a}\rangle + |\widetilde{\psi}_{W,b}\rangle + |\widetilde{\psi}_{W,c}\rangle = |\psi_W\rangle \end{aligned} \tag{5.59}$$

La medida proyectiva es medir $|\psi_W\rangle$ sobre la base computacional de $\mathscr{H}_W$ que es un espacio de dos qubits (aunque la dimensión $|10\rangle$ no se usa) y de nuevo se reproducen las probabilidades correctas. Los estados postmedida no son de $\mathscr{H}'$ y no tienen identificación natural con un estado del qubit inicial.

El espacio inicial de un qubit, dimensión 2, no puede soportar tres proyectores ortogonales. Se ha llevado el estado $|\psi\rangle$ (la información sobre sus amplitudes α, β) a un espacio de dimensión mayor de modo que el estado $|\psi_W\rangle$ ya se puede descomponer en tres trozos ortogonales para medirlos proyectivamente con los pesos correctos.

En la realización física el operador de evolución debe ser unitario (invertible además de isométrico) por eso es necesario un qubit auxiliar *en el estado inicial*. [5.10] Por otro lado aunque es un sistema de dos qubits, al variar $|\psi\rangle$ (es decir, α y β) $|\psi_W\rangle$ sólo ocupa un subespacio $\mathscr{H}' \subset \mathscr{H}_W$ de dimensión 2, igual que $\mathscr{H}$. Es importante notar que la información del estado puede tomar distintos aspectos, realizarse en distintos espacios (soportes físicos) y fluir de uno a otro. Las distintas realizaciones son equivalentes mientras se mantenga la coherencia cuántica (evolución lineal unitaria). Por ejemplo, el circuito se puede invertir y recuperar $|\psi\rangle$ a partir de $|\psi_W\rangle$, o bien producir $|0\rangle \otimes |\psi\rangle$ en vez de $|\psi\rangle \otimes |0\rangle$, aunque el soporte físico es distinto la información es la misma.

5.4. Ejemplos: estrategias para discriminación de estados

El problema de identificación de estados, en su forma más simple, es que nos presentan un estado de un espacio $\mathscr{H}$ conocido, que puede ser uno cualquiera de dos posibles $|\psi_1\rangle$ y $|\psi_2\rangle$, conocidos, pero no sabemos cuál. Si los estados $|\psi_1\rangle$ y $|\psi_2\rangle$ son ortogonales una identificación perfecta es posible. Basta hacer una medida proyectiva en una base ortonormal de la que formen parte. [5.11] Veamos que una identificación perfecta *no es posible cuando los estados no son ortogonales* (se entiende con una sola medida, sólo tenemos una copia del estado, que puede quedar alterada al medir). Es un teorema de imposibilidad, análogo al de no clonación.

Supongamos un POVM tal que $\Pi_1 + \Pi_2 = I$, $\Pi_m \geq 0$, tal que resulte Π_1 si y sólo si llega $|\psi_1\rangle$ y resulte Π_2 si y sólo si llega $|\psi_2\rangle$. Esa identificación perfecta requiere

$$\begin{aligned} \langle\psi_1|\Pi_1|\psi_1\rangle &= 1, \qquad \langle\psi_1|\Pi_2|\psi_1\rangle = 0, \\ \langle\psi_2|\Pi_2|\psi_2\rangle &= 1, \qquad \langle\psi_2|\Pi_1|\psi_2\rangle = 0. \end{aligned} \tag{5.60}$$

Entonces (usando que $\Pi_m \geq 0$)

$$0 = \langle\psi_2|\Pi_1|\psi_2\rangle = \|\Pi_1^{1/2}|\psi_2\rangle\|^2 \implies \Pi_1^{1/2}|\psi_2\rangle = 0 \implies \Pi_1|\psi_2\rangle = 0 \tag{5.61}$$

[5.10] El mismo papel lo desempeña $\mathscr{H}_B$ en la construcción de Naimark. En la construcción con fotones, el espacio $\mathscr{H}_P$ completo se pone de manifiesto haciendo incidir fotones con amplitudes arbitrarias por todos los canales que en nuestra construcción son canales salientes. Puesto que todas las puertas son reversibles, esos fotones acabarán saliendo por algún lado (y no sólo por el canal inicial de nuestra construcción).

[5.11] Lo usual es rotar dichos estados para medirlos cómodamente en una base estándar.

Igualmente $\Pi_2|\psi_1\rangle = 0$, entonces

$$0 = \Pi_2|\psi_1\rangle = (I - \Pi_1)|\psi_1\rangle \implies \Pi_1|\psi_1\rangle = |\psi_1\rangle \tag{5.62}$$

Puesto que $|\psi_1\rangle$ y $|\psi_2\rangle$ son ambos vectores propios con valores propios distintos de un operador hermítico, deben ser ortogonales. □

Ante esta imposibilidad hay dos estrategias principales:

i) **Discriminación inequívoca**. Nunca cometer un error en la identificación, a costa de dejar casos sin identificar. Aquí se trata minimizar el número de casos sin identificar.

ii) **Discriminación con mínimo error**. Dar siempre una estimación de cuál es el estado, con cierta posibilidad de error en la identificación, e intentar minimizar ese error. Esta opción se estudia más adelante.

5.4.1. Discriminación inequívoca entre estados

5.4.1.1. Ejemplo de detección sin error

Empezamos examinando esta estrategia con un ejemplo.

Tenemos dos estados posibles de un qubit, $|\psi_1\rangle = |0\rangle$ y $|\psi_2\rangle = |+\rangle$, que llegan con igual probabilidad. Los estados y sus probabilidades son conocidas, pero no sabemos concretamente cuál de los dos estados ha llegado. Se trata de identificar el estado recibido *sin error*. Es decir, las opciones son:

1) Identificar con seguridad el estado como $|0\rangle$.

2) Identificar con seguridad el estado como $|+\rangle$.

3) No identificar el estado.

No se admite una asignación incorrecta de modo que si la identificación no es completamente segura automáticamente estamos en el caso 3. Eso ocurrirá con una probabilidad Q. Obviamente el protocolo será más eficiente cuanto menor sea Q.

Usando medidas proyectivas (PVM) se puede usar por ejemplo

$$\Pi_1' = |-\rangle\langle -|, \qquad \Pi_0' = |+\rangle\langle +|, \qquad \Pi_1' + \Pi_0' = I\,. \tag{5.63}$$

La medida va a dar uno y sólo uno de los dos resultados. Claramente si se dispara el detector Π_1' el estado no puede ser $|+\rangle$, y por tanto tiene que ser $|0\rangle$. En cambio si se dispara Π_0' no se puede decir con seguridad si es $|0\rangle$ o $|+\rangle$.

La probabilidad de observar Π_1' es la probabilidad de que llegue $|0\rangle$ (que es $1/2$) por la probabilidad condicionada de que si llega $|0\rangle$ se dispare Π_1' que es $\langle 0|\Pi_1'|0\rangle = \frac{1}{2}$. Es decir, la probabilidad de identificación segura es $1/4$ y la de no identificación es $Q = 3/4 = 0.75$

Para este problema el método óptimo es un POVM con

$$\Pi_1 = \lambda|-\rangle\langle -|, \qquad \Pi_2 = \lambda|1\rangle\langle 1|, \qquad \Pi_0 = I - \Pi_1 - \Pi_2, \qquad \lambda = 2 - \sqrt{2} = 0.59 \tag{5.64}$$

Si se dispara Π_1 el estado es con seguridad $|0\rangle$, ya que no puede ser $|+\rangle$. Análogamente, si se dispara Π_2 el estado es con seguridad $|+\rangle$, ya que no puede ser $|0\rangle$. En ambos casos hay una identificación inequívoca. Sólo si se dispara Π_0 no hay identificación. Eso ocurre con probabilidad

$$Q = \frac{1}{2}\langle 0|\Pi_0|0\rangle + \frac{1}{2}\langle +|\Pi_0|+\rangle \tag{5.65}$$

Introduciendo la matriz densidad del estado que llega

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|+\rangle\langle +| \tag{5.66}$$

también se puede escribir como

$$\begin{aligned} Q &= \mathrm{Tr}(\rho\Pi_0) = 1 - \mathrm{Tr}(\rho\Pi_1) - \mathrm{Tr}(\rho\Pi_2) = 1 - \lambda\langle -|\rho|-\rangle - \lambda\langle 1|\rho|1\rangle \\ &= 1 - \frac{\lambda}{2}|\langle -|0\rangle|^2 - \frac{\lambda}{2}|\langle 1|+\rangle|^2 = 1 - \frac{\lambda}{2} = \frac{1}{\sqrt{2}} = 0.71 \end{aligned} \tag{5.67}$$

Evidentemente se podría reducir $\mathrm{Tr}(\rho\Pi_0)$ tomando un λ mayor, sin embargo eso no es posible sin violar la condición $\Pi_0 \geq 0$.

Otra observación es que aquí se ha supuesto que ambos estados llegan con probabilidad $1/2$. Si uno de los estados, digamos $|0\rangle$, llegara con mucha más probabilidad (conocida) que $|+\rangle$, el POVM seguiría dando $Q = 0.71$ mientras que el PVM sería más eficiente, con $Q \approx 0.5$.

El circuito descrito en Fig. 5.2 se puede usar para realizar este POVM añadiendo antes una rotación para llevar los estados $|0\rangle$ y $|+\rangle$ a la forma estándar

$$|\psi_1\rangle = \begin{pmatrix}\cos(\theta/2)\\ \text{sen}(\theta/2)\end{pmatrix} \qquad |\psi_2\rangle = \begin{pmatrix}\cos(\theta/2)\\ -\text{sen}(\theta/2)\end{pmatrix} \qquad 0 \le \theta \le \frac{\pi}{2}. \tag{5.68}$$

La condición

$$\frac{1}{\sqrt{2}} = |\langle 0|+\rangle| = |\langle \psi_1|\psi_2\rangle| = \cos(\theta) \tag{5.69}$$

se satisface con $\theta = \frac{\pi}{4}$. Ahora

$$\Pi_1 = \lambda |\psi_2^\perp\rangle\langle\psi_2^\perp|, \qquad \Pi_2 = \lambda |\psi_1^\perp\rangle\langle\psi_1^\perp|, \qquad \Pi_0 = I - \Pi_1 - \Pi_2 \tag{5.70}$$

$\lambda = \dfrac{1}{1+\cos(\theta)}$ es el valor crítico tal que $\Pi_0 \ge 0$ y tiene un autovalor 0.

En estas condiciones, el circuito se aplica con $r = \tan(\theta/2)$. Como se puede comprobar

$$\begin{aligned} &\Pi_1 = \Pi_b = |\widetilde{\Phi_b}\rangle\langle\widetilde{\Phi_b}|, \qquad |\Phi_b\rangle = \begin{pmatrix}\text{sen}(\theta/2)\\ \cos(\theta/2)\end{pmatrix} = |\psi_2^\perp\rangle, \qquad \lambda = \frac{|r|^2+1}{2} \\ &\Pi_2 = \Pi_c = |\widetilde{\Phi_c}\rangle\langle\widetilde{\Phi_c}|, \qquad |\Phi_c\rangle = \begin{pmatrix}-\text{sen}(\theta/2)\\ \cos(\theta/2)\end{pmatrix} = |\psi_1^\perp\rangle. \end{aligned} \tag{5.71}$$

Como se puede ver en (5.59), la elección $r = \tan(\theta/2)$ es la que hace que se anulen las componentes c o b según llegue $|\psi_1\rangle$ o $|\psi_2\rangle$. Por otro lado en este ejemplo $Q = |\alpha t|^2 = \cos(\theta)$ independientemente de cuál de los dos estados llegue. [5.12]

5.4.1.2. Consideraciones generales

Nos envían uno y sólo uno de los estados ρ_j, $j = 1,\ldots,N$ en $\mathscr{H}$, con probabilidades η_j, $\sum_j \eta_j = 1$. Los ρ_j y los η_j son conocidos, lo que no sabemos es cuál de ellos llega concretamente.

Se busca un POVM

$$I = \Pi_0 + \sum_{j=1}^{N} \Pi_j \tag{5.72}$$

[5.12] Toda la construcción es válida para estados del tipo (5.68), no sólo $\theta = \pi/4$. Todo par de estados se puede llevar a la forma (5.68), salvo fase, mediante una transformación de SU(2).

tal que si se dispara el detector Π_j necesariamente ha llegado el estado ρ_j. Π_0 corresponde a no identificación. Se quiere que esto ocurra con una probabilidad Q mínima.

$$Q = \mathrm{Tr}(\rho\Pi_0), \qquad \rho = \sum_{j=1}^{N} \eta_j \rho_j \,. \tag{5.73}$$

La condición de detección sin error requiere $\mathrm{Prob}(\Pi_k|\rho_j) = 0$ si $k \neq j$ y $k \neq 0$,

$$\mathrm{Tr}(\rho_j \Pi_k) = 0 \quad \text{si } k \neq j, \qquad j,k \in \{1,\ldots,N\}. \tag{5.74}$$

Si $\mathscr{V}_j = \rho_j \mathscr{H}$ (el subespacio imagen de ρ_j como operador) se deduce que para $k \neq j$ necesariamente $\Pi_k \mathscr{V}_j = 0$. En efecto, $\rho_j = \sum_\alpha |\widetilde{\varphi}_{j\alpha}\rangle\langle\widetilde{\varphi}_{j\alpha}|$ y $\langle\widetilde{\varphi}_{j\alpha}|\Pi_k|\widetilde{\varphi}_{j\alpha}\rangle = 0$ implica $\Pi_k|\widetilde{\varphi}_{j\alpha}\rangle = 0$.

En particular, esto implica que si $\mathscr{V}_j = \mathscr{H}$ para algún estado, los demás no se podrán identificar y si hay dos estados con imagen igual a todo $\mathscr{H}$ no sería posible una identificación sin error (más allá de la solución trivial $\Pi_0 = I$, $Q = 1$).

La identificación será peor cuanto más se parezcan los estados ρ_j. Así por ejemplo, si $\mathscr{V}_1 = \mathscr{V}_2$ se tendrá $0 = \Pi_1 \mathscr{V}_2 = \Pi_1 \mathscr{V}_1 \implies \mathrm{Tr}(\rho_1 \Pi_1) = 0$, es decir, los estados ρ_1 y ρ_2 no se pueden identificar. En cambio si los $\mathscr{V}_j$ son ortogonales no hay problema en identificar todos los estados (a saber, con una medida proyectiva sobre los $\mathscr{V}_j$) y $Q = 0$.

Nótese que sólo se usan los valores de $\mathrm{Tr}(\rho_j \Pi_m)$, y bastaría $\Pi_m \geq 0$ en $\mathscr{V}_j$, sin embargo es necesario imponer $\Pi_m \geq 0$ sobre todo $\mathscr{H}$ para que las medidas sean físicamente realizables.

Para aligerar la discusión suponemos $N = 2$. Π_1 se anula sobre $\mathscr{V}_2$ y Π_2 se anula sobre $\mathscr{V}_1$.

Si se define $\mathscr{V} = \mathscr{V}_1 + \mathscr{V}_2$, se tendrá $\mathscr{H} = \mathscr{V} \oplus \mathscr{V}^\perp$, y todo ocurre en el subespacio $\mathscr{V}$. Sean I_V e $I_{V^\perp} = I - I_V$ los proyectores ortogonales sobre $\mathscr{V}$ y $\mathscr{V}^\perp$. En principio se quiere Π_1 y Π_2 lo mayores posibles (en $\mathscr{V}$) y Π_0 lo menor posible (en $\mathscr{V}$) para minimizar Q. Un POVM natural es del tipo

$$\begin{aligned} &\Pi_1 = c_1(I_V - P_2), \qquad \Pi_2 = c_2(I_V - P_1), \\ &\Pi_0 = I - \Pi_1 - \Pi_2 = I_{V^\perp} + I_V - c_1(I_V - P_2) - c_2(I_V - P_1), \end{aligned} \tag{5.75}$$

siendo P_1 y P_2 los proyectores ortogonales sobre $\mathscr{V}_1$ y $\mathscr{V}_2$, respectivamente. La idea es que $I_V - P_2$ es el operador más grande en $\mathscr{V}$ que se anula sobre $\mathscr{V}_2$, ídem $I_V - P_1$. [5.13] Las condiciones de positividad son entonces

$$c_1, c_2 \geq 0, \qquad c_1(I_V - P_2) + c_2(I_V - P_1) \leq I_V \,. \tag{5.76}$$

[5.13] Sería ineficiente usar $I - P_2$ e $I - P_1$, ya que entonces $\Pi_0 = (1 - c_1 - c_2)I_{V^\perp} + I_V - c_1(I_V - P_2) - c_2(I_V - P_1)$ y la condición $1 - c_1 - c_2 \geq 0$ impondría condiciones innecesariamente restrictivas sobre $c_{1,2}$.

5.4.1.3. Caso de dos estados puros

Queremos identificar sin error dos estados puros $|\psi_1\rangle$ y $|\psi_2\rangle$ de un espacio $\mathscr{H}$, con probabilidades de llegada η_1 y η_2 $(\eta_1+\eta_2=1)$. Los estados y sus probabilidades son conocidos. Sin pérdida de generalidad se puede trabajar en el subespacio *bidimensional* que subtienden los dos estados y usar la forma (5.68). También podemos suponer

$$0 \leq \eta_2 \leq \frac{1}{2} \leq \eta_1 \leq 1\,. \tag{5.77}$$

(llamando $|\psi_1\rangle$ al estado más probable).

Definimos

$$0 \leq \cos(\theta) \equiv |\langle\psi_1|\psi_2\rangle| \leq 1, \qquad 0 \leq \theta \leq \frac{\pi}{2} \tag{5.78}$$

En el caso degenerado $\eta_2 = 0$, es decir, siempre llega $|\psi_1\rangle$, el POVM $\Pi_1 = |\psi_1\rangle\langle\psi_1|$, $\Pi_2 = 0$ y $\Pi_0 = |\psi_1^\perp\rangle\langle\psi_1^\perp|$ no produce identificaciones incorrectas y $Q = 0$, por tanto es óptimo, independientemente del solapamiento entre los estados.

En lo que sigue suponemos $\eta_2 > 0$. Consideramos un POVM $\Pi_0+\Pi_1+\Pi_2 = I$ de la forma

$$\Pi_1 = c_1|\psi_2^\perp\rangle\langle\psi_2^\perp|, \qquad \Pi_2 = c_2|\psi_1^\perp\rangle\langle\psi_1^\perp|, \qquad 0 \leq c_{1,2} \leq 1. \tag{5.79}$$

Los $c_{1,2}$ también van a estar acotados superiormente por la condición $\Pi_0 \geq 0$.

En este caso (dos estados puros, $\eta_2 > 0$) esta familia de POVM's contiene el POVM óptimo, y se trata de optimizar $c_{1,2}$ para minimizar Q.

Si definimos $p_j = \langle\psi_j|\Pi_j|\psi_j\rangle = c_j\,\mathrm{sen}^2(\theta)$ $(j=1,2)$ los operadores de medida se pueden expresar como

$$\Pi_1 = \frac{p_1}{\mathrm{sen}^2(\theta)}|\psi_2^\perp\rangle\langle\psi_2^\perp|, \qquad \Pi_2 = \frac{p_2}{\mathrm{sen}^2(\theta)}|\psi_1^\perp\rangle\langle\psi_1^\perp|. \tag{5.80}$$

p_j es la probabilidad de identificar el estado $|\psi_j\rangle$ cuando ése es el estado que llega. Entonces $q_j = 1-p_j$ es la probabilidad de que llegue $|\psi_j\rangle$ y no se identifique (el resultado de la medida sea Π_0). Por tanto la probabilidad de no identificación es

$$Q = \eta_1 q_1 + \eta_2 q_2. \tag{5.81}$$

Nótese que las condiciones $c_j \leq 1$ implican $p_j \leq \mathrm{sen}^2(\theta)$ y por tanto $q_j \geq \cos^2(\theta)$. Entonces para estos POVM $Q \geq |\langle\psi_1|\psi_2\rangle|^2$.

Usamos $q_j \in [0,1]$ como parámetros. En este caso $\Pi_{1,2} \geq 0$ automáticamente. Como puede verificarse, la condición $\Pi_0 \geq 0$ requiere

$$q_1 q_2 \geq \cos^2(\theta). \tag{5.82}$$

La desigualdad se verifica cuando los dos autovalores de $\Pi_0 = I - \Pi_1 - \Pi_2$ son positivos, y la igualdad cuando uno de los autovalores se anula. [5.14]

El problema es minimizar $Q = \eta_1 q_1 + \eta_2 q_2$ respecto de $q_{1,2}$ sujeto a las condiciones $q_j \in [0,1]$ y $q_1 q_2 \geq \cos^2(\theta)$. Los parámetros η_1 y θ son datos.

Cuando $\cos(\theta) = 0$, el mínimo corresponde a $q_1 = q_2 = 0$, $\Pi_0 = 0$ y $Q = 0$.

En otro caso $q_1, q_2 > 0$. Si el par (q_1, q_2) satisface las ligaduras y $q_1 q_2 > \cos^2(\theta)$ es claro que siempre se puede reducir Q reduciendo q_1 hasta que cumpla $q_1 q_2 = \cos^2(\theta)$. Por tanto a partir de ahora podemos suponer que

$$q_1 = \frac{\cos^2(\theta)}{q_2}, \qquad \cos^2(\theta) \leq q_2 \leq 1\,, \tag{5.83}$$

(la cota inferior procede de $q_1 \leq 1$) y

$$Q = \eta_1 \frac{\cos^2(\theta)}{q_2} + \eta_2 q_2\,. \tag{5.84}$$

El mínimo absoluto de Q imponiendo sólo $q_2 > 0$ es

$$\bar{q}_2 = \sqrt{\frac{\eta_1}{\eta_2}} \cos\theta\,. \tag{5.85}$$

Este valor siempre satisface la condición $\cos^2(\theta) \leq \bar{q}_2$, ya que $\cos(\theta) \leq 1 \leq \sqrt{\frac{\eta_1}{\eta_2}}$ (por $\eta_1 \geq \eta_2$). Sin embargo $\bar{q}_2$ puede violar la condición $\bar{q}_2 \leq 1$, en cuyo caso valor el óptimo corresponderá a $q_2 = 1$. Por tanto hay que distinguir dos casos:

a) $|\langle \psi_1 | \psi_2 \rangle|^2 \leq \dfrac{\eta_2}{\eta_1}$ (solapamiento pequeño o probabilidades parecidas)

[5.14] Aquí se utiliza que el espacio es bidimensional. Nótese que las probabilidades de llegada $\eta_{1,2}$ no desempeñan ningún papel en esta desigualdad.

En este caso los valores óptimos son

$$q_1 = \sqrt{\frac{\eta_2}{\eta_1}}\,|\langle\psi_1|\psi_2\rangle|, \qquad q_2 = \sqrt{\frac{\eta_1}{\eta_2}}\,|\langle\psi_1|\psi_2\rangle|, \qquad Q = 2\sqrt{\eta_1\eta_2}\,|\langle\psi_1|\psi_2\rangle|. \tag{5.86}$$

Cuando $\eta_1 = \eta_2$ se está en este caso y

$$Q = |\langle\psi_1|\psi_2\rangle| \qquad (\eta_1 = \eta_2 = \frac{1}{2}). \tag{5.87}$$

b) $\dfrac{\eta_2}{\eta_1} \leq |\langle\psi_1|\psi_2\rangle|^2$ (solapamiento grande o probabilidades desiguales)

En este caso los valores óptimos son

$$q_1 = |\langle\psi_1|\psi_2\rangle|^2, \qquad q_2 = 1, \qquad Q = \eta_1|\langle\psi_1|\psi_2\rangle|^2 + \eta_2. \tag{5.88}$$

Corresponde a la medida proyectiva $\Pi_1 = |\psi_2^\perp\rangle\langle\psi_2^\perp|,\ \Pi_2 = 0,\ \Pi_0 = |\psi_2\rangle\langle\psi_2|$.

Cuando $|\langle\psi_1|\psi_2\rangle| \to 1$, $Q \to 1$, y no hay posibilidad de identificación segura. Y viceversa, si $\langle\psi_1|\psi_2\rangle \to 0,\ Q \to 0$, siempre se identifican (usando el POVM óptimo, que de hecho es PVM).

Cuando $\eta_1 = \eta_2,\ Q = |\langle\psi_1|\psi_2\rangle|$, en cambio cuando $\eta_1 \to 1,\ Q \to |\langle\psi_1|\psi_2\rangle|^2$.

5.4.2. Identificación de estados con error mínimo

5.4.2.1. Ejemplo de identificación con error mínimo

Tenemos tres estados de un qubit (el llamado **conjunto trino**)

$$|\psi_1\rangle = |0\rangle, \qquad |\psi_2\rangle = \frac{1}{2}(|0\rangle + \sqrt{3}|1\rangle), \qquad |\psi_3\rangle = \frac{1}{2}(|0\rangle - \sqrt{3}|1\rangle). \tag{5.89}$$

Forman un triángulo equilátero en el plano xz del espacio de Bloch.

Me envían uno cualquiera de estos estados con igual probabilidad y se trata de acertar cuál es. Si se apuesta al azar la probabilidad de acertar es $P_{\text{corr}} = 1/3$, la de error $P_{\text{err}} = 2/3 = 0.67$.

Más sofisticado es hacer una medida. Si insisto en una medida proyectiva, se puede tomar por ejemplo

$$P_0 = |0\rangle\langle 0|, \qquad P_1 = |1\rangle\langle 1|. \tag{5.90}$$

Si sale P_0 apuesto por $|\psi_1\rangle$, en otro caso apuesto al 50 % por $|\psi_2\rangle$ o $|\psi_3\rangle$, al azar. [5.15]

$1/3$ de las veces llega $|\psi_1\rangle$, entonces sale P_0 seguro y acierto.
$1/3$ de las veces llega $|\psi_2\rangle$, entonces sale P_1 con probabilidad $\langle\psi_2|P_1|\psi_2\rangle = \frac{3}{4}$ y acierto con probabilidad $1/2$.
Ídem para $|\psi_3\rangle$.

En conjunto, la probabilidad de acertar es

$$P_{\text{corr}} = \frac{1}{3}\times 1 + \frac{1}{3}\times\frac{3}{4}\times\frac{1}{2} + \frac{1}{3}\times\frac{3}{4}\times\frac{1}{2} = \frac{7}{12}, \qquad P_{\text{err}} = \frac{5}{12} = 0.42 \tag{5.91}$$

La estrategia más eficiente en este problema es usar un POVM con

$$\Pi_j = \frac{2}{3}|\psi_j\rangle\langle\psi_j|, \qquad j = 1,2,3 \tag{5.92}$$

y apostar por $|\psi_j\rangle$ si sale Π_j. Como se puede comprobar los operadores satisfacen la condición de normalización,

$$\sum_{j=1}^{3}\Pi_j = \frac{2}{3}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} + \frac{2}{3}\begin{pmatrix}\frac{1}{4} & \frac{\sqrt{3}}{4}\\ \frac{\sqrt{3}}{4} & \frac{3}{4}\end{pmatrix} + \frac{2}{3}\begin{pmatrix}\frac{1}{4} & -\frac{\sqrt{3}}{4}\\ -\frac{\sqrt{3}}{4} & \frac{3}{4}\end{pmatrix} = I. \tag{5.93}$$

Cuando llega un estado $|\psi_j\rangle$ cualquiera, la probabilidad de acertar es $\langle\psi_j|\Pi_j|\psi_j\rangle = \frac{2}{3}$. Es decir, en conjunto, la probabilidad de acertar es $P_{\text{corr}} = 2/3$, la de error $P_{\text{err}} = 1/3 = 0.33$ Es una mejora sustancial con respecto del mejor PVM.

El circuito descrito en Fig. 5.2 permite reproducir este POVM eligiendo $r = 1/\sqrt{3}$ (por tanto $|t|^2 = \frac{2}{3}$). En este caso

$$|\psi_1\rangle = |\Phi_a\rangle, \qquad |\psi_2\rangle = |\Phi_b\rangle, \qquad |\psi_3\rangle = |\Phi_c\rangle. \tag{5.94}$$

[5.15] Equivale al POVM $\Pi'_1 = P_0$, $\Pi'_2 = \Pi'_3 = \frac{1}{2}P_1$ y apostar por $|\psi_j\rangle$ si sale Π'_j.

5.4.2.2. Consideraciones generales

Tenemos N estados ρ_j, $j = 1, \ldots, N$, en un espacio $\mathscr{H}$ cualquiera, con probabilidades η_j, $\sum_j \eta_j = 1$. Los ρ_j y las η_j son conocidos.

Elegimos un POVM con N operadores Π_j y apostamos por ρ_j si el resultado de la medida es Π_j.

Si llega ρ_j la probabilidad de que la medida resulte en Π_j (y acertemos en nuestra apuesta) es $\text{Prob}(\Pi_j|\rho_j) = \text{Tr}(\rho_j\Pi_j)$. Entonces la probabilidades de acertar/errar son

$$P_{\text{corr}} = \sum_j \eta_j \,\text{Tr}(\rho_j\Pi_j), \qquad P_{\text{err}} = 1 - P_{\text{corr}}. \tag{5.95}$$

Se trata elegir el POVM tal que se minimice la probabilidad de error de identificación.

Teorema (Helstrom) Sea $\Gamma \equiv \sum_{j=1}^{N} \eta_j\rho_j\Pi_j$. Una condición necesaria y suficiente para que el POVM $\{\Pi_j\}_{j=1}^{N}$ sea óptimo (P_{err} mínimo) es [5.16]

$$\Gamma \geq \eta_j\rho_j \quad \forall j. \tag{5.96}$$

(La demostración puede verse en el libro de Barnett, pág. 100.) Nótese que

$$P_{\text{corr}} = \text{Tr}(\Gamma), \qquad P_{\text{err}} = 1 - \text{Tr}(\Gamma)\,. \tag{5.97}$$

En general esta condición no determina unívocamente el POVM y la solución puede no ser única (aunque todas los POVM óptimos producen el mismo operador Γ).

Para valores arbitrarios de N, la condición de Helstrom no proporciona directamente un POVM. Pero sí hay solución explícita cuando $N = 2$, y el POVM es de hecho una medida proyectiva (PVM).

5.4.2.3. Caso de dos estados

Para $N = 2$, definimos el operador hermítico

$$\Lambda \equiv \eta_2\rho_2 - \eta_1\rho_1 \tag{5.98}$$

[5.16] Es decir, $\langle\psi|\Gamma|\psi\rangle \geq \eta_j\langle\psi|\rho_j|\psi\rangle \quad \forall j, |\psi\rangle$.

de modo que

$$P_{\rm err} = 1 - {\rm Tr}(\eta_1\rho_1\Pi_1 + \eta_2\rho_2\Pi_2) = {\rm Tr}(\eta_1\rho_1\Pi_2 + \eta_2\rho_2\Pi_1) = \eta_1 + {\rm Tr}(\Lambda\Pi_1)\,. \tag{5.99}$$

Igualmente (intercambiando 1 y 2)

$$P_{\rm err} = \eta_2 - {\rm Tr}(\Lambda\Pi_2). \tag{5.100}$$

Estas fórmulas sugieren que el mínimo error se hallará cuando Π_1 proyecte sobre el subespacio $\Lambda < 0$ y Π_2 sobre el subespacio $\Lambda > 0$, y en efecto es así.

Sea $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$ el espectro de Λ, siendo $d = \dim \mathscr{H}$, y $\{|\phi_k\rangle\}$ los autoestados normalizados. Definimos el POVM

$$\Pi_1 = \sum_{\lambda_k<0} |\phi_k\rangle\langle\phi_k|, \qquad \Pi_2 = \sum_{\lambda_k\geq 0} |\phi_k\rangle\langle\phi_k|\,. \tag{5.101}$$

Π_1 es el proyector sobre $\Lambda < 0$ y Π_2 sobre $\Lambda \geq 0$. Obviamente $\Pi_1 + \Pi_2 = I$. [5.17]

Para ver que este POVM es óptimo basta verificar la condición de Helstrom:

$$\Gamma - \eta_1\rho_1 = \eta_1\rho_1\Pi_1 + \eta_2\rho_2\Pi_2 - \eta_1\rho_1 = \eta_2\rho_2\Pi_2 - \eta_1\rho_1\Pi_2 = \Lambda\Pi_2 \geq 0\,. \tag{5.102}$$

($\Lambda\Pi_2 \geq 0$ ya que Π_2 proyecta sobre el subespacio $\Lambda \geq 0$.) Igualmente

$$\Gamma - \eta_2\rho_2 = \eta_1\rho_1\Pi_1 - \eta_2\rho_2\Pi_1 = -\Lambda\Pi_1 \geq 0\,. \qquad \square \tag{5.103}$$

Con este POVM

$$\begin{aligned} P_{\rm err} &= \eta_1 + {\rm Tr}(\Lambda\Pi_1) = \eta_1 + \sum_{\lambda_k<0} \lambda_k \\ P_{\rm err} &= \eta_2 - {\rm Tr}(\Lambda\Pi_2) = \eta_2 - \sum_{\lambda_k\geq 0} \lambda_k\,. \end{aligned} \tag{5.104}$$

Sumando ambas expresiones se obtiene

$$P_{\rm err} = \frac{1}{2}(1 - \sum_k |\lambda_k|) = \frac{1}{2}(1 - {\rm Tr}(|\Lambda|))\,. \tag{5.105}$$

[5.17] Se ha elegido poner el subespacio $\Lambda = 0$ (si existe) en Π_2, pero eso es irrelevante.

La simple estrategia de apostar siempre por el caso más probable, da una probabilidad de acierto máx(η_1,η_2) y de error mín(η_1,η_2). Por tanto, para el protocolo óptimo debe cumplirse [5.18]

$$P_{\text{err}} \leq \text{mín}(\eta_1,\eta_2) \leq \frac{1}{2}\,. \tag{5.106}$$

Todo esto es para estados mezcla cualesquiera. En el caso especial de que *se trate de dos estados puros* $|\psi_1\rangle$ y $|\psi_2\rangle$, todo ocurre en el espacio bidimensional subtendido por los dos estados. Λ es una matriz hermítica 2×2 y se puede calcular su espectro en forma explícita. Se obtiene la fórmula para el mínimo error:

$$P_{\text{err}} = \frac{1}{2}\left(1-\sqrt{1-4\eta_1\eta_2|\langle\psi_1|\psi_2\rangle|^2}\right)\,. \tag{5.107}$$

Cuando los estados son ortogonales $P_{\text{err}}=0$, y en efecto, se pueden identificar perfectamente con una medida proyectiva.

Cuando los estados son iguales $P_{\text{err}} = \frac{1}{2}\left(1-\sqrt{1-4\eta_1\eta_2}\right) = \text{mín}(\eta_1,\eta_2)$. [5.19]

Para estados mezcla ρ_1, ρ_2 cualesquiera puede probarse la relación

$$Q \geq 2P_{\text{err}}\,, \tag{5.108}$$

donde Q es la probabilidad de no identificación (pero no errar) con un protocolo óptimo, [5.20] y P_{err} la probabilidad de error también con un protocolo óptimo.

Para $N>2$ no es nada trivial obtener el POVM óptimo aunque hay resultados para casos particulares. Un protocolo no necesariamente óptimo pero empleado a menudo es tomar el POVM

$$\Pi_j = \eta_j\rho^{-1/2}\rho_j\rho^{-1/2}, \qquad \rho = \sum_j \eta_j\rho_j\,. \tag{5.109}$$

Este protocolo es óptimo cuando los estados son puros, equiprobables y además $\rho = \dfrac{1}{d}I$.

$$\eta_j = \frac{1}{N}, \qquad \rho_j^2 = \rho_j, \qquad \sum_j \rho_j = \frac{N}{d}I\,. \tag{5.110}$$

[5.18] Obviamente debe ser $P_{\text{err}} \leq \dfrac{1}{2}$, de otro modo sería mejor, al final del protocolo, intercambiar la identificación.

[5.19] Teniendo en cuenta que $\eta_1+\eta_2=1$.

[5.20] Nosotros sólo hemos estudiado en detalle Q mínimo para el caso de estados puros.

Entonces

$$\Pi_j = \frac{d}{N}\rho_j, \qquad \Gamma = \sum_j \eta_j \rho_j \Pi_j = \frac{1}{N}\frac{d}{N}\sum_j \rho_j = \frac{1}{N} I, \qquad \eta_j \rho_j = \frac{1}{N}\rho_j \leq \Gamma, \tag{5.111}$$

y en este caso el POVM es óptimo ya que satisface la condición de Helstrom. Justamente ésta es la estrategia que se ha aplicado al conjunto trino en el ejemplo discutido al principio.

6. Criptografía cuántica

6.1. Introducción

La comunicación cuántica es quizá la parte más desarrollada a nivel práctico de la información cuántica. Ya hemos visto dos ejemplos de aplicación, a saber, la codificación densa y la teleportación.

Dentro de comunicación cuántica, la criptografía cuántica es una de las áreas más avanzadas y prometedoras. La criptografía (clásica) está sujeta a la posibilidad de que se descubra el código secreto usado para la encriptación sin que los usuarios lo sepan. La gran ventaja de la criptografía cuántica es que en principio es invulnerable ante tales ataques, es decir, el código puede ser enviado por el canal cuántico [6.1] y no puede ser interceptado sin dejar un rastro que emisor y receptor pueden detectar. Una vez el código ha llegado de forma segura se puede enviar el mensaje codificado.

El tema de criptografía cuántica está muy desarrollado. Nosotros veremos algunas de las formas de codificación más estudiadas.

6.2. Claves de un uso

6.2.1. Claves clásicas de un uso

La criptografía clásica se ha usado para diferentes fines desde la antigüedad. Por ejemplo el código César, que se usaba para transmitir órdenes o información de tipo militar, consistía en cambiar cada letra por otra desplazada un cierto número de lugares en el alfabeto. [6.2] Así si lo desplazamos 3 lugares $a \to d$, $b \to e$, etc y por ejemplo *EQVVS* pasa a *HTZZX*.

El número de lugares desplazados, en este caso 3 es la **clave** de esta codificación, y cómo usar ese número es el **algoritmo de encriptación y desencriptación**.

El mensaje o texto a enviar es el **texto en abierto** que se codifica usando el algoritmo junto con la clave para producir el **texto encriptado**. Se supone que la clave y el texto en abierto son **secretos** (privados) y sólo los conocen el que envía el mensaje y el que lo recibe. [6.3] En este caso la clave

[6.1] Nota: En comunicación cuántica, la expresión "canal cuántico" puede referirse a enviar información usando un soporte cuántico, y no a "canal cuántico" en el sentido de superoperador (aunque ambos usos están relacionados).

[6.2] *ABCDEFGHIKLMNOPQRSTVXYZ*, faltan *JÑUW*.

[6.3] No es la única posibilidad. En el sistema de claves pública/privada, el receptor tiene una clave pública (un número

debe ser compartida entre emisor y receptor. El algoritmo mismo, así como el mensaje encriptado se consideran **públicos**, es decir, conocido por todo el mundo, el que envía, el que recibe y los posibles espías. (No es que los hagan públicos pero no pueden garantizar que se mantengan secretos frente a actividades de espionaje.) Por el mismo motivo, los canales de intercambio de mensajes (sean clásicos o cuánticos) también se deben considerar públicos.

Un código basado en desplazar letras no es seguro ya que hay un número pequeño de claves posibles (tantas como letras). Más generalmente, un código basado en cambiar cada letra por otra (u otro símbolo, pero siempre igual en todo el mensaje) no es nada seguro, porque las letras se van a poder identificar por su frecuencia típica (en el idioma en el que esté escrito el texto, si éste se desconoce no es difícil probar varios, todo depende del esfuerzo que se quiera dedicar). Especialmente si el mensaje es suficientemente largo.

Un método completamente seguro es por ejemplo el **cifrado de un solo uso** o **de Vernam**.

Los posibles mensajes a codificar (el texto en abierto) son o equivalen a cadenas de bits, 0's y 1's, de longitud menor o igual a n. Este n es público. Si el mensaje es más corto se añaden 0's hasta tener n bits. (Si el mensaje fuera más largo hay que trocearlo y enviarlo como varios mensajes.) La **clave** es una cadena **aleatoria** de n bits. Aleatoria quiere decir que está tomada al azar con igual probabilidad entre las $N = 2^n$ cadenas posibles. Equivalentemente, cada nuevo bit (hasta tener n) tiene probabilidad $1/2$ de ser 0 o 1 y además *no está correlacionado* con los anteriores.

Si $A = (a_1, \ldots, a_n)$ es el mensaje en abierto, $K = (k_1, \ldots, k_n)$ la clave y $C = (c_1, \ldots, c_n)$ el mensaje codificado, el algoritmo de codificación es

$$C = A \oplus K, \qquad c_i = a_i \oplus k_i, \tag{6.1}$$

y se descodifica mediante

$$A = C \oplus K, \qquad a_i = c_i \oplus k_i. \tag{6.2}$$

La operación $\oplus$ representa la suma módulo 2 bit a bit. Por ejemplo $(0110) \oplus (1100) = (1010)$.

Este método es demostrablemente seguro suponiendo:

1) Nadie más tiene la clave K (o acceso al método de generación de K)

2) K sólo se usa una vez para cada mensaje. La clave no se puede reutilizar. En otro caso se podría obtener información sobre K. Dado que $C \oplus A = K$, si el espía tiene acceso a C e información parcial sobre A (usualmente conoce el tema) tiene información parcial sobre K.

---

natural obtenido como el producto de varios números primos grandes) que el emisor utiliza para que codificar el mensaje. El receptor lo puede descodificar con su clave privada (la factorización del número).

Es relativamente obvio que $c_i = a_i \oplus k_i$ es un bit arbitrario al azar dado que k_i lo es. Más en detalle, que sea una encriptación segura quiere decir que

$$\text{Prob}(A|C) = \text{Prob}(A) \tag{6.3}$$

de modo que conocer C no proporciona ninguna información sobre A. (A no es un mensaje aleatorio, trata de un tema conocido y por tanto cada A tiene una cierta probabilidad $\text{Prob}(A)$ de ser el mensaje real.)

Equivalentemente $\text{Prob}(C|A) = \text{Prob}(C)$ (A y C son independientes). Esto se cumple en el método de Vernam:

$$\text{Prob}(C|A) = \sum_K \text{Prob}(C,K|A) = \sum_K \text{Prob}(C|K,A)\,\text{Prob}(K|A) = \sum_K \delta_{C,K\oplus A}\frac{1}{2^n} = \frac{1}{2^n} \tag{6.4}$$

Por tanto $\text{Prob}(C|A) = 1/2^n$ y en efecto no depende de A. [6.4]

Obviamente se pueden usar variantes del método para hacerlo más intrincado.

6.2.2. Claves cuánticas "de un uso"

Es instructivo reformular el protocolo completamente clásico de Vernam usando el formalismo cuántico. Basta considerar el caso $n = 1$ (un bit). El bit a es el mensaje abierto. El bit $a = 0$ se codifica en un qubit como el estado mezcla (puro) $\rho = |0\rangle\langle 0|$ y el bit $a = 1$ como $\rho = |1\rangle\langle 1|$. En el método de Vernam el bit clave k equivale a tirar una moneda y según salga cruz ($k = 1$) o cara ($k = 0$) se voltea el bit a o no. Voltear el bit es $0 \leftrightarrow 1$ y es aplicar el operador X sobre el qubit. Esto equivale al canal cuántico

$$\rho \to \rho' = T(\rho) = \frac{1}{2}\rho + \frac{1}{2}X\rho X. \tag{6.5}$$

Claramente

$$T(|0\rangle\langle 0|) = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}I, \qquad T(|1\rangle\langle 1|) = \frac{1}{2}I, \tag{6.6}$$

es decir,

$$\forall \rho = p_0|0\rangle\langle 0| + p_1|1\rangle\langle 1|, \qquad \rho' = \frac{1}{2}I. \tag{6.7}$$

[6.4] No sería correcto repetir el argumento para concluir $\text{Prob}(A|C) = \text{Prob}(A) = 1/2^n$, que no es cierto en general. La falacia está en que $\text{Prob}(K|A) = 1/2^n$ pero $\text{Prob}(K|C)$ no tiene que ser $1/2^n$; C da información sobre K si los A no son equiprobables. Por ejemplo si $A = (0\cdots 0)$ con seguridad, $K = C$ siempre.

Por tanto el estado ρ está totalmente enmascarado (encriptado) para el que no conozca la clave. Si el público mide ρ' se encontrará un valor 0 o 1 bien definido, pero no sabrá qué significa ese valor y la medida no aportará información sobre ρ.

Es fácil ahora extender el protocolo de Vernam para encriptar qubits (en vez de bits). [6.5] De nuevo basta considerar el caso de un solo qubit ($n = 1$). El estado del qubit es

$$\rho = \frac{1}{2}(1 + \boldsymbol{u}\cdot\boldsymbol{\sigma}), \qquad \|\boldsymbol{u}\| \leq 1. \tag{6.8}$$

(El caso anterior, de un bit, corresponde a $\boldsymbol{u} = (0,0,u_z)$ siendo $u_z = p_0 - p_1$.)

Ahora generamos una clave aleatoria de dos bits $K = (k_1 k_2)$ (hay que tirar dos veces la moneda). Primero, se aplica o no Z, según que $k_1 = 1$ o $k_1 = 0$, y a continuación se aplica o no X, según que $k_2 = 1$ o $k_2 = 0$. [6.6]

$$\begin{aligned} \boldsymbol{u} = (u_x, u_y, u_z) \quad &\underset{(0,0)=I}{\longrightarrow} \quad (+u_x, +u_y, +u_z) \\ &\underset{(0,1)=X}{\longrightarrow} \quad (+u_x, -u_y, -u_z) \\ &\underset{(1,0)=Z}{\longrightarrow} \quad (-u_x, -u_y, +u_z) \\ &\underset{(1,1)=XZ}{\longrightarrow} \quad (-u_x, +u_y, -u_z) \end{aligned} \tag{6.9}$$

Al combinar las cuatro opciones con el mismo peso se obtiene $\boldsymbol{u}' = 0$ y resulta $\rho' = \frac{1}{2}I$. El estado ρ del qubit queda totalmente enmascarado para los que ignoren K. Pueden medir el qubit pero no sabrán como interpretar el resultado para inferir ρ.

6.3. Protocolo B92 de distribución de claves cuánticas

El escenario más usual es que Andrea quiere enviar a Benito un mensaje, que equivale a una cadena de bits. El método de Vernam es seguro siempre que la clave permanezca siendo privada y se utilice sólo una vez, entonces Andrea codifica el mensaje y envía públicamente el mensaje codificado por el canal clásico. El principal problema a resolver es que Andrea y Benito compartan la clave de manera segura. Una opción es que ambos tengan una copia de las claves a usar y la mantengan a buen recaudo (pero nunca se puede estar seguro de que algún espía no haya tenido acceso a la libreta de claves).

[6.5] P. O. Boykin, V. Roychowdhury, *Optimal encryption of quantum bits,* Phys. Rev. A **67** (2003), 042317. [7] .

[6.6] En realidad $XZ = -ZX$ y se obtiene el mismo resultado si primero es Z y luego X o al revés.

Usar el canal cuántico ofrece una posibilidad de compartir claves de forma segura, dado que cualquier interferencia se notaría. Si hay interferencia se repite el envío de un nuevo código hasta que sea seguro y entonces se puede enviar el mensaje codificado por el canal clásico.

Es importante notar que no se envía el secreto mismo por el canal cuántico sino una clave aleatoria. El motivo es que la seguridad se basa en que a veces hay que descartar qubit enviados y volver a enviar otros. Si son aleatorios para establecer una clave eso no tiene problema (cambian aleatoriamente en cada intento) pero si fueran parte del mensaje mismo y hubiera que enviar el mismo qubit repetidamente el secreto se vería comprometido.

B92 (Bennet 1992) es un protocolo para compartir una clave, que es una cadena de bits.

1) Andrea genera una secuencia de bits aleatorios (que deben permanecer secretos). [6.7] Codifica un bit 0 mediante un qubit en el estado $|0\rangle$ y 1 en el estado $|+\rangle$. Se tiene entonces una cadena aleatoria de qubits que es enviada a Benito.

2) Benito aplica el protocolo de identificación inequívoca de qubits (Tema 5.4.1.3) sobre cada qubit recibido. La probabilidad de identificación (ec. (5.87)) de cada qubit es $1 - Q = 1 - 1/\sqrt{2} = 0.29$

   A partir de este punto ya sólo se necesita usar comunicación clásica.

3) Benito comunica a Andrea qué qubits fueron correctamente identificados (por ej. los qubits número 1, 4 y 7) pero no cuáles fueron los resultados ($|0\rangle$ ó $|+\rangle$). Los resultados los conoce Andrea, pero no el público. Entonces Andrea y Benito guardan y usan en su clave sólo esos bits que ambos conocen y descartan el resto. Estos bits compartidos pero privados son lo que se conoce como **clave en basto**. (En basto porque puede ser sometida a posterior refinamiento frente a espionaje o errores.)

Supongamos que hay una espía, Eva, que intercepta los qubits enviados a Benito. Si Eva también emplea el método de **identificación inequívoca**, un 29 % de las veces identificará el estado (y lo reenviará a Benito que no notará ninguna diferencia) pero un 71 % de las veces no sabrá qué estado es y simplemente enviará al azar $|0\rangle$ o $|+\rangle$ a Benito (al 50 %). Es decir, Benito recibirá un bit distinto de que debería un 35 % de la veces. Para tener esto en cuenta se puede aplicar el siguiente protocolo adicional:

[6.7] Por ejemplo, si utiliza un generador de números pseudoaleatorios, la semilla debe ser secreta, de otro modo se podría reproducir la secuencia de bits supuestamente aleatorios.

4′) La presencia de Eva se puede detectar si Andrea y Benito se comunican un cierto número de bits de la clave que deberían ser iguales para los dos. Si hay una discrepancia del orden del 35 % o más, no se puede descartar que la clave esté comprometida, entonces es clave se desecha y se repite el proceso.

Sin embargo Eva lo puede hacer mejor. Si en vez de identificación inequívoca Eva usa **identificación con mínimo error** (ec. (5.107)) su identificación será incorrecta $\frac{1}{2}(1-\frac{1}{\sqrt{2}}) = 15\,\%$ de las veces, y acertará 85 % de las veces. Benito recibirá un bit incorrecto sólo el 15 % de las veces (mucho menos que antes, 4′) ya no sirve) y Eva conocerá el 85 % de la clave.

En las transmisiones hay errores inevitables. Si la proporción de error supera el 15 % el protocolo B92 no es viable. Si es menor, la presencia de Eva se puede tener en cuenta, usando, en vez de 4′)

4) Andrea y Benito anuncian públicamente un cierto número de bits de la clave que deberían ser iguales para los dos. Si hay una discrepancia del orden del 15 % o más, no se puede descartar que la clave esté comprometida. En ese caso se desecha y se repite el proceso.

Por supuesto los bits que Andrea y Benito se comunican para hacer la comprobación se descartan y no forman parte de la clave compartida final, ya que no hay forma de garantizar que sean secretos. Si por ejemplo se quiere una clave final de 100 bits y se van a necesitar 50 para comprobación, hay que generar una clave de 150 bits.

6.4. Protocolo BB84

BB84 (Bennet y Brassard, 1984) es el protocolo más antiguo y conocido para establecer una clave compartida segura.

1) Andrea elige al azar con igual probabilidad entre los cuatro estados de un qubit $|0\rangle$, $|1\rangle$, $|+\rangle$ y $|-\rangle$ y envía ese qubit a Benito. [6.8] El convenio es que los estados $|0\rangle$ y $|+\rangle$ codifican un bit 0 mientras que $|1\rangle$ y $|-\rangle$ codifican un bit 1 de la clave.

[6.8] Equivalentemente, Andrea elige al 50 % entre las dos bases $\{|0\rangle, |1\rangle\}$ (base Z) y $\{|+\rangle, |-\rangle\}$ (base X) y luego al 50 % entre el primer o segundo elemento de la base.

bit	Z	X
0	$\|0\rangle$	$\|+\rangle$
1	$\|1\rangle$	$\|-\rangle$

2) Benito elige al azar con igual probabilidad si medir el qubit recibido en la base $\{|0\rangle, |1\rangle\}$ o en la base $\{|+\rangle, |-\rangle\}$. Si ha elegido la misma base que Andrea obtendrá como resultado el mismo qubit que le enviaron. En caso contrario obtendrá un resultado de su base al azar con probabilidad 50 %. (Por ejemplo, Andrea envía $|0\rangle$ y Benito mide en la base $\{|0\rangle, |1\rangle\}$, entonces Benito encontrará $|0\rangle$ con seguridad, si en cambio mide en la base $\{|+\rangle, |-\rangle\}$ obtendrá uno de estos dos estados con igual probabilidad.)

3) *Después de hacer las medidas*, Benito comunica a Andrea en qué base midió cada qubit (pero no el resultado de la medida), y Andrea le dice para qué qubits coinciden las dos bases. Los qubits en los que las bases son distintas se descartan. Entonces Andrea y Benito tienen una clave compartida privada formada por los bits correspondientes a los qubits en los que los dos eligieron la misma base. Sólo ellos saben el estado de cada qubit y por tanto del bit.

qubit	1	2	3	4	5	$\cdots$
A	$\|0\rangle$	$\|-\rangle$	$\|0\rangle$	$\|1\rangle$	$\|+\rangle$	$\cdots$
B	Z	Z	X	Z	X	$\cdots$
	$\|0\rangle$	—	—	$\|1\rangle$	$\|+\rangle$	$\cdots$
bit	0	—	—	1	0	$\cdots$

Tabla 1: En esta tirada se utilizan los qubits 1, 4, 5, ... y se descartan los qubits 2, 3, ...

Veamos el efecto de un espía. Podemos suponer que Andrea y Benito usan la misma base. [6.9] Eva intercepta el qubit enviado por Andrea, lo mide en alguna de las dos bases [6.10] y reenvía el resultado a Benito, que desconoce la presencia de Eva (cree que es el qubit enviado por Andrea).

Inevitablemente, en promedio, Eva va a medir la mitad de las veces en la base correcta y la otra mitad en la incorrecta (y sabrá cuál es después, cuando Benito lo comunique a Andrea).

Si Eva mide en la base correcta obtendrá el resultado correcto con toda seguridad y por tanto Benito recibirá el mismo estado que le envío Andrea. Si Eva mide en la base incorrecta obtendrá

[6.9] Cuando mide el qubit Benito no sabe cuál es la base que ha usado Andrea, pero luego Andrea y Benito se comunican y sólo se quedan con los qubits en los que sus bases coinciden. Esos qubits son los que consideramos aquí.

[6.10] Podría decidir medir en otras bases pero eso sólo puede aumentar la aleatoriedad y por tanto que se note más su presencia.

un estado de esa base al azar y a su vez Benito obtendrá un estado de la base correcta pero al azar. (Por ejemplo, Andrea envía $|0\rangle$, Eva mide $|\pm\rangle$ y Benito obtiene $|0\rangle$ ó $|1\rangle$ con igual probabilidad.) En conjunto Benito tendrá el resultado correcto con probabilidad $\frac{1}{2}+\frac{1}{2}\times\frac{1}{2}=\frac{3}{4}$ y el incorrecto con probabilidad $\frac{1}{4}$.

De nuevo, Andrea y Benito pueden detectar ese 25 % de errores comparando públicamente parte de los bits (que por supuesto no se usan en la clave final) y descartar la clave si ha habido interferencias. (Puesto que el error mínimo que inevitablemente introduce Eva en este protocolo es mayor que en B92, y por tanto es más fácil de detectar, en principio BB84 es preferible a B92. Además aprovecha el 50 % de los qubits enviados en vez de sólo el 29 %).

Se pueden considerar ataques más sofisticados pero la conclusión es que cuánta más información gane Eva sobre la clave menos información exacta recibe Benito, y por tanto aumenta el número de errores que delatan la presencia de un espía. No puede haber una transferencia neta de información a Eva sin efecto sobre la información que recibe Benito (la información cuántica no se puede clonar).

Por ejemplo, supongamos que Eva intercepta un estado ρ_A enviado por Andrea, y lo procesa para extraer información, pero de tal modo que el estado que luego reenvía a Benito es el mismo estado ρ_A, para no introducir ninguna distorsión en lo que recibe Benito. Lo más general es tener un canal cuántico

$$\rho_A \to \rho_{AE} = T(\rho_A)\,. \tag{6.10}$$

Aquí ρ_A está en $\mathscr{H}_A$ y ρ_{AE} en $\mathscr{H}_A \otimes \mathscr{H}_E$. Lo que Eva reenvía a Benito, y lo que le queda es, respectivamente

$$\rho_B = \mathrm{Tr}_E(\rho_{AE}), \qquad \rho_E = \mathrm{Tr}_A(\rho_{AE}). \tag{6.11}$$

Si ahora se postula que $\rho_B = \rho_A$ siempre (es decir, $\forall\rho_A$), no es difícil probar (ver apéndice 6.7) que necesariamente $\rho_{AE} = \rho_A \otimes \rho_E$, donde ρ_E es una matriz densidad fija, independiente de ρ_A. Eso implica que Eva no obtiene ninguna información sobre ρ_A.

En la misma línea (Eva no quiere permitir ninguna distorsión en lo que reenvía a Benito) pero restringiéndonos a estados concretos (en vez $\forall\rho_A$). Los estados interceptables pueden ser $|0\rangle$ y $|1\rangle$ (en realidad dos estados ortogonales cualesquiera) y también

$$|\phi\rangle = \alpha|0\rangle + \beta|1\rangle, \qquad |\phi^\perp\rangle = -\beta^*|0\rangle + \alpha^*|1\rangle\,, \qquad \alpha\beta \neq 0\,. \tag{6.12}$$

(La condición $\alpha\beta \neq 0$ indica que $\{|\phi\rangle, |\phi^\perp\rangle\}$ no son los mismos estados que $\{|0\rangle, |1\rangle\}$.

Lo que Eva puede hacer es manipular los estados mediante un operador unitario en $\mathscr{H}_A \otimes \mathscr{H}_E$

(añadiendo estados auxiliares).

$$U|i\rangle_A \otimes |0\rangle_E = \sum_j |j\rangle_A \otimes |\xi_{ij}\rangle_E, \tag{6.13}$$

pero para no introducir error sólo debe haber amplitud saliente para $j = i$

$$U|i\rangle_A \otimes |0\rangle_E = |i\rangle_A \otimes |\xi_i\rangle_E \tag{6.14}$$

Tampoco se quiere introducir error si llega $|\phi\rangle_A$,

$$\begin{aligned} 0 &= {}_A\langle\phi^\perp|U|\phi\rangle_A \otimes |0\rangle_E = (-\beta\langle 0| + \alpha\langle 1|)(\alpha|0\rangle \otimes |\xi_0\rangle + \beta|1\rangle \otimes |\xi_1\rangle \\ &= -\alpha\beta(|\xi_0\rangle - |\xi_1\rangle) \implies |\xi_0\rangle = |\xi_1\rangle. \end{aligned} \tag{6.15}$$

Es decir, $U|\psi\rangle_A \otimes |0\rangle_E = |\psi\rangle_A \otimes |\xi\rangle_E$. El resultado no está entrelazado y no hay ninguna transferencia de información al sector E, no se aprende nada sobre el estado $|\psi\rangle_A$.

6.5. Protocolo E91

En el protocolo E91 (Ekert 1991) se utiliza entrelazamiento en vez del envío de qubits de Andrea a Benito para establecer una clave segura compartida.

Supongamos que Andrea y Benito disponen de un acopio de ebits que comparten, por ejemplo en el estado de Bell $|\Psi_+\rangle_{AB}$. Para establecer una clave compartida simplemente miden cada uno su qubit en la base computacional. Al hacer la medida los dos obtendrán el mismo valor $|0\rangle$ ó $|1\rangle$ lo cual establece una clave compartida. Más generalmente, una tercera parte, Carlos, puede ser el servidor de ebits que envía a Andrea y Benito conforme los vayan necesitando.

Si realmente Andrea y Benito comparten el estado $|\Psi_+\rangle$, ninguna interferencia (actividad de espionaje) puede afectarles, ya que no hay ningún qubit viajando entre ellos. Sin embargo, tal y como está, este protocolo es vulnerable: en cualquiera de las dos versiones los qubits del par tienen que haber sido entrelazados juntos y al menos uno de ellos tiene que haber viajado. Si Eva intercepta los qubits, puede enviarles al azar $|00\rangle$ y $|11\rangle$ sin que Andrea y Benito lo detecten ya que no verán nada raro: reciben 0 y 1 al 50 % y además si usan algunos bits como comprobación verán que sus bits siempre coinciden, si Andrea mide $|0\rangle$ o $|1\rangle$ Benito también. Haciendo medidas exclusivamente en la base computacional no es posible distinguir el estado puro entrelazado $|\Psi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ del estado mezcla separable $\frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11|$.

El remedio es que antes de usar los ebits Andrea y Benito comprueben que los estados están correctamente entrelazados *mediante desigualdades de Bell*. Gastando parte de los pares compartidos, pueden hacer medidas en direcciones elegidas independientemente y comprobar que $\langle S \rangle$ tiene el valor correcto. Si no es así deben descartar los supuestos ebits. (Por otro lado este tipo de comprobación es necesaria siempre, no sólo en este protocolo, dado un acopio de ebits, o ebits remitidos por un servidor, para ver que no se han deteriorado o han sido manipulados.)

6.6. Compartición cuántica de secretos

Andrea quiere compartir un secreto suyo (por ejemplo la clave de una caja fuerte) con Benito y Carlos, de tal modo que cada uno de los dos tenga información parcial y sólo colaborando los dos accedan al secreto. De este modo, si al menos uno de los dos es honesto el secreto está a salvo.

Supongamos que el secreto es una cadena de bits A. Clásicamente un método es que Andrea genere una clave aleatoria K. A Benito le da $A \oplus K$ y a Carlos le da K. Una versión segura basada en criptografía cuántica es que Andrea use el protocolo BB84 para establecer claves K_1 con Benito y K_2 con Carlos. Entonces Andrea codifica su secreto con la clave $K = K_1 \oplus K_2$. El secreto codificado $C = A \oplus K$ es conocido por Benito y Carlos, si colaboran pueden obtener K y reconstruir A.

Otro enfoque es el basado en entrelazamiento.

1) Andrea quiere compartir bits para establecer una clave. Codifica 0 ó 1 respectivamente como

$$|\Psi_0\rangle = \cos(\theta/2)|00\rangle + \mathrm{sen}(\theta/2)|11\rangle, \qquad |\Psi_1\rangle = \cos(\theta/2)|00\rangle - \mathrm{sen}(\theta/2)|11\rangle, \tag{6.16}$$

   y envía uno de los qubits a Benito y el otro a Carlos.

2) Benito mide su qubit en la base X. Si obtiene $|\pm\rangle$ sabe que Carlos tiene $|\xi_\pm\rangle$ cuando el bit secreto es 0 y $|\xi_\mp\rangle$ cuando es 1,

$$\begin{gathered} |\xi_\pm\rangle := \cos(\theta/2)|0\rangle \pm \mathrm{sen}(\theta/2)|1\rangle, \\ {}_B\langle\pm|\Psi_0\rangle = |\xi_\pm\rangle_C, \qquad {}_B\langle\pm|\Psi_1\rangle = |\xi_\mp\rangle_C. \end{gathered} \tag{6.17}$$

3) Carlos aplica el método de identificación inequívoca a su qubit para ver si es $|\xi_+\rangle$ o $|\xi_-\rangle$. Con probabilidad $|\cos(\theta)|$ no habrá identificación. En ese caso se repite el proceso (Andrea vuelve a enviar dos qubits). Si hay identificación Benito y Carlos tienen dos bits correlacionados, pero no saben cómo. Sólo colaborando pueden reconstruir la clave que tiene Andrea.

		B	
	C	$\|+\rangle$	$\|-\rangle$
A	$\|\Psi_0\rangle$	$\|\xi_+\rangle$	$\|\xi_-\rangle$
	$\|\Psi_1\rangle$	$\|\xi_-\rangle$	$\|\xi_+\rangle$

Si Eva intercepta los qubits puede aplicar identificación con mínimo error y reenviar el resultado a Benito y Carlos, pero eso deja un rastro que Andrea, Benito y Carlos pueden detectar comparando

parte de sus bits. Obsérvese que la eficiencia en la identificación $(1-|\cos(\theta)|)$ es máxima cuando $|\Psi_0\rangle$ y $|\Psi_1\rangle$ son ortogonales. Pero entonces también Eva puede medir perfectamente el estado entrelazado y sin dejar rastro. Por lo tanto hay que tomar un valor de compromiso (entre eficiencia y seguridad) para θ. Típicamente $\theta = \pi/4$.

6.7. Apéndice: No replicación de la información

Al hilo de (6.10), Eva intercepta ρ_A y lo procesa añadiendo estados auxiliares E

$$\rho_A \to \rho_{AE} = T(\rho_A) = \sum_\mu A_\mu \rho_A A_\mu^\dagger, \qquad \sum_\mu A_\mu^\dagger A_\mu = I_A\,. \tag{6.18}$$

Aquí $A_\mu : \mathscr{H}_A \to \mathscr{H}_A \otimes \mathscr{H}_E$. Lo que Eva reenvía a Benito es

$$\rho_B = \mathrm{Tr}_E(\rho_{AE}). \tag{6.19}$$

Si $\{|e\rangle\}$ es una base ortonormal de $\mathscr{H}_E$,

$$\rho_B = \sum_\mu \sum_e {}_E\langle e|A_\mu \rho_A A_\mu^\dagger |e\rangle_E = \sum_{\mu,e} B_{\mu e}\rho_A B_{\mu e}^\dagger\,, \qquad B_{\mu e} := {}_E\langle e|A_\mu\,. \tag{6.20}$$

$B_{\mu e}$ son operadores en $\mathscr{H}_A$. Si se postula que $\rho_B = \rho_A\ \forall \rho_A$ el canal $\mathrm{Tr}_E(T(\cdot))$ es la identidad en A, por tanto deben ser una rotación unitaria de operadores de Kraus $(I_A, 0, 0, \ldots)$, es decir,

$$B_{\mu e} = u_{\mu e} I_A, \qquad \sum_{\mu,e} |u_{\mu e}|^2 = 1\,. \tag{6.21}$$

Esto implica

$$A_\mu = \sum_e u_{\mu e} I_A \otimes |e\rangle_E = I_A \otimes |\tilde{u}_\mu\rangle_E, \qquad \sum_\mu \||\tilde{u}_\mu\rangle_E\|^2 = 1\,. \tag{6.22}$$

Entonces

$$T(\rho_A) = \rho_A \otimes \rho_E, \qquad \rho_E = \sum_\mu |\tilde{u}_\mu\rangle\langle\tilde{u}_\mu|_E\,. \tag{6.23}$$

Como ρ_E no depende de ρ_A Eva no extrae ninguna información.

7. Algoritmos cuánticos

7.0.1. Circuitos clásicos y cuánticos

Hacer cálculos con un **circuito clásico** (que podemos suponer utiliza bits) requiere combinar puertas lógicas, así como usar bits auxiliares (o de trabajo) y copiar bits. Esto lo hace la denominada puerta FANOUT,

$$\text{FANOUT}(u) = (u,u) \qquad u \in \{0,1\} \tag{7.1}$$

Todos los circuitos clásicos se pueden hacer con unas pocas puertas lógicas universales, tales como AND, OR, NOT, de 2, 2 y 1 bit, respectivamente. La relación booleana $\neg(u \vee v) = \neg u \wedge \neg v$ indica que OR no es imprescindible. De hecho todo se puede hacer con la puerta NAND :

$$\text{NAND}(u,v) = \neg(u \wedge v) = 1 \oplus uv, \qquad u,v \in \{0,1\} \tag{7.2}$$

usando bits auxiliares. Así por ejemplo [7.1]

$$\text{NOT}(u) = \text{NAND}(u,1), \qquad \text{AND}(u,v) = \text{NOT}(\text{NAND}(u,v))\,. \tag{7.3}$$

La puerta NOT es **reversible**, en cambio otras puertas clásicas, tales como AND, NAND y OR, no lo son. Es posible hacer toda la computación clásica con puertas reversibles, lo cual implica que no hay borrado de información. De ese modo se puede sortear el **principio de Landauer**: borrar un bit a una temperatura absoluta T requiere gastar al menos una energía $kT\ln(2)$, que se disipa al ambiente. [7.2] La reversibilidad es imprescindible en el caso cuántico.

Clásicamente las puertas reversibles de 1 o 2 bits no bastan para formar todas las demás, pero sí lo hace la puerta C^2-NOT o $C^2(X)$ o TOFFOLI$(u,v,w) = (u,v,w \oplus uv)$, que es claramente reversible (de hecho coincide con su inverso).

Esta puerta es suficiente para reproducir NAND y FANOUT, que son las dos puertas universales clásicas,

$$\begin{aligned} &\text{TOFFOLI}(u,v,1) = (u,v,1 \oplus uv) = (u,v,\text{NAND}(u,v)\,), \\ &\text{TOFFOLI}(u,1,0) = (u,1,0 \oplus u) = (u,1,u)\,. \end{aligned} \tag{7.4}$$

[7.1] El bit 1 se puede obtener a partir del 0 (que es el inicio estándar) mediante NAND$(0,0)$.

[7.2] Con las configuraciones de ordenadores actuales, el gasto por resistencia óhmica es mucho mayor que el mínimo impuesto por el principio de Landauer.

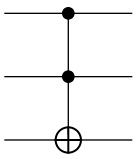


Figura 7.1: Puerta de Toffoli.

Todas las puertas clásicas reversibles tienen automáticamente su versión cuántica (son unitarias) identificando un bit $x = 0, 1$ con un estado $|x\rangle$ de la base computacional. Así por ejemplo NOT corresponde a $X = \sigma_x$.

Una puerta del tipo $x \to y = f(x)$ (aquí x, y pueden ser uno o más bits) es posible clásicamente pero en general no es reversible y por tanto no es realizable cuánticamente. Sin embargo toda función $f : A \to G$, donde A es un conjunto cualquiera y G tiene estructura de grupo, se puede reformular de manera **reversible** mediante la construcción

$$U_f : A \times G \to A \times G, \qquad U_f(x, y) = (x, f(x)y^{-1}) . \tag{7.5}$$

Claramente la aplicación $f \to U_f$ es inyectiva (U_f identifica unívocamente a f). La función U_f es invertible, de hecho cuando el grupo es abeliano $U_f^{-1} = U_f$. En particular para bits, si $f : \mathbb{Z}_2^n \to \mathbb{Z}_2^m$, $y = f(x)$, (x tiene n bits, y tiene m bits) la función reversible asociada U_f es $(x, y) \mapsto (x, f(x) \oplus y)$. [7.3] Se suele denominar el **oráculo** (que devuelve el valor de $f(x)$ cuando se invoca) o la **caja negra**. La puerta U_f tiene la forma

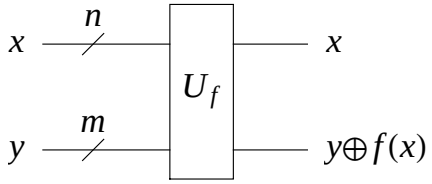


Figura 7.2: Oráculo de una función $f : \mathbb{Z}_2^n \to \mathbb{Z}_2^m$.

Automáticamente U_f define una puerta cuántica de $n+m$ qubits. Es una puerta f-CNOT, siendo x los bits o qubits de control (a través del valor de $f(x)$) e y los controlados,

$$U_f|x\rangle \otimes |y\rangle = |x\rangle \otimes |y \oplus f(x)\rangle \tag{7.6}$$

[7.3] Si $x, y \in \mathbb{Z}_2^n$, $x \oplus y$ es la suma bit a bit módulo 2. Equivale a decir $x + y$ considerando $\mathbb{Z}_2^n$ como un espacio vectorial sobre el cuerpo $\mathbb{Z}_2$.

Cuánticamente hacer un circuito implica construir un operador U unitario $N \times N$, $U \in \mathrm{SU}(N)$. Sin pérdida de generalidad se puede suponer $N = 2^n$, siendo n el número de qubits que forman la base computacional (digamos qubits computacionales). Se puede demostrar que cualquier U se puede construir componiendo un número finito de rotaciones bidimensionales involucrando dos qubits computacionales cada vez. Y a su vez, estas rotaciones de $\mathrm{SU}(2)$ siempre se pueden construir con puertas de 1 qubit computacional y puertas CNOT sobre dos qubits computacionales, ambos en número finito. [7.4]

En consecuencia, con puertas de un qubit y CNOT se puede construir cualquier circuito cuántico. No hay contradicción con lo dicho anteriormente de que clásicamente no bastan puertas reversibles de 1 y 2 bits, porque para construir el circuito con CNOT se necesitan puertas de un qubit sin versión clásica, tal como Hadamard, $H = (Z+X)/\sqrt{2}$, es decir se hace uso de la superposición cuántica. Un circuito genérico requiere $O(N^2)$ puertas CNOT.

Las puertas de un qubit referidas forman un continuo. Si sólo se quiere usar un conjunto finito fijo de puertas universales, cualquier U se puede aproximar tanto como se quiera (tolerando un error máximo ε). Un tal conjunto de puertas universales es por ejemplo $\{H, T, \mathrm{CNOT}\}$ [7.5]

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}, \quad \mathrm{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{7.7}$$

(T también se denomina puerta $\pi/8$.) El número de puertas requerido en el circuito, en el peor de los casos (para U genérica) crece como $O\left(N\dfrac{\log(1/\varepsilon)}{\log(\log(N))}\right)$. [7.6] Es un crecimiento exponencial en el número de qubits n.

Es importante notar que la terminología "circuitos" y "puertas" puede llevar a confusión, ya que sugiere que los qubits circularían espacialmente pasando por ciertos dispositivos físicos que serían las puertas. Eso es así en el caso de fotones pasando por cristales varios, pero no es la situación estándar. El circuito debe entenderse en sentido temporal, no espacial. Lo usual es que los qubits estén codificados en un sistema físico, por ejemplo iones en una trampa, y sobre ese sistema las puertas se realizan aplicando distintas acciones físicas, tales como campos magnéticos, láseres, etc.

[7.4] Nielsen & Chuang, pág. 191.

[7.5] Por comodidad suele añadirse la puerta PHASE o $\pi/4$ $S = T^2 = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$, aunque es redundante.

[7.6] Este el óptimo teórico, usando construcciones sistemáticas se requiere un número de puertas $O(N^2 \log^2(N) \log^c(N^2 \log^2(N)/\varepsilon))$, siendo c aproximadamente 2.

Más que pasar los qubits por las puertas, son las puertas las que pasan por los qubits, no es esencial que éstos se desplacen espacialmente.

De ese modo un circuito puede tener una cantidad muy grande de puertas, que simplemente significa que se aplica un láser (u otra acción) de varias formas repetidamente, sin que ello requiera una cantidad ingente de copias de unos dispositivos físicos (las puertas) para materializar el circuito lógico.

Por tanto los qubits son materiales pero las puertas son más bien virtuales. Las puertas son transformaciones que "leen" y "escriben" un estado u otro sobre el sistema de qubits, que desde este punto de vista forma un sistema de registros. En ese sentido la codificación del circuito y sus puertas se hace a nivel clásico. *Cuántico* es que a varias opciones se les asigna amplitudes de probabilidad, *clásico* que se les asigna probabilidades o certezas. En el circuito sólo el estado de los qubits es cuántico, lo demás es clásico. Un ordenador clásico ejecuta un programa y lee (o escribe) de una base de datos, mandando señales que se traducen en acciones de dispositivos (ej. láser) que actúan físicamente sobre los qubits. [7.7]

Debe comentarse que aparte de circuitos, se puede hacer computación cuántica por otros medios. Por ejemplo, dado un hamiltoniano $H(\vec{\lambda})$ parametrizado, un cambio adiabático (es decir, muy lento) en los parámetros, $\vec{\lambda}_1 \to \vec{\lambda}_2$, permite pasar del estado fundamental $|\Psi_0(\vec{\lambda}_1)\rangle$ a $|\Psi_0(\vec{\lambda}_2)\rangle$. Es una transformación unitaria que constituye un tipo de computación cuántica. La evolución del estado $|\Psi(t)\rangle$ se puede hacer eficientemente con circuitos cuánticos para $H = \sum_k H_k$ si los términos H_k se pueden implementar de modo eficiente en los circuitos y el número de términos crece a lo sumo polinómicamente.

7.1. Algoritmo de Deutsch-Jozsa

El algoritmo de Deutsch-Jozsa generaliza el de Deutsch. El problema que resuelve es el siguiente, dada una función $f : \{0,1\}^n \to \{0,1\}$, se sabe que o bien f es **constante** (todas las imágenes son iguales) o f es **equilibrada** (la mitad de imágenes son 0 y la otra mitad 1). Se trata de determinar sin posibilidad de error a cuál de las dos clases pertenece f.

Hay $N = 2^n$ posibles $x \in \{0,1\}^n$. Clásicamente hay que evaluar la función $N/2+1$ veces (en el peor de los casos) para determinar si es constante o equilibrada. El algoritmo cuántico de Deutsch-

[7.7] Esas acciones no son sólo sobre un qubit, se requiere también inducir interacción entre pares de qubits para producir una puerta CNOT.

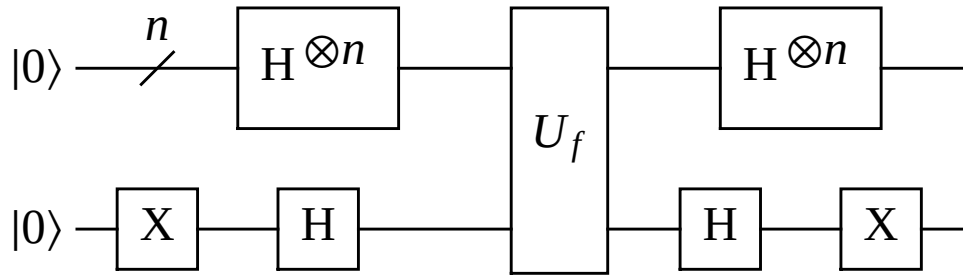


Figura 7.3: Circuito para el algoritmo de Deutsch-Jozsa.

Jozsa lo determina con *una sola llamada al oráculo* U_f. Técnicamente es una mejora exponencial sobre la algoritmo clásico.

Se puede hacer con el circuito de la Fig. 7.3. U_f actúa sobre $n+1$ qubits, los n primeros controlan, a través del valor de $f(x)$, al qubit $n+1$-ésimo.

$x \in \{0,1\}^n$ se puede ver como un vector en el espacio vectorial $\mathbb{Z}_2^n$ (sobre el cuerpo $\mathbb{Z}_2$). También se puede ver como un número entero x, $0 \leq x \leq N-1$ tal que $x = x_1 x_2 \ldots x_n$ donde los $x_j \in \{0,1\}$ son sus dígitos en binario ($x = 2^{n-1}x_1 + \cdots 2x_{n-1} + x_n$). A x le corresponde un estado de la base computacional

$$|x\rangle = |x_1 \ldots x_n\rangle = |x_1\rangle \otimes \cdots \otimes |x_n\rangle = \bigotimes_{j=1}^{n} |x_j\rangle_j, \qquad x \in \mathbb{Z}_2^n, \tag{7.8}$$

del espacio de n-qubits $(\mathbb{C}^2)^{\otimes n} = \mathbb{C}^N$, con dimensión $N = 2^n$.

La acción de X y luego H sobre el qubit controlado es

$$HX|0\rangle = H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = |-\rangle\,. \tag{7.9}$$

Se podía haber puesto directamente el qubit controlado en estado $|-\rangle$ pero es más usual inicializar los circuitos con estados $|0\rangle$ de la base computacional y que sea el circuito el que produzca los estados necesarios. Igualmente las últimas puertas H y X no son realmente necesarias en nuestro caso, pero es usual dejar los qubits auxiliares otra vez en estado $|0\rangle$ para reutilizarlos.

Evaluemos la acción de $H^{\otimes n}$ sobre los n qubits de control. La acción de H es $H|0\rangle = |+\rangle$, $H|1\rangle = |-\rangle$, y se puede expresar como

$$u \in \{0,1\} \qquad H|u\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + (-1)^u|1\rangle\right). \tag{7.10}$$

Entonces, $H^{\otimes n}$ sobre un estado genérico de la base computacional produce

$$\begin{aligned} x \in \mathbb{Z}_2^n \qquad H^{\otimes n}|x\rangle &= \bigotimes_{j=1}^{n} H|x_j\rangle_j = \left(\frac{1}{\sqrt{2}}\right)^n \bigotimes_{j=1}^{n} \left(|0\rangle_j + (-1)^{x_j}|1\rangle_j\right) \\ &= \frac{1}{2^{n/2}} \sum_{z=0}^{N-1} \left(\prod_{j=1}^{n} (-1)^{x_j z_j}\right) |z_1 \dots z_n\rangle \end{aligned} \tag{7.11}$$

Cada $|z_j\rangle_j$ aporta un factor (-1) si y sólo si $z_j = x_j = 1$.

El resultado se puede reescribir como

$$H^{\otimes n}|x\rangle = \frac{1}{\sqrt{N}} \sum_{z=0}^{N-1} (-1)^{x\cdot z}|z\rangle \tag{7.12}$$

donde se ha definido

$$\forall x, z \in \mathbb{Z}_2^n \qquad x \cdot z \equiv \sum_{j=1}^{n} x_j z_j \pmod 2. \tag{7.13}$$

No es más que el producto escalar usual pero en el espacio vectorial $\mathbb{Z}_2^n$. También se ha usado la propiedad

$$\prod_{j=1}^{n} (-1)^{x_j z_j} = (-1)^{\sum_{j=1}^{n} x_j z_j} = (-1)^{x\cdot z}. \tag{7.14}$$

Para el circuito en consideración, y para los n qubits de control después aplicar las puertas de Hadamard se tiene el estado

$$|\psi_1\rangle = H^{\otimes n}|0\rangle = \frac{1}{2^{n/2}} \sum_{x=0}^{N-1} |x\rangle\,. \tag{7.15}$$

Es una superposición sobre todos los estados de la base computacional por igual. De hecho es $|+\rangle \otimes \cdots \otimes |+\rangle$ y es separable.

A continuación hay que aplicar el oráculo que actúa también sobre el qubit controlado. Teniendo en cuenta que

$$U_f|x\rangle \otimes |y\rangle = |x\rangle \otimes |y \oplus f(x)\rangle, \tag{7.16}$$

para un estado de la base computacional se tendría

$$\begin{aligned} U_f|x\rangle \otimes |-\rangle &= \frac{1}{\sqrt{2}}(U_f|x\rangle \otimes |0\rangle - U_f|x\rangle \otimes |1\rangle) \\ &= \frac{1}{\sqrt{2}}|x\rangle \otimes \left(|f(x)\rangle - |f(x)\oplus 1\rangle\right) \\ &= (-1)^{f(x)}|x\rangle \otimes |-\rangle. \end{aligned} \tag{7.17}$$

(Esto vale para cualquier función $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$.) O sea efectivamente el oráculo actúa como

$$|x\rangle \underset{U_f}{\to} (-1)^{f(x)}|x\rangle \equiv \hat{U}_f|x\rangle\,. \tag{7.18}$$

De hecho se podría haber definido así directamente (y prescindir del qubit auxiliar) pero la puerta

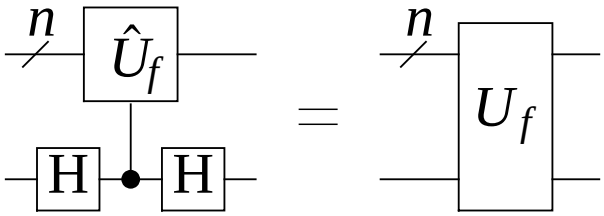


Figura 7.4: Reconstrucción U_f a partir de $\hat{U}_f$.

$\hat{U}_f$ no existe a nivel clásico y el oráculo U_f es la forma estándar de introducir una función. Como acabamos de ver la puerta U_f permite obtener $\hat{U}_f$, y también al revés: la construcción se muestra en la Fig. 7.4. [7.8]

El estado $|-\rangle$ del qubit auxiliar factoriza y puede obviarse en lo que sigue. Entonces después de aplicar U_f se tiene para los n qubits

$$|\psi_1\rangle \underset{\hat{U}_f}{\to} |\psi_2\rangle = \frac{1}{2^{n/2}} \sum_{x=0}^{N-1} (-1)^{f(x)}|x\rangle\,. \tag{7.19}$$

A continuación se aplican otra vez las puertas de Hadamard

$$|\Psi\rangle = H^{\otimes n}|\psi_2\rangle = \frac{1}{2^n} \sum_{x,z=0}^{N-1} (-1)^{f(x)+x\cdot z}|z\rangle\,. \tag{7.20}$$

Finalmente podemos medir el PVM $P_0 = |0\rangle\langle 0|$, $P_1 = I - P_0$ sobre el estado $|\Psi\rangle$. La amplitud de $|0\rangle$ es

$$\langle 0|\Psi\rangle = \frac{1}{2^n} \sum_{x}^{N-1} (-1)^{f(x)} = \begin{cases} (-1)^{f(0)} & \text{si } f \text{ es constante} \\ 0 & \text{si } f \text{ es equilibrada} \end{cases} \tag{7.21}$$

Por lo tanto la probabilidad de P_0 es 1 si f es constante y 0 si es equilibrada y P_1 al revés, y se sabe con una sola llamada a U_f.

[7.8] Las funciones f y $f \oplus 1$ producen $\hat{U}_{f\oplus 1} = -\hat{U}_f$. La diferencia de signo sólo se puede ver si $\hat{U}_f$ aparece controlado, no insertado. Por tanto $\hat{U}_f$ insertado no permitiría reconstruir U_f.

Como resultado colateral se deduce la importante identidad

$$\sum_{x=0}^{N-1}(-1)^{x\cdot z}=N\delta_{z,0}\,, \tag{7.22}$$

ya que cuando $f(x)\equiv 0$ $|\Psi\rangle=\dfrac{1}{N}\displaystyle\sum_{x,z=0}^{N-1}(-1)^{x\cdot z}|z\rangle$ pero $\langle 0|\Psi\rangle=1$ implica $|\Psi\rangle=|0\rangle$. Se deduce que todas las componentes de estados $|z\rangle$ con $z\neq 0$ se anulan. El mismo resultado se obtiene factorizando:

$$\sum_{x=0}^{N-1}(-1)^{x\cdot z}=\sum_{x=0}^{N-1}(-1)^{\sum_{j=1}^{n}x_jz_j}=\prod_{j=1}^{n}\sum_{x_j=0}^{1}(-1)^{x_jz_j}=\prod_{j=1}^{n}(1+(-1)^{z_j}) \tag{7.23}$$

y el resultado se anula si y sólo si algún $z_j=1$, es decir, si $z\neq 0$. Cuando $z=0$ la suma es $2^n=N$.

7.2. Algoritmo de Berstein-Vazirani

Se puede utilizar el mismo circuito para resolver otro problema, planteado por Berstein y Vazirani. En este caso también se tiene una función $f:\mathbb{Z}_2^n\to\mathbb{Z}_2$ y se sabe que es de la forma

$$f(x)=a\cdot x+b\ \ (\text{mód}\,2) \tag{7.24}$$

siendo $a\in\mathbb{Z}_2^n,\ b\in\mathbb{Z}_2$. [7.9] No se conocen ni a ni b y es trata de determinar a. El parámetro a contiene n bits y cada llamada a f proporciona un bit por lo que harán falta al menos n llamadas. Por ejemplo, $f(0)=b$ proporciona b, luego los bits de a se obtienen calculando $a_j=f(x)-b$ para $x=0\ldots1\ldots0$ (un 1 en la posición j-ésima) para $j=1,\ldots,n$.

Para esta función f, el estado final después del circuito es (usando (7.20))

$$|\Psi\rangle=\frac{(-1)^b}{N}\sum_{z=0}^{N-1}\sum_{x=0}^{N-1}(-1)^{x\cdot(z+a)}|z\rangle\,. \tag{7.25}$$

Usando ahora la propiedad (7.22),

$$|\Psi\rangle=(-1)^b\sum_{z=0}^{N-1}\delta_{z,a}|z\rangle=(-1)^b|a\rangle\,. \tag{7.26}$$

[7.9] No es la función más general, de hecho hay $2^N=2^{2^n}$ funciones posibles (2^n bits) mientras que las del tipo afín se especifican con $n+1$ bits de información, hay $2N$ funciones afines. En general para dos conjuntos $A^B=\{f\,|\,f:B\to A\}$ y $|A^B|=|A|^{|B|}$.

Por tanto una medida en la base computacional proporciona el valor de a, con una sola llamada a la función.

7.3. Algoritmo de búsqueda de Grover

7.3.1. Algoritmo de Grover

Los algoritmos vistos hasta ahora pretendían probar que el tratamiento cuántico puede ser más eficiente que el mejor algoritmo clásico en algún problema. Pero son problemas ad hoc y con poca aplicación práctica. El algoritmo de Grover es cualitativamente distinto en el sentido de que es de uso genérico. Se trata de encontrar una entrada marcada en una base de datos no estructurada. También se conoce como el problema de encontrar una aguja en un pajar. Matemáticamente el problema se puede formular como sigue: Sea $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$, tal que para cierto valor $a \in \mathbb{Z}_2^n$

$$f(x) = \delta_{x,a} = \begin{cases} 1 & x = a \\ 0 & x \neq a \end{cases} \tag{7.27}$$

No se conoce el valor de a y se trata de determinar su valor. La información sobre f y a nos la proporcionan exclusivamente a través de su caja negra u oráculo U_f. La Fig. 7.5 ilustra el circuito para el oráculo de la función correspondiente a $n = 3$ y $a = (0,1,1)$.

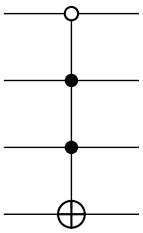


Figura 7.5: Oráculo para $f : \mathbb{Z}_2^3 \to \mathbb{Z}_2$ tal que $f(0,1,1) = 1$, y $f(x) = 0$ si $x \neq (0,1,1)$.

Clásicamente se requiere un número medio de llamadas $N/2$ a la función, es decir $O(N)$ evaluaciones de f. Cuánticamente se puede hacer con $O(\sqrt{N})$ invocaciones a U_f mediante el algoritmo de Grover.

Podría parecer que si se tiene una caja negra que proporciona $f(x)$ para cada x que se quiera, automáticamente se tiene a. Aunque es cierto que el valor de a está completamente determinado, extraerlo requiere un trabajo. Así como *comprobar* si un candidato x es realmente a es fácil usando el

oráculo, *determinar* el valor de a dado el oráculo no es inmediato. Técnicamente es un problema de clase NP, es decir, comprobación polinómica (la comprobación es eficiente). La búsqueda en sí no es de clase P sino exponencial (ineficiente) en n, tanto clásica como cuánticamente.

Veamos cómo funciona el algoritmo de Grover.

Dado un subespacio $\mathscr{V} \subset \mathscr{H}$ se puede definir el operador R_V de **reflexión** respecto de $\mathscr{V}$,

$$\forall|\psi\rangle \in \mathscr{V} \quad |\psi\rangle = |\psi_{\|}\rangle + |\psi_{\perp}\rangle \xrightarrow[R_V]{} |\psi_{\|}\rangle - |\psi_{\perp}\rangle, \qquad |\psi_{\|}\rangle \in \mathscr{V}, \quad |\psi_{\perp}\rangle \in \mathscr{V}^{\perp}. \tag{7.28}$$

R_V es unitario y $R_V^2 = I$, además $R_{V^\perp} = -R_V$. En función de los proyectores ortogonales sobre $\mathscr{V}$ y $\mathscr{V}^{\perp}$, $R_V = P_V - P_{V^\perp}$.

Los operadores $\hat{U}_f|x\rangle = (-1)^{f(x)}|x\rangle$ para funciones f de n en 1 bit, se pueden ver como operadores de reflexión, siendo $\mathscr{V}$ el subespacio subtendido por los $|x\rangle$ con $f(x) = 0$ y $\mathscr{V}^{\perp}$ el subespacio correspondiente a $f(x) = 1$.

Podemos definir las siguientes reflexiones

$$R_{a^\perp} = I - 2|a\rangle\langle a|, \qquad R_h = 2|h\rangle\langle h| - I, \tag{7.29}$$

donde

$$|h\rangle \equiv H^{\otimes n}|0\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle\,. \tag{7.30}$$

$R_{a^\perp}$ produce $\psi_a + \psi_{a^\perp} \to -\psi_a + \psi_{a^\perp}$ mientras que R_h produce $\psi_h + \psi_{h^\perp} \to \psi_h - \psi_{h^\perp}$.

El estado $|h\rangle$ se reparte por igual sobre todos los estados de la base computacional. Va a implementar el paralelismo cuántico, igual que en los algoritmos de Deutsch-Jozsa o Bernstein-Vazirani.

El algoritmo de Grover va a consistir en aplicar repetidamente el **operador de Grover**

$$G = R_h R_{a^\perp} \tag{7.31}$$

sobre el estado $|h\rangle$ y luego medir en la base computacional. Después de aplicar G un número de veces $O(\sqrt{N})$ la probabilidad de obtener $|a\rangle$ como resultado de la medida es próxima a 1.

Veamos que tenemos todos los elementos para hacer el cálculo:

i) El estado $|h\rangle$ sabemos construirlo usando puertas de Hadamard a partir de $|0\rangle$.

ii) El operador $R_{a^\perp}$ actúa según:

$$R_{a^\perp}|x\rangle = \begin{cases} -|x\rangle & x = a \\ +|x\rangle & x \neq a \end{cases} \tag{7.32}$$

Por tanto $R_{a^\perp} = \hat{U}_f$. Este operador lo tenemos a través del oráculo de la función f.

iii) El operador R_h se puede construir mediante $R_h = H^{\otimes n} R_0 H^{\otimes n}$, donde $R_0 = 2|0\rangle\langle 0| - I$ que también es factible:

$$-R_0|x\rangle = \begin{cases} -|x\rangle & x = 0 \\ +|x\rangle & x \neq 0 \end{cases}. \tag{7.33}$$

R_0 es esencialmente el oráculo de la función $x \mapsto \delta_{x,0}$.

Una fase extra (en este caso el signo en $-R_0$) en un operador unitario U **insertado** en un circuito produce una fase extra en el estado final y no tiene ningún efecto físico, todos los operadores $e^{i\varphi}U$ son equivalentes. Nótese que esto ya no es cierto si el operador U aparece **controlado** por otros qubits y no simplemente insertado en el circuito, como ocurre en Fig. 7.4.

Sea $\mathscr{V}_R$ el *espacio vectorial real* subtendido por las combinaciones lineales reales de $|h\rangle$ y $|a\rangle$. Teniendo en cuenta que $\langle a|h\rangle = \dfrac{1}{\sqrt{N}}$, se tiene

$$\begin{aligned} R_{a^\perp}|a\rangle &= -|a\rangle, \qquad R_{a^\perp}|h\rangle = |h\rangle - 2\gamma|a\rangle, \\ R_h|h\rangle &= |h\rangle, \qquad R_h|a\rangle = -|a\rangle + 2\gamma|h\rangle, \qquad \gamma \equiv \frac{1}{\sqrt{N}}. \end{aligned} \tag{7.34}$$

Se deduce que el espacio $\mathscr{V}_R$ es invariante bajo la acción de G. Todos los estados $G^k|h\rangle,\ k = 0,1,2,\ldots$, están en $\mathscr{V}_R$ que es un plano isomorfo a $\mathbb{R}^2$.

El caso de interés $N \gg 1$, que es cuando se requiere un método eficiente de búsqueda. En este caso $|h\rangle$ y $|a\rangle$ son casi perpendiculares, ya que $\langle h|a\rangle = \gamma \ll 1$. El espacio de Hilbert completo tiene una dimensión $N = 2^n$, y $|a\rangle$ es uno de los N vectores (casi) perpendiculares a $|h\rangle$. Sin embargo se ha conseguido reducir el problema a uno mucho más asequible en el espacio $\mathbb{R}^2$ en el que actúa el operador de Grover.

En el plano $\mathscr{V}_R$ se puede definir un vector $|a^\perp\rangle$ ortogonal a $|a\rangle$,

$$|h\rangle = \gamma|a\rangle + \sqrt{1-\gamma^2}|a^\perp\rangle \tag{7.35}$$

Denotamos α el ángulo entre $|a^{\perp}\rangle$ y $|h\rangle$ (entonces $\frac{\pi}{2}-\alpha$ será el ángulo entre $|a\rangle$ y $|h\rangle$)

$$\gamma = \langle h|a\rangle = \cos(\pi/2-\alpha) = \text{sen}(\alpha), \qquad \alpha = \text{arc sen}(1/\sqrt{N}) \tag{7.36}$$

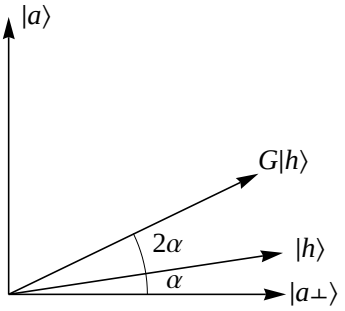


Figura 7.6: Acción del operador de Grover en el plano $\mathscr{V}_R$.

Veamos ahora que operador de Grover es precisamente una rotación de ángulo 2α en el plano $\mathscr{V}_R$.

Dado que el operador G conserva el producto escalar tiene que ser una **rotación propia** (conserva la orientación) o **impropia** (invierte la orientación). [7.10] En el plano, las rotaciones impropias son reflexiones respecto de una recta que pasa por el origen. Puesto que $R_{a^{\perp}}$ y R_h son ambas reflexiones, se sigue que G conserva la orientación y *es una rotación* en $\mathscr{V}_R$. El ángulo de dicha rotación se puede determinar viendo su acción sobre un vector cualquiera, elegimos $|a\rangle$:

$$G|a\rangle = -R_h|a\rangle = |a\rangle - 2\gamma|h\rangle, \qquad \langle a|G|a\rangle = 1-2\gamma^2 = 1-2\sin^2(\alpha) = \cos(2\alpha) \tag{7.37}$$

que implica

$$G = \text{Rotación}(2\alpha)\,. \tag{7.38}$$

Otra forma de obtener la misma conclusión es ver $\mathscr{V}_R$ como el plano complejo siendo $|a^{\perp}\rangle = 1$ y $|a\rangle = i$, cada $|\psi\rangle = z$ tiene un argumento θ. α es el argumento de $|h\rangle$. Bajo $R_{a^{\perp}}$ $\theta \to -\theta$, mientras que bajo R_h $\theta \to 2\alpha - \theta$. En conjunto, bajo G $\theta \to -\theta \to \theta + 2\alpha$ que es una rotación de ángulo 2α.

Un punto crucial es que el ángulo α no depende del $|a\rangle$ concreto (que no conocemos) ya que la idea es ir rotando el vector con G empezando con $|h\rangle$ hasta llegar a $|a\rangle$ o casi. Puesto que en cada aplicación el ángulo disminuye en 2α siempre se puede conseguir un estado final $G^k|h\rangle$ formando un

[7.10] En un espacio vectorial real, dos bases tienen la misma orientación cuando la matriz de cambio de base tiene determinante positivo.

ángulo no mayor que α con $|a\rangle$. Inicialmente el ángulo es $\pi/2-\alpha$, el exponente ideal k que produciría un ángulo nulo es

$$k_{\text{ideal}} = \frac{\frac{\pi}{2}-\alpha}{2\alpha} = \frac{\pi}{4\,\text{arc sen}(1/\sqrt{N})} - \frac{1}{2} = O(\sqrt{N}). \tag{7.39}$$

Como k debe ser entero, el valor óptimo $\bar{k}$ se obtiene por redondeo al entero más próximo:

$$\bar{k} = k_{\text{ideal}} + \xi, \qquad |\xi| \leq \frac{1}{2}, \qquad \bar{k} \in \mathbb{Z}. \tag{7.40}$$

El estado final $|\bar{a}\rangle \equiv G^{\bar{k}}|h\rangle$ se mide en la base computacional lo cual nos proporcionará un valor $|x\rangle$. Dado que el ángulo entre $|\bar{a}\rangle$ y $|a\rangle$ es $2\alpha\xi$, la probabilidad de que x no sea a viene dada por

$$\text{Prob}(x \neq a) = 1 - |\langle \bar{a}|a\rangle|^2 = \text{sen}^2(2\alpha\xi) = O(1/N). \tag{7.41}$$

Que x sea a o no se puede comprobar a posteriori, directamente si se tiene la función, o bien invocando el oráculo: $U_f|x\rangle \otimes |0\rangle = |x\rangle \otimes |f(x)\rangle$, una medida del bit auxiliar permite saber si $f(x) = 1$. En el improbable caso de que $f(x) = 0$ simplemente se repite el proceso. Para N grande $\bar{k} \approx \frac{\pi}{4}\sqrt{N}$. Es una **mejora cuadrática** sobre el número de llamadas a la función en la búsqueda clásica, $N/2$ (en promedio).

El método de Grover se puede extender para buscar M valores marcados, $a_1, \ldots, a_M$ de entre N (antes $M = 1$) supuesto M conocido. [7.11]

7.3.2. Implementación del algoritmo

El algoritmo de búsqueda cuántica que se acaba de describir tiene todo el sentido si se nos proporciona el oráculo y se trata de adivinar el $x = a$ marcado, pero puede parecer paradójico si se intenta aplicar a un caso práctico en el que nosotros construimos el oráculo, porque da la sensación de que para hacerlo es necesario saber ya la solución de antemano, es decir, necesitamos tener la solución para empezar a buscarla.

Veamos que no es así con una aplicación concreta. Lo que se quiere ver es que el oráculo y todo el algoritmo se puede construir de una manera automatizada, sin saber la solución y sin perder paralelismo cuántico. [7.12]

[7.11] Nielsen y Chuang, pág. 248.

[7.12] Otra cuestión es que el método que se va a describir sea de interés práctico. Probablemente el algoritmo de Grover sea útil no tanto para búsquedas en bases de datos como en optimización de búsqueda en problemas difíciles, de clase NP (Nielsen y Chuang, pág. 265).

El caso que estudiamos es el siguiente. Se tiene una base de datos (una lista) con $N = 2^n$ palabras de ℓ bits cada una. Se supone que no hay palabras repetidas (requiere $n \leq \ell$). $x \in \{0, \ldots, N-1\}$ o equivalentemente $x \in \mathbb{Z}_2^n$ es el índice del registro x-ésimo en la base de datos y $d_x \in \mathbb{Z}_2^\ell$ es la palabra en el registro x. Nos dan una palabra $s \in \mathbb{Z}_2^\ell$ que está en la base de datos y queremos encontrar su índice x_s, tal que $s = d_{x_s}$. s es una variable dinámica, el proceso de búsqueda debe poder repetirse según llegan los distintos valores de s.

Vamos a usar 3 registros (al menos explícitamente):

i) $|x\rangle_1$ formado por n qubits va a contener el índice x.

ii) $|d\rangle_2$ formado por ℓ qubits. Se usará para contener una palabra de la base de datos d_x.

iii) $|-\rangle_3$ es un qubit auxiliar para usar el oráculo. El estado $|-\rangle_3$ se escribe al principio y no va a cambiar durante toda la ejecución del algoritmo.

El valor de s y la base de datos $\{x \mapsto d_x\}$ pueden estar contenidos en qubits o bien codificarse mediante puertas lógicas. Suponemos esto último ya que sólo requiere una memoria clásica (más estable) y además es más fácil de hacer.

El algoritmo de Grover va a actuar sobre $|x\rangle_1$. La idea es inicializar a $|h\rangle_1 \otimes |0\rangle_2 \otimes |-\rangle_3$, y aplicar repetidamente el operador G de Grover, pasando a $(G^k|h\rangle_1) \otimes |0\rangle_2 \otimes |-\rangle_3$, para eventualmente medir $|x\rangle_1$.

$G = R_h \hat{U}_f$ actúa sobre $|x\rangle_1$ y no modifica los demás registros. El operador R_h es conocido. La acción de $\hat{U}_f$ requiere tres pasos. El primero es un operador U_L que carga el valor d_x en $|0\rangle_2$ dado x en $|x\rangle_1$:

$$U_L|x\rangle_1 \otimes |d\rangle_2 = |x\rangle_1 \otimes |d \oplus d_x\rangle_2 \tag{7.42}$$

U_L tiene que leer de la base de datos. Como se ha dicho, los valores d_x están codificados en puertas lógicas. En la Fig. 7.7 se ilustra el caso $n = 2$ y $\ell = 3$.

La puerta d_x contiene la palabra d_x de la base de datos. Así si por ejemplo $d = 011$, la puerta aplica los operadores $I \otimes X \otimes X$ sobre $|000\rangle_2$ para producir $|011\rangle_2$ (Fig. 7.8).

Como inicialmente el estado del segundo registro era 0, después de aplicar U_L se tendrá $|x\rangle_1 \otimes |d_x\rangle_2 \otimes |-\rangle_3$. En realidad el primer registro estaba en estado $|h\rangle_1$ pero se aplica paralelismo cuántico y U_L actúa sobre todos los posibles x a la vez. Por supuesto es esencial preservar la linealidad cuántica aunque la base de datos está codificada clásicamente, tal y como hace U_L. [7.13]

[7.13] No serviría un protocolo clásico que dado x lea de la base de datos y ponga d_x en $|\,\rangle_2$.

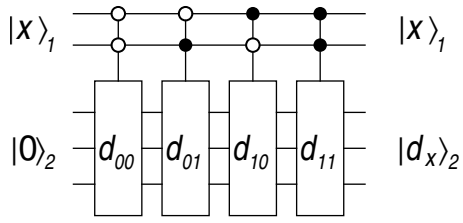


Figura 7.7: Circuito para $U_L|x\rangle_1 \otimes |0\rangle_2 = |x\rangle_1 \otimes |d_x\rangle_2$, para $n = 2$ y $\ell = 3$.

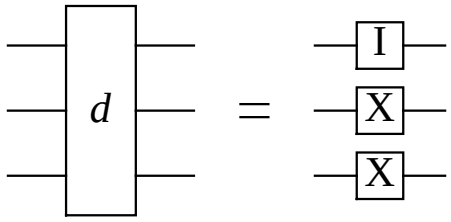


Figura 7.8: Forma de una puerta d que contiene la palabra 011 ($\ell = 3$).

El segundo paso es aplicar un operador U_C que compare el contenido del registro $|d_x\rangle_2$ con el valor de s y produzca una fase (-1) si y sólo si $d_x = s$,

$$U_C|d\rangle_2 = (-1)^{\delta_{d,s}}|d\rangle_2\,. \tag{7.43}$$

Este operador no depende del valor de n. De nuevo, la cadena s está codificada en una puerta lógica, del mismo tipo que las usadas para contener d_x. Una realización de U_C se muestra en la Fig. 7.9, utilizando el qubit auxiliar $|-\rangle_3$.

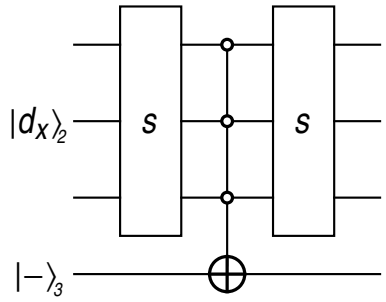


Figura 7.9: Puerta U_C para $\ell = 3$: produce un signo menos cuando $d_x = s$.

El operador $\hat{U}_f$ se completa aplicando nuevamente $U_L = U_L^{-1}$ de modo que $|\,\rangle_2$ vuelva a quedar en

estado $|0\rangle_2$. El efecto neto de $\hat{U}_f = U_L U_C U_L$ sobre $|x\rangle_1 \otimes |0\rangle_2 \otimes |-\rangle_3$ es un signo $(-1)^{\delta_{x,x_s}}$, tal como se requiere en el algoritmo de Grover.

Tal y como se quería, $\hat{U}_f$ se ha construido sin conocer x_s de antemano y sin perder paralelismo cuántico, aunque la construcción concreta de U_L presentada aquí no tiene interés práctico (es costosa). El efecto de U_L es producir un estado entrelazado que contiene la base de datos:

$$|\Psi_0\rangle := |h\rangle_1 \otimes |0\rangle_2 = \frac{1}{\sqrt{N}} \sum_x |x\rangle_1 \otimes |0\rangle_2 \xrightarrow[U_L]{} |\Psi\rangle := \frac{1}{\sqrt{N}} \sum_x |x\rangle_1 \otimes |d_x\rangle_2 \,. \tag{7.44}$$

El factor $\hat{U}_f$ del operador de Grover podría hacerse con $|\Psi\rangle$, en vez de $|\Psi_0\rangle$, lo cual requeriría aplicar U_L sólo una vez al principio. Desgraciadamente la parte R_h tiene que hacerse sobre el estado separable $|\Psi_0\rangle$ y hay que aplicar U_L repetidamente.

7.3.3. Mejora cuántica en algoritmos de búsqueda

La reducción de $O(N)$ a $O(\sqrt{N})$ conseguida por el algoritmo de Grover (aunque quizá no el coeficiente $\frac{\pi}{4}\sqrt{N}$) es de hecho la máxima posible por métodos cuánticos. Esto se puede demostrar como sigue. El valor de a está codificado en el oráculo $\hat{U}_f = R_{a^\perp} = I - 2|a\rangle\langle a|$, entonces partiendo de un estado inicial $|0\rangle$ (sin pérdida de generalidad) hacemos k llamadas al oráculo con evolución unitaria por en medio para obtener

$$|\psi_k^a\rangle = U_k \hat{U}_f \cdots U_1 \hat{U}_f U_0 |0\rangle. \tag{7.45}$$

Los operadores U_j no dependen de a (que no se conoce) y lo mismo k. Esto es lo más general que se puede hacer. [7.14] Que en el caso óptimo $k = O(\sqrt{N})$ se va a deducir de requerir que $\forall a \; |\langle a|\psi_k^a\rangle|^2 \geq F$ para cierta fidelidad mínima $F > 0$ (independiente de N y k).

Si se define

$$|\psi_k\rangle = U_k \cdots U_1 U_0 |0\rangle, \qquad D_k = \sum_{a=0}^{N-1} \| |\psi_k^a\rangle - |\psi_k\rangle \|^2, \tag{7.46}$$

es posible demostrar por inducción [7.15] la siguiente cota superior para D_k:

$$\forall k \quad D_k \leq 4k^2 \,, \tag{7.47}$$

[7.14] O casi. Las puertas $\hat{U}_f$ podrían aparecer en el circuito controladas en vez de insertadas (o combinaciones de ambas opciones). Como se ve en Fig. 7.4 esto es equivalente a insertar U_f. También se podrían hacer medidas intermedias. Presumiblemente la demostración se puede extender al caso en que el oráculo se invoca k veces en todas sus versiones.

[7.15] Bergou y Hillery, pág. 100.

independientemente de cuáles sean los U_j. La idea es obtener una cota inferior *dependiente de N pero no de k* a base de requerir que el protocolo garantice una fidelidad mínima entre $|a\rangle$ y $|\psi_k^a\rangle$ para cualquier a. De ese modo se acotará k inferiormente.

Para establecer una cota inferior a D_k, usamos la siguiente desigualdad válida para tres estados normalizados $|\psi\rangle$, $|\phi\rangle$ y $|\chi\rangle$ cualesquiera, [7.16]

$$\||\psi\rangle - |\phi\rangle\|^2 \geq 2\left(1 - \sqrt{1-|\langle\chi|\psi\rangle|^2} - |\langle\chi|\phi\rangle|\right). \tag{7.48}$$

La aplicamos a $|\psi_k^a\rangle$, $|\psi_k\rangle$ y $|a\rangle$, sumando sobre a:

$$\forall k \qquad D_k \geq 2\sum_a \left(1 - \sqrt{1-|\langle a|\psi_k^a\rangle|^2} - |\langle a|\psi_k\rangle|\right). \tag{7.49}$$

Se puede aplicar aquí la desigualdad $\langle x\rangle^2 \leq \langle x^2\rangle$ (Cauchy-Schwarz)

$$\left(\frac{1}{N}\sum_a |\langle a|\psi_k\rangle|\right)^2 \leq \frac{1}{N}\sum_a |\langle a|\psi_k\rangle|^2 = \frac{1}{N} \implies \sum_a |\langle a|\psi_k\rangle| \leq \sqrt{N}, \tag{7.50}$$

para obtener la desigualdad

$$\forall k \qquad D_k \geq 2(N-\sqrt{N}) - 2\sum_a \sqrt{1-|\langle a|\psi_k^a\rangle|^2}. \tag{7.51}$$

Para U_j genéricos el estado final $|\psi_k^a\rangle$ no se parecerá especialmente a $|a\rangle$, es decir $|\langle a|\psi_k^a\rangle|^2 = O(1/N)$. Pero supongamos que con una elección adecuada de U_j y k, que pueden depender de N pero no de a, el método funciona para cualquier valor de a, es decir, se garantiza una fidelidad mínima $F > 0$ (independiente de N o acotada inferiormente por un valor no nulo independiente de N) en el estado final:

$$\forall a \quad |\langle a|\psi_k^a\rangle|^2 \geq F. \tag{7.52}$$

(Aunque la fidelidad no sea 1, si es finita independiente de N basta repetir el proceso. En todo caso añadiría un factor $O(1)$ que no cambia el comportamiento $k = O(\sqrt{N})$.) Esto implica $\sqrt{1-|\langle a|\psi_k^a\rangle|^2} \leq \sqrt{1-F}$. Juntando las cotas inferior y superior

$$4k^2 \geq D_k \geq 2\left(1-\sqrt{1-F}\right)N - 2\sqrt{N}. \tag{7.53}$$

[7.16] Ibid. pág. 100.

Se obtiene entonces una cota inferior $O(\sqrt{N})$ al número k de llamadas al oráculo si el algoritmo ha de garantizar una fidelidad mínima F para todos los a. Para N grande

$$k \geq C_F\sqrt{N} + O(1), \qquad C_F = \sqrt{(1-\sqrt{1-F})/2}. \tag{7.54}$$

Si el algoritmo es tal que garantiza $F \to 1$ para N grande, la cota es $k > \sqrt{N/2}$. Para Grover es $\bar{k} = \frac{\pi}{4}\sqrt{N}$ y es consistente con la desigualdad.

7.4. Algoritmo de Simon (determinación de periodos)

Estudiamos ahora el algoritmo de Simon que ilustra métodos de determinar el periodo de una función periódica con un caso muy simple. Existen protocolos más sofisticados como el que se usa como parte del algoritmo de factorización de Shor.

7.4.0.1. Apartado matemático: espacios $\mathbb{C}^A$

Sea A un conjunto finito (para evitar problemas de convergencia). Se puede asociar a cada elemento $y \in A$ un elemento de una base ortonormal $|y\rangle$ de un espacio de Hilbert complejo de dimensión $d = |A|$, que podemos denominar $\mathscr{H}_A$.

Si esto parece arbitrario, la construcción se puede hacer de modo sistemático como sigue. Definimos $\mathscr{H}_A$ como $\mathbb{C}^A$, que es el espacio de funciones $g : A \to \mathbb{C}$. Claramente es un espacio vectorial de dimensión d. Para $y_0 \in A$, $|y_0\rangle$ es la función $y \in A \to \delta_{y,y_0}$. Los estados son $|g\rangle = \sum_{y\in A} g(y)|y\rangle$. El producto escalar se define imponiendo que $\{|y\rangle, y \in A\}$ sea una base ortonormal y convierte a $\mathbb{C}^A$ en un espacio de Hilbert.

Si por ejemplo $A = \{\text{pera}, \text{manzana}, \text{plátano}\}$, es posible cómodamente sumar peras con manzanas, por ejemplo, $7|\text{pera}\rangle + 4|\text{manzana}\rangle$. O incluso $|\text{pera}\rangle \otimes |\text{manzana}\rangle$.

Tenemos una función $f : \mathbb{Z}_2^n \to A$, donde A es un conjunto finito. Podemos asociar a A un espacio de Hilbert $\mathscr{H}_A$ con base ortonormal $\{|y\rangle, y \in A\}$. Se puede entonces definir el operador U_f tal que

$$\forall x \in \mathbb{Z}_2^n \qquad U_f|x\rangle \otimes |y_0\rangle = |x\rangle \otimes |f(x)\rangle. \tag{7.55}$$

(El primer factor es el espacio asociado a $\mathbb{Z}_2^n$ y el segundo el asociado a A.) Este operador es isométrico y se puede extender a unitario. Si por ejemplo, $A = \mathbb{Z}_2^m$, $U_f|x\rangle \otimes |y\rangle = |x\rangle \otimes |y \oplus f(x)\rangle$ es una solución. Lo que importa enfatizar aquí es que el algoritmo no requiere ninguna estructura algebraica en A.

Se dice que un $\xi \in \mathbb{Z}_2^n$ tal que

$$\forall x \in \mathbb{Z}_2^n \qquad f(x) = f(x+\xi) \tag{7.56}$$

es un **periodo** de la función f. (La suma $x+\xi$ es en el espacio vectorial $\mathbb{Z}_2^n$, coincide con $x \oplus \xi$.)

Es inmediato comprobar que el conjunto de periodos de una función forman un subespacio vectorial, $\mathscr{V}_p \subset \mathbb{Z}_2^n$ de dimensión r. Cuando $r = 0$ f es una función aperiódica, si $\dim \mathscr{V}_p = 1$ f es periódica con periodo ξ (siendo $\mathscr{V}_p = \{0, \xi\}$). [7.17] Si $r > 1$ la función es multiperiódica.

El problema que se plantea es determinar $\mathscr{V}_p$ empleando un mínimo de llamadas al oráculo. Veamos que cuánticamente se puede hacer con un número de llamadas de $O(n)$ (con probabilidad 1).

De hecho lo hace un circuito similar al de la Fig. 7.3 del problema de Deutsch-Jozsa. Tenemos un primer registro con n-qubits que contiene x y un segundo registro que contiene $|y_0\rangle$. El circuito comienza con

$$\begin{aligned} |0\rangle \otimes |y_0\rangle \longrightarrow (H^{\otimes n}|0\rangle) \otimes |y_0\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle \otimes |y_0\rangle \\ \xrightarrow[U_f]{} \frac{1}{\sqrt{N}} \sum_x |x\rangle \otimes |f(x)\rangle \equiv |\Psi_1\rangle \,. \end{aligned} \tag{7.57}$$

El disponer de copias del estado $|\Psi_1\rangle$ es todo lo que realmente se necesita para llevar a cabo el algoritmo.

Aplicamos de nuevo las puertas de Hadamard sobre el primer registro, usando la expresión obtenida anteriormente

$$H^{\otimes n}|x\rangle = \frac{1}{\sqrt{N}} \sum_{z=0}^{N-1} (-1)^{x\cdot z} |z\rangle \,. \tag{7.58}$$

Tenemos entonces

$$|\Psi_1\rangle \longrightarrow |\Psi_2\rangle \equiv \frac{1}{\sqrt{N}} \sum_z |z\rangle \otimes |\phi_z\rangle, \qquad |\phi_z\rangle \equiv \frac{1}{\sqrt{N}} \sum_x (-1)^{x\cdot z} |f(x)\rangle \,. \tag{7.59}$$

Para cualquier $\xi \in \mathbb{Z}_2^n$ la traslación $x \to x+\xi$ es una biyección en $\mathbb{Z}_2^n$, de modo que cuando x recorre todos los valores posibles una y una sola vez, también $x+\xi$ recorre todos los valores posibles

[7.17] En $\mathbb{Z}_2^n$ el único múltiplo de un vector ξ es él mismo o 0.

una y una sola vez. Por tanto podemos reescribir $|\phi_z\rangle$ cambiando $x \to x+\xi$ en los sumandos. Si además ξ es un periodo de f se tendrá

$$\begin{aligned} \forall \xi \in \mathscr{V}_p \quad |\phi_z\rangle &= \frac{1}{\sqrt{N}} \sum_x (-1)^{(x+\xi)\cdot z} |f(x+\xi)\rangle = \frac{1}{\sqrt{N}} \sum_x (-1)^{(x+\xi)\cdot z} |f(x)\rangle \\ &= (-1)^{\xi\cdot z} |\phi_z\rangle \,. \end{aligned} \tag{7.60}$$

Se concluye entonces que $|\phi_z\rangle = 0$ a menos que $\xi \cdot z = 0 \pmod 2$. Como ξ es cualquier periodo se sigue que los únicos $|\phi_z\rangle$ no nulos corresponden a z ortogonales a todos los periodos, es decir, $z \in \mathscr{V}_p^\perp$:

$$|\Psi_2\rangle = \frac{1}{\sqrt{N}} \sum_{z \in \mathscr{V}_p^\perp} |z\rangle \otimes |\phi_z\rangle \,. \tag{7.61}$$

Nótese que $\mathscr{V}_p^\perp = \{z \in \mathbb{Z}_2^n \mid \forall \xi \in \mathscr{V}_p \ \xi \cdot z = 0\}$ es un subespacio perfectamente bien definido de Z_2^n. Sin embargo debe tenerse en cuenta que en $\mathbb{Z}_2^n$ hay vectores isótropos, $x \neq 0$ tales que $x \cdot x = 0$ (por ejemplo $x = (1,1,0)$ para $n = 3$). Por tanto algunas otras propiedades también son inusuales, así en general $\mathscr{V}_p \cap \mathscr{V}_p^\perp \neq 0$ y $\mathbb{Z}_2^n$ no será suma directa de $\mathscr{V}_p$ y $\mathscr{V}_p^\perp$. En todo caso las propiedades $\dim \mathscr{V}_p + \dim \mathscr{V}_p^\perp = n$ y $(\mathscr{V}_p^\perp)^\perp = \mathscr{V}_p$ sí se satisfacen. [7.18]

Para evitar confusiones es crucial distinguir entre el espacio $\mathbb{Z}_2^n$ (donde reside z) y el espacio de Hilbert cuántico $\mathbb{C}^{2^n}$ donde reside $|z\rangle$: $\mathbb{Z}_2^n$ es un espacio vectorial de dimensión n sobre el cuerpo $\mathbb{Z}_2$, hay a lo sumo n vectores linealmente independientes (y en particular ortogonales). Es importante notar que los z no son linealmente independientes como vectores de $\mathbb{Z}_2^n$, en total hay $N = 2^n$ vectores z distintos y sólo n independientes. No debe confundirse $\mathbb{Z}_2^n$ con el espacio $\mathscr{H}$ sobre el cuerpo $\mathbb{C}$ subtendido por los estados $|z\rangle$. Hay N estados $|z\rangle$ todos linealmente independientes y de hecho ortonormales como vectores de $\mathscr{H}$. El espacio $\mathscr{H} = \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2 = \mathbb{C}^N$ tiene dimensión N mucho mayor que n en general, y tiene un número infinito de vectores (con un máximo de N linealmente independientes).

Al iterar el proceso de obtener el estado $|\Psi_2\rangle$ y medir $|z\rangle$ (el primer registro) se irán obteniendo valores de $z \in \mathbb{Z}_2^n$ siempre ortogonales a todos los periodos ξ (que no se conocen).

Para una función multiperiódica f, pero por lo demás genérica, los z_ℓ obtenidos al medir tenderán a llenar (en el sentido de subtender) todo el espacio $\mathscr{V}_p^\perp$. Es decir, después de un número de llamadas $O(n)$ se podrá extraer un número $n-r$ de z_ℓ linealmente independientes que deje de crecer

[7.18] La segunda se deduce de la primera y $\mathscr{V}_p \subset (\mathscr{V}_p^\perp)^\perp$, y la primera es porque las ecuaciones $\xi_k \cdot z = 0 \ \ (k = 1, \ldots, r)$, siendo ξ_k una base de $\mathscr{V}_p$, eliminan r parámetros libres en z y quedan $n-r$. Se deduce que cuando $\mathscr{V}_p$ tiene vectores isótropos $\mathscr{V}_p + \mathscr{V}_p^\perp \subsetneq \mathbb{Z}_2^n$.

con más llamadas. El espacio ortogonal $\mathscr{V}_p$ se puede determinar encontrando todas las soluciones independientes del sistema de ecuaciones (a resolver en $\mathbb{Z}_2^n$ con operaciones en el cuerpo $\mathbb{Z}_2$)

$$\xi \cdot z_\ell = 0, \qquad \ell = 1, \ldots, n-r. \tag{7.62}$$

Aunque probable, sin más hipótesis sobre f no es posible afirmar con seguridad que el r así estimado es correcto (podría ser menor) y tampoco es posible verificar sin posibilidad de error que un ξ dado es un periodo de f sin hacer un número de llamadas sustancialmente mayor que $O(n)$ (más bien $O(N^\gamma)$ para $\gamma > 0$). Incluso en ese caso el conocimiento parcial de vectores en $\mathscr{V}_p^\perp$ permite acelerar la búsqueda de periodos por fuerza bruta.

Si se conoce r (este es el caso ideal) y se encuentran $n-r$ valores linealmente independientes z_ℓ entonces $\mathscr{V}_p$ queda unívocamente identificado con seguridad.

Podría ocurrir que no se obtengan los $n-r$ valores requeridos porque sistemáticamente algunos coeficientes factores $|\phi_z\rangle$ se anulen para ciertos z pero no por motivos de periodicidad (una degeneración accidental). Eso se puede excluir imponiendo más condiciones sobre la función. Supongamos que se sabe que cada valor imagen $f(x)$ aparece exactamente 2^r dos veces (es decir, $f(x) = f(y)$ implica $x + y \in \mathscr{V}_p$). En este caso, si w denota cualquiera de las imágenes en $\mathrm{ran}(f)$, todo x se podrá escribir de forma única como $x = x_w + \xi$, eligiendo un x_w canónico para cada w (tal que $f(x_w) = w$) y $\xi \in \mathscr{V}_p$. Entonces, si $z \in \mathscr{V}_p^\perp$

$$|\phi_z\rangle = \frac{1}{\sqrt{N}} \sum_x (-1)^{x\cdot z} |f(x)\rangle = \frac{1}{\sqrt{N}} \sum_{w \in \mathrm{ran}(f)} \sum_{\xi \in \mathscr{V}_p} (-1)^{z\cdot(x_w+\xi)} |w\rangle = \frac{2^r}{\sqrt{N}} \sum_{w \in \mathrm{ran}(f)} (-1)^{z\cdot x_w} |w\rangle \tag{7.63}$$

y por tanto el coeficiente $|\phi_z\rangle$ no se anula. De hecho todos los $z \in \mathscr{V}_p^\perp$ aparecen con la misma probabilidad $2^{(r-n)}$.

El esfuerzo para resolver las ecuaciones $n-1$ (posprocesado clásico) es a lo sumo polinómico en n. Esta es una importante ganancia con respecto del protocolo clásico. Por ejemplo, sabiendo que $r = 1$ y cada imagen $f(x)$ aparece exactamente dos veces, es necesario encontrar un par de valores x e y en $\mathbb{Z}_2^n$ tales que $f(x) = f(y)$ lo cual proporciona el periodo $\xi = x + y$. Esa búsqueda clásica requiere $O(\sqrt{N})$ llamadas a la función. [7.19]

[7.19] Buscar un x con un valor concreto $f(x)$ requiere $O(N)$ pero buscar x, y con el mismo valor, sea cual sea, sólo requiere $O(\sqrt{N})$ (paradoja del cumpleaños).

7.5. Transformada de Fourier y estimación de fases

7.5.1. Transformada de Fourier discreta

La **transformada de Fourier discreta** es una aplicación lineal unitaria $U_F : \mathbb{C}^N \to \mathbb{C}^N$,

$$\tilde{\psi} = U_F \psi\,, \qquad \tilde{\psi}_k = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} e^{2\pi i k x/N} \psi_x, \qquad k = 0, 1, \ldots, N-1\,. \tag{7.64}$$

Aquí $N \in \mathbb{Z}^+$ cualquiera. [7.20] Por conveniencia se puede suponer que ψ_x o $\tilde{\psi}_k$ están definidos $\forall x, k \in \mathbb{Z}$ como *funciones periódicas* con periodo N, $\psi_{x+N} = \psi_x$, $\tilde{\psi}_{k+N} = \tilde{\psi}_k$. (Y en efecto $e^{2\pi i k x/N}$ es invariante bajo $x \to x+N$ o $k \to k+N$.)

Podemos definir

$$\omega_N \equiv e^{2\pi i/N} \tag{7.65}$$

tal que

$$\tilde{\psi}_k = \frac{1}{\sqrt{N}} \sum_x \omega_N^{kx} \psi_x\,, \tag{7.66}$$

donde se sobreentiende que x se suma sobre un periodo (es decir, N enteros consecutivos)..

Conviene establecer la propiedad

$$\frac{1}{N} \sum_x \omega_N^{kx} = \delta_{k,0} \tag{7.67}$$

(donde se sobreentiende $k = 0 \pmod N$).

Demostración: Dado que es una serie geométrica se puede usar (7.93), o también

$$S = \sum_{x=0}^{N-1} \omega_N^{kx} = \sum_{x=1}^{N} \omega_N^{kx} = \sum_{x'=0}^{N-1} \omega_N^{k(x'+1)} = \omega_N^k S \implies 0 = (1-\omega_N^k) S, \tag{7.68}$$

que implica que $S = 0$ a menos que $k = 0 \pmod N$. □

Usando esta propiedad se comprueba que la transformada inversa es

$$\psi = U_F^{-1} \tilde{\psi}\,, \qquad \psi_x = \frac{1}{\sqrt{N}} \sum_k \omega_N^{-kx} \tilde{\psi}_k\,. \tag{7.69}$$

[7.20] Aquí kx es el producto de dos números enteros. No el producto escalar $k \cdot x$ en $\mathbb{Z}_2^n$.

De paso esta fórmula verifica que U_F es unitaria, ya que

$$(U_F^{-1})_{xk} = \frac{\omega_N^{-kx}}{\sqrt{N}} = \left(\frac{\omega_N^{kx}}{\sqrt{N}}\right)^* = (U_F^\dagger)_{xk}\,. \tag{7.70}$$

Alternativamente,

$$\|\tilde{\psi}\|^2 = \sum_k |\tilde{\psi}_k|^2 = \frac{1}{N}\sum_k \sum_{x,x'} \omega_N^{-kx}\omega_N^{kx'}\psi_x^*\psi_{x'} = \frac{1}{N}\sum_{x,x'}\delta_{xx'}\psi_x^*\psi_{x'} = \sum_x |\psi_x|^2 = \|\psi\|^2\,. \tag{7.71}$$

7.5.1.1. Apartado matemático: Serie y transformada de Fourier

Por conveniencia trabajamos ahora en el periodo $[-N/2+1, N/2]$ tanto para x como para k (suponemos N par).

Si definimos la variable reescalada $q_x = \Delta q\, x, \quad \Delta q \equiv L/N$ y se considera el límite de $N \to \infty$ con $L > 0$ fijo, q tiende a ser una variable continua en el intervalo $[-L/2, L/2]$. En ese intervalo definimos una función $f(q)$ tal que $f(q_x) = \psi_x$. Como $N \to \infty$ ahora $k \in \mathbb{Z}$. Definimos la función reescalada $\tilde{f}_k = \tilde{\psi}_k/\sqrt{N}$. Se tiene entonces

$$\tilde{f}_k = \frac{1}{N}\frac{1}{\Delta q}\sum_x \Delta q\, e^{2\pi i k q_x/L} f(q_x) \longrightarrow \tilde{f}_k = \frac{1}{L}\int_{-L/2}^{L/2} e^{2\pi i k q/L} f(q)\, dq \quad \forall k \in \mathbb{Z} \tag{7.72}$$

$\tilde{f}_k$ son los coeficientes de la serie de Fourier de $f(q)$ en $[-L/2, L/2]$. Igualmente para la serie de Fourier

$$\psi_x = \sum_k e^{-2\pi i k x/N}\frac{1}{\sqrt{N}}\tilde{\psi}_k \longrightarrow f(q) = \sum_{k\in\mathbb{Z}} e^{-2\pi i k q/L}\tilde{f}_k\,. \tag{7.73}$$

Si ahora se introduce una variable reescalada $p_k = \Delta p\, k$, con $\Delta p = 2\pi/L$, se define la función $\tilde{f}(p)$ tal que $\tilde{f}(p_k) = L\tilde{f}_k$, y se toma el límite $L \to \infty$, (de modo que $\Delta p \to 0$ y p_k tiende a una variable continua p)

$$L\tilde{f}_k = \int_{-L/2}^{L/2} e^{ip_k q} f(q)\, dq \longrightarrow \tilde{f}(p) = \int_{\mathbb{R}} e^{ipq} f(q)\, dq \qquad \forall p \in \mathbb{R}\,, \tag{7.74}$$

e igualmente

$$f(q) = \frac{1}{\Delta p}\sum_{k\in\mathbb{Z}} \Delta p\, e^{-ip_k q}\tilde{f}_k \longrightarrow f(q) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-ipq}\tilde{f}(p)\, dp, \tag{7.75}$$

que corresponden a la transformada de Fourier y su inversa.

Al igual que la serie y transformada de Fourier, la transformada discreta de Fourier también se puede hacer en varias dimensiones.

7.5.2. Transformada de Fourier cuántica

El espacio $\mathbb{C}^N$ se puede ver como un espacio de Hilbert de dimensión N de un sistema cuántico, y ψ_x son las componentes en una base ortonormal $|x\rangle$ de ese espacio.

$$|\psi\rangle = \sum_x \psi_x |x\rangle, \qquad |\tilde{\psi}\rangle = \hat{U}_F |\psi\rangle = \sum_k \tilde{\psi}_k |k\rangle \tag{7.76}$$

U_F es la matriz del operador unitario $\hat{U}_F$ en la base $|x\rangle$

$$\langle k|\hat{U}_F|x\rangle = (U_F)_{kx} = \frac{1}{\sqrt{N}} \omega_N^{kx}. \tag{7.77}$$

En este contexto a $\hat{U}_F$ se le denomina **transformada de Fourier cuántica**.

Aquí se ha adoptado un punto de vista activo: el estado $|x = j\rangle$ es el mismo que $|k = j\rangle$, y es el estado $|\psi\rangle$ el que ha cambiado por efecto de $\hat{U}_F$. El punto de vista pasivo sería que $|\psi\rangle$ no cambia, lo hacen sus componentes al cambiar de base. [7.21]

En lo que sigue denotamos $\hat{U}_F$ simplemente como U_F y seguimos un puntos de vista activo.

U_F obedece al principio de incertidumbre, ψ_x y $\tilde{\psi}_k$ no pueden estar ambos muy localizados. Así por ejemplo, si $|\psi\rangle = |x\rangle$ (totalmente localizado en x)

$$U_F|x\rangle = \frac{1}{\sqrt{N}} \sum_k \omega_N^{kx} |k\rangle \,. \tag{7.78}$$

El estado resultante $\tilde{\psi}_k$ es de la forma $|\tilde{\psi}_k| = \dfrac{1}{\sqrt{N}}$ (totalmente deslocalizado en k).

A partir de ahora suponemos un espacio de n qubits, $N = 2^n$ y $|x\rangle = |x_1 x_2 \ldots x_n\rangle$, $x_j \in \{0,1\}$ ($x = 2^{n-1}x_1 + \cdots + 2^{n-j}x_j + \cdots + 2^0 x_n$ y $0 \le x \le N-1$).

Queremos implementar el operador U_F mediante un circuito. Resulta que esto puede hacerse de manera eficiente con puertas de 1 y 2 qubits. Introducimos la notación

$$0.a_1 a_2 \ldots a_j \equiv \frac{1}{2} a_1 + \frac{1}{2^2} a_2 + \cdots \frac{1}{2^j} a_j \,, \tag{7.79}$$

[7.21] En física cuántica se habla de representación de posiciones y momentos, de un mismo estado, ese es el punto de vista pasivo.

es decir es un número fraccionario en binario. En particular $0.x_1x_2\ldots x_n = 2^{-n}x$.

La acción de U_F sobre la base computacional toma la forma

$$U_F|x\rangle = \frac{1}{\sqrt{N}}\left(|0\rangle + e^{2\pi i 0.x_n}|1\rangle\right)_1 \otimes \cdots \left(|0\rangle + e^{2\pi i 0.x_{n-j+1}\ldots x_n}|1\rangle\right)_j \otimes \cdots \left(|0\rangle + e^{2\pi i 0.x_1\ldots x_n}|1\rangle\right)_n . \tag{7.80}$$

Esta fórmula es notable porque U_F no es un operador separable pero demuestra que $U_F|x\rangle$ sí lo es, al igual que $|x\rangle$.

La demostración se basa en notar que multiplicar por 2^{-j} desplaza los dígitos binarios j lugares a la derecha, entonces $x = x_1\ldots x_n.0$ pasa a

$$2^{-j}x = \underbrace{x_1x_2\ldots x_{n-j}}_{n-j} . \underbrace{x_{n-j+1}\ldots x_n}_{j}, \tag{7.81}$$

esto implica que, $0.x_{n-j+1}\ldots x_n = \dfrac{x}{2^j} + m,\ m\in\mathbb{Z}$ y por tanto

$$e^{2\pi i 0.x_{n-j+1}\ldots x_n} = e^{2\pi i x/2^j} . \tag{7.82}$$

Aplicamos esta observación en la expresión de la derecha en (7.80)

$$\begin{aligned} \text{RHS} &= \frac{1}{\sqrt{N}}\bigotimes_{j=1}^{n}\left(|0\rangle_j + e^{2\pi i x/2^j}|1\rangle_j\right) = \frac{1}{\sqrt{N}}\bigotimes_{j=1}^{n}\left(\sum_{k_j=0}^{1} e^{2\pi i x k_j/2^j}|k_j\rangle_j\right) \\ &= \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} e^{2\pi i x\sum_{j=1}^{n}k_j2^{-j}}|k_1\ldots k_n\rangle = \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} e^{2\pi i x k/N}|k\rangle = U_F|x\rangle . \qquad \square \end{aligned} \tag{7.83}$$

Usando la identidad (7.80) se puede implementar el operador U_F mediante el circuito de la Fig. 7.10. Ahí se ha definido $R_k \equiv \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^k} \end{pmatrix}$. El circuito mostrado produce $U_F|x\rangle$ pero con los factores en orden traspuesto. Hay que completarlo con puertas SWAP para ordenarlos correctamente. [7.22]

El circuito para U_F^{-1} es completamente similar, también vale aplicar la construcción anterior invirtiendo el orden.

[7.22] Nielsen y Chuang, pág. 218.

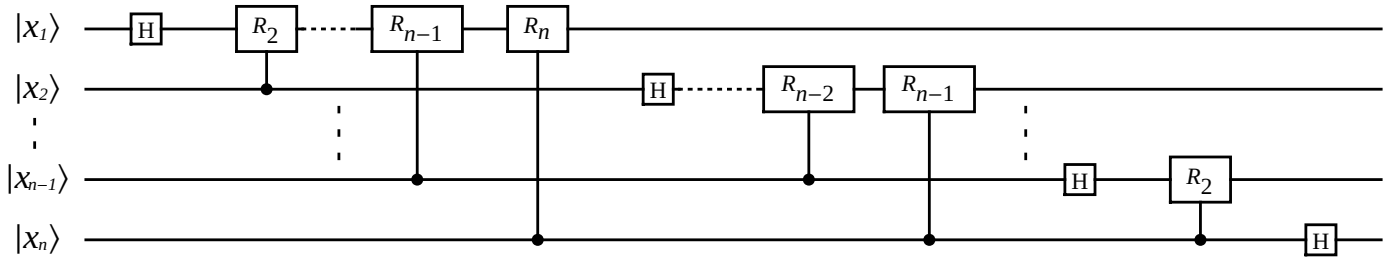


Figura 7.10: Circuito para la transformada de Fourier cuántica. (Posteriormente hay que invertir el orden de los qubits.)

La implementación requiere $O(n^2)$ puertas. Clásicamente calcular $\sum_x \omega_N^{kx} \psi_x$ requiere $O(N^2)$ operaciones (N por la suma sobre x y N al hacerlo para cada valor k). Esto es si se hace directamente. Aprovechando que N es de la forma 2^n y la estructura de la matriz $(U_F)_{kx} \propto \omega_N^{kx}$, se puede usar el algoritmo de FFT (**Fast Fourier Transform**) que requiere sólo $O(N \log N)$ operaciones, esto es $O(n2^n)$. En principio la transformada de Fourier cuántica es mucho más eficiente, es una ganancia de exponencial (clásico) a polinómico (cuántico). En la práctica hay dos dificultades para para aprovechar esta ventaja de la transformada cuántica, una es la codificación de ψ_x en el circuito, la otra es la extracción de $\tilde{\psi}_k$ una vez computado como una superposición cuántica. La mayor aplicación de la transformada de Fourier cuántica es como parte de otros algoritmos, tales como el de factorización de Shor.

7.5.3. Estimación de fases

Veamos cómo utilizar la transformada de Fourier cuántica para estimar autovalores de operadores unitarios.

Se tiene un operador unitario U que actúa en un espacio $\mathscr{H}$, y tal que

$$U|\psi\rangle = \Omega|\psi\rangle \qquad |\Omega| = 1\,. \tag{7.84}$$

Disponemos de una copia del estado $|\psi\rangle$ y de puertas $C(U^k)$ (esto es, U^k controlado) para $k = 1, 2, \ldots, 2^{m-1}$. El problema es determinar el autovalor al nivel de m bits (m dígitos binarios de ϕ, siendo $\Omega = e^{2\pi i\phi}$). El circuito a utilizar se muestra en la Fig. 7.11

Tenemos m qubits de control que actúan sobre el espacio $\mathscr{H}$. El registro $\mathscr{H}$ se inicializa en el

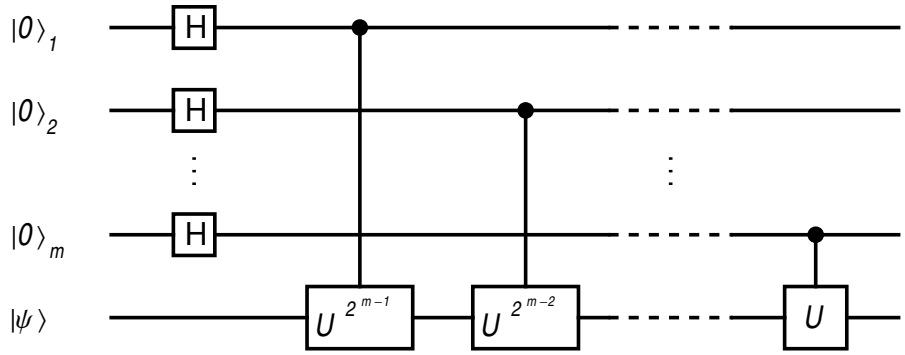


Figura 7.11: Circuito para estimación de fase.

estado $|\psi\rangle$. Los m qubits empiezan en estado $|0\rangle$ y sobre ellos actúa $H^{\otimes m}$. Se obtiene entonces

$$|\Psi_1\rangle = H^{\otimes m}|0\rangle \otimes |\psi\rangle = \frac{1}{\sqrt{M}} \bigotimes_{j=1}^{m} (|0\rangle_j + |1\rangle_j) \otimes |\psi\rangle, \qquad M \equiv 2^m . \tag{7.85}$$

Sobre este estado se aplican los operadores $C_j(U^{2^{m-j}})$ es decir $U^{2^{m-j}}$ en $\mathscr{H}$ controlado por el qubit j, para $j = 1, \ldots, m$. Estos operadores conmutan y, teniendo en cuenta que $U^\ell|\psi\rangle = \Omega^\ell|\psi\rangle$, producen

$$\begin{aligned} |\Psi_2\rangle &= \frac{1}{\sqrt{M}} \bigotimes_{j=1}^{m} (|0\rangle_j + \Omega^{2^{m-j}}|1\rangle_j) \otimes |\psi\rangle = \frac{1}{\sqrt{M}} \sum_{k=0}^{M-1} \Omega^{\sum_{j=1}^{m} 2^{m-j} k_j} |k_1 \ldots k_m\rangle \otimes |\psi\rangle \\ &= \frac{1}{\sqrt{M}} \sum_{k=0}^{M-1} \Omega^k |k\rangle \otimes |\psi\rangle = |\chi\rangle \otimes |\psi\rangle , \end{aligned} \tag{7.86}$$

usando $|k\rangle = \bigotimes_{j=1}^{m} |k_j\rangle_j = |k_1 \ldots k_m\rangle$ y $k = 2^{m-1}k_1 + \cdots + 2^0 k_m$.

En el estado final, $|\psi\rangle$ queda separado (no se entrelaza con los qubits de control). El factor en $\mathbb{C}^M$ es [7.23]

$$|\chi\rangle \equiv \frac{1}{\sqrt{M}} \sum_{k=0}^{M-1} \Omega^k |k\rangle . \tag{7.87}$$

Después de aplicar el circuito, el protocolo continúa aplicando U_F^{-1} a $|\chi\rangle$ y midiendo los m qubits.

[7.23] Nótese que aquí la suma es sobre $0 \le k < M$ y no otro rango de k de longitud M, porque esos son los valores que se obtienen al escribir k como $k_1 \ldots k_m$ en binario.

Recordemos la acción de U_F y U_F^{-1},

$$U_F|x\rangle = \frac{1}{\sqrt{M}} \sum_{k=0}^{M-1} e^{2\pi i k x/M} |k\rangle, \qquad U_F^{-1}|k\rangle = \frac{1}{\sqrt{M}} \sum_{x=0}^{M-1} e^{-2\pi i k x/M} |x\rangle. \tag{7.88}$$

En el supuesto de que

$$\Omega = e^{2\pi i \bar{x}/M} \qquad \bar{x} \in \mathbb{Z} \qquad 0 \leq \bar{x} \leq M-1\,, \tag{7.89}$$

Ω sencillamente se determina aplicando U_F^{-1} sobre $|\chi\rangle$, ya que

$$|\chi\rangle = \frac{1}{\sqrt{M}} \sum_{k=0}^{M-1} e^{2\pi i k \bar{x}/M} |k\rangle = U_F|\bar{x}\rangle. \tag{7.90}$$

$U_F^{-1}|\chi\rangle = |\bar{x}\rangle$ y toda la amplitud se concentra en el estado $x = \bar{x}$.

La condición $\Omega = e^{2\pi i \bar{x}/M}$ equivale a decir que $\Omega^M = 1$. En el caso general Ω^M no será exactamente 1, pero para M suficientemente grande habrá un $\bar{\Omega}$ tan próximo como se desee a Ω tal que $\bar{\Omega}^M = 1$. Como $\bar{x}$ tiene m bits, ϕ se va determinar con una precisión de a lo sumo m dígitos binarios.

Definimos

$$\Omega = e^{2\pi i \phi}, \qquad \phi_x := \frac{x}{M}, \qquad \delta_x := \phi - \phi_x. \tag{7.91}$$

ϕ sólo está bien definido módulo 1. Podemos ahora calcular el efecto de U_F^{-1},

$$|w\rangle := U_F^{-1}|\chi\rangle = \frac{1}{M} \sum_{x,k=0}^{M-1} e^{-2\pi i k x/M} \Omega^k |x\rangle = \frac{1}{M} \sum_{x=0}^{M-1} \sum_{k=0}^{M-1} e^{2\pi i k \delta_x} |x\rangle. \tag{7.92}$$

La amplitud del estado $|x\rangle$ es una serie geométrica finita, que puede sumarse con

$$\sum_{\ell=\alpha}^{\beta} \xi^\ell = \frac{\xi^{\beta+1} - \xi^\alpha}{\xi - 1} \qquad \xi \in \mathbb{C} \setminus \{1\}, \quad \alpha, \beta \in \mathbb{Z}, \quad \alpha \leq \beta\,. \tag{7.93}$$

Entonces

$$\langle x|w\rangle = \frac{1}{M} \sum_{k=0}^{M-1} e^{2\pi i k \delta_x} = \frac{1}{M} \frac{e^{2\pi i M \delta_x} - 1}{e^{2\pi i \delta_x} - 1} = \frac{1}{M} \frac{\operatorname{sen}(\pi M \delta_x)}{\operatorname{sen}(\pi \delta_x)} e^{i\pi(M-1)\delta_x}\,, \tag{7.94}$$

y al medir x

$$\text{Prob}(x) = |\langle x|w\rangle|^2 = \frac{1}{M^2} \frac{\operatorname{sen}^2(\pi(\phi M - x))}{\operatorname{sen}^2(\pi(\phi - x/M))}\,. \tag{7.95}$$

Esta función tendría un máximo (alcanzaría la unidad) en $x = M\phi$, si x fuera una variable continua. De entre los valores permitidos el máximo corresponde al $\bar{x} \in \{0, \ldots, M-1\}$ más próximo, de modo que $|M\phi - \bar{x}| \leq \frac{1}{2}$. [7.24] Es decir,

$$\bar{\phi} := \frac{\bar{x}}{M}, \qquad \delta := \phi - \bar{\phi}, \qquad |\delta| \leq \frac{1}{2M}. \tag{7.96}$$

(Por tanto $\bar{\phi}$ tiene m dígitos binarios correctos de ϕ.) La probabilidad satisface dos propiedades interesantes

$$\begin{aligned} &\text{Prob}(x = \bar{x}) \geq \frac{4}{\pi^2} = 0.405 \\ &\text{Prob}(|x - \bar{x}| > k) \leq \frac{1}{2(k-1)} \qquad \forall k > 0. \end{aligned} \tag{7.97}$$

Demostramos sólo la primera desigualdad. [7.25] Usando las relaciones

$$0 \leq \alpha \leq \frac{\pi}{2} \implies \frac{2\alpha}{\pi} \leq \text{sen}(\alpha) \leq \alpha, \tag{7.98}$$

se deduce

$$\pi M|\delta| \leq \frac{\pi}{2} \implies \text{sen}(\pi M|\delta|) \geq 2M|\delta|, \qquad \text{sen}(\pi|\delta|) \leq \pi\delta, \tag{7.99}$$

que implican

$$\text{Prob}(\bar{x}) = \frac{1}{M^2}\frac{\text{sen}^2(\pi M\delta)}{\text{sen}^2(\pi\delta)} \geq \frac{1}{M^2}\frac{(2M\delta)^2}{(\pi\delta)^2} = \frac{4}{\pi^2}. \tag{7.100}$$

Nótese que en este problema no se puede "comprobar la solución" a posteriori (que el x medido sea $\bar{x}$) al menos directamente. Se puede repetir el proceso para disminuir la probabilidad de error.

Podemos usar la segunda desigualdad reescalando todo con $1/M$ y $\varepsilon := k/M$,

$$\text{Prob}(|\phi_x - \bar{\phi}| > \varepsilon) \leq \frac{1}{2(M\varepsilon - 1)}. \tag{7.101}$$

Así, si quiere $\varepsilon = 1/2^n$ (n dígitos binarios de precisión en ϕ) con probabilidad al menos $1 - \eta$

$$\text{Prob}(|\phi_x - \bar{\phi}| > 2^{-n}) \leq \eta, \tag{7.102}$$

[7.24] Si un valor de ϕ está muy próximo a 1 siempre se puede se puede usar $\phi - 1$ de modo que la diferencia $|M\phi - \bar{x}|$ no supere $1/2$.

[7.25] Para la segunda véase Nielsen y Chuang, pág. 224.

basta elegir m tal que

$$\text{Prob} \leq \frac{1}{2(2^m 2^{-n}-1)} \leq \eta \implies m > n + \log(1+\frac{1}{2\eta}). \tag{7.103}$$

También es interesante notar que si $|\psi\rangle$ no es un estado propio sino una combinación de ellos

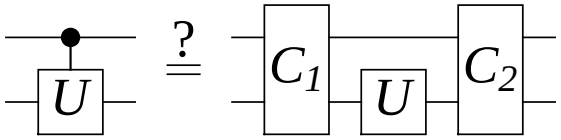
$$|\psi\rangle = \sum_u \psi_u |u\rangle, \qquad U|u\rangle = \Omega_u |u\rangle \tag{7.104}$$

al aplicar el algoritmo se tendrá (ya que todo es lineal, paralelismo cuántico)

$$U_F^{-1} \sum_u \psi_u |\chi_u\rangle \otimes |u\rangle \approx \sum_u \psi_u |\bar{\phi}_u\rangle \otimes |u\rangle \tag{7.105}$$

siendo $e^{2\pi i \bar{\phi}_u}$ una buena estimación de Ω_u.

Respecto del circuito Fig. 7.11, hay que hacer una observación. Matemáticamente, la acción de un operador U determina unívocamente la acción de $C(U)$ (U controlado). Podría parecer entonces que dada una realización física de la puerta U (U es una caja negra) automáticamente se tiene también $C(U)$. En realidad no es así, no hay una construcción (circuito fijo) en la que uno introduzca U y produzca el efecto de $C(U)$. No hay algo del tipo



con estructuras C_1, C_2 universales. Esto es obvio porque implicaría $C(e^{i\alpha}U) \stackrel{?}{=} e^{i\alpha}C(U)$, que no es cierto. Las puertas son objetos clásicos que actúan o no (o con cierta probabilidad) pero no condicionados a una amplitud de probabilidad. [7.26]

Para un U actuando *sobre un qubit*, un método de construcción de $C(U)$ es expresar $U = e^{i\alpha}AXBXC$ donde X es la puerta NOT y A, B, C son operadores unitarios tales que $ABC = I$ (se puede probar que siempre existen $\forall U$). [7.27] A continuación se puede aplicar la construcción de la Fig. 7.12 donde $S_\alpha = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{pmatrix}$.

[7.26] Sin embargo sí es posible hacer $C(U)$ con un U caja negra usando fotones en vez de circuitos.
[7.27] Nielsen y Chuang, pág. 180.

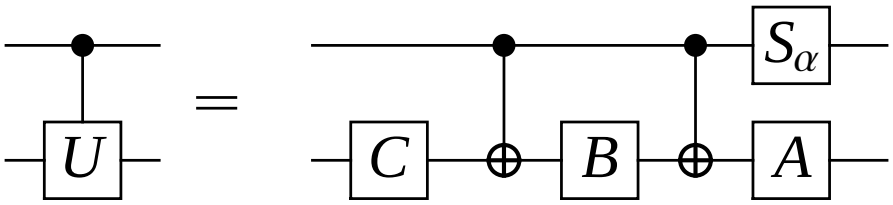


Figura 7.12: Construcción de $C(U)$ para $U = e^{i\alpha}AXBXC,\ ABC = I$ en el espacio de un qubit.

Una vez construido $C(U)$ se tiene también $C(U^k)$ iterando $C(U)$: [7.28]

$$\begin{aligned} C(U) &= |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U \\ C(U)|x\rangle \otimes |\psi\rangle &= |x\rangle \otimes (\delta_{x0}|\psi\rangle + \delta_{x1}U|\psi\rangle) \\ C(U^k)|x\rangle \otimes |\psi\rangle &= |x\rangle \otimes (\delta_{x0}|\psi\rangle + \delta_{x1}U^k|\psi\rangle) = (C(U))^k|x\rangle \otimes |\psi\rangle\,. \end{aligned} \tag{7.106}$$

También se ve escribiendo $C(U)$ como un operador diagonal el bloques $\begin{pmatrix} I & 0 \\ 0 & U \end{pmatrix}$, el primer bloque corresponde al estado $|0\rangle$ y el segundo a $|1\rangle$ del qubit de control.

7.5.4. Recuento cuántico

Un tema relacionado con los algoritmos de Grover y de estimación de fases es el problema de **recuento cuántico**. Dada una función $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$, se trata de contar el número de valores M de x tal que $f(x) = 1$. Consideremos brevemente el caso en el que se sabe que o bien $M = 0$ o bien $M = 1$.

Un método es aplicar Grover. Si se obtiene consistentemente la misma estimación $x = a$ debe ser $M = 1$ (y se puede comprobar a posteriori). Si el valor del hipotético a varía cada vez, deber ser $M = 0$.

Un método alternativo es el siguiente. El operador de Grover es de la forma $G = R_h\hat{U}_f$. Cuando $f \equiv 0\ \ \hat{U}_f = I$ y por tanto G es una reflexión, con autovalores ± 1 y concretamente $|h\rangle$ es un autovector con valor propio $+1$. En cambio, cuando $f(x) = \delta_{x,a}$ (para cierto a) G es una rotación de ángulo $2\alpha = O(1/\sqrt{N})$. En el espacio complejo bidimensional subtendido por $|h\rangle$ y $|a\rangle$, que es el único que interviene (y también vale cuando $M = 0$) la matriz de G es (en la base $|h\rangle, |h^\perp\rangle$)

$$G = \begin{pmatrix} \cos(2\alpha) & -\operatorname{sen}(2\alpha) \\ \operatorname{sen}(2\alpha) & \cos(2\alpha) \end{pmatrix} \qquad (M = 1) \tag{7.107}$$

[7.28] Aunque esto es correcto, en el caso de un qubit podría ser más eficiente construir $C(U^k)$ usando directamente la descomposición ABC concreta de U^k.

con autovalores $e^{\pm 2i\alpha}$, y vectores propios $|\alpha_\pm\rangle = (|h\rangle \pm i|h^\perp\rangle)/\sqrt{2}$.

Se puede entonces aplicar el algoritmo de determinación de fases con $|\psi\rangle = |h\rangle$ y $U = G$. El vector $|h\rangle$ no es propio pero es superposición de dos vectores propios, $|h\rangle = (|\alpha_+\rangle + |\alpha_-\rangle)/\sqrt{2}$. El algoritmo producirá un estado $(|\bar{\phi}_+\rangle \otimes |\alpha_+\rangle + |\bar{\phi}_-\rangle \otimes |\alpha_-\rangle)/\sqrt{2}$, con estimaciones $\bar{\phi}_\pm$ de las fases de los autovalores. Al hacer una medida se obtendrá uno de los dos. Dado que $\alpha = O(1/\sqrt{N})$, es decir, los primeros $n/2$ dígitos "decimales" de α se anulan, es necesario usar una precisión $m > n/2$ para ver si α es 0 o no. Si $\alpha = 0$ estamos en el caso $M = 0$, si no en $M = 1$.

El interés de este método alternativo a Grover es que en vez de $O(2^{n/2})$ evaluaciones, basta un número polinómico en n. A cambio es necesario disponer de la puertas controladas $C(U^{2^j})$. El método se puede extender a otros valores de M. En ese caso, la forma del espectro de G depende de M y su determinación permite estimar M. [7.29]

[7.29] Nielsen y Chuang, pág. 261.

8. Máquinas cuánticas

8.1. Introducción

Una máquina cuántica es una colección de puertas actuando sobre un sistema de qubits para realizar una serie de tareas dentro de un circuito mayor (es como un subprograma). La máquina puede ser fija (siempre hace la misma tarea sobre un estado cuántico que contiene los datos) o ser programable (además de datos acepta un estado cuántico como programa).

En particular veremos clonadores aproximados y limitaciones sobre el rendimiento en máquinas programables.

8.2. Clonadores y puerta UNOT

8.2.1. Máquina para clonación aproximada

La clonación perfecta (y universal) no es viable (no se puede clonar una superposición) pero sí es posible hacer clonaciones aproximadas. Como se vio al hablar del teorema de no clonación (Tema 4.3.1) se puede hacer una clonación exacta para estados de una base, (4.76), que funcionará bien (mayor fidelidad) para estados próximos a la base computacional y peor para estados $|\pm\rangle$ en el caso de un qubit. Queremos ahora examinar un esquema que tenga una fidelidad uniforme, independientemente del estado a clonar.

Supongamos que Carlos tiene un estado $|\psi\rangle$ de un qubit y querría enviar una copia a Andrea y otra a Benito. Definamos las matrices de estados puros

$$\rho_\psi = |\psi\rangle\langle\psi|, \qquad \rho_{\psi^\perp} = |\psi^\perp\rangle\langle\psi^\perp| = I - \rho_\psi\,, \tag{8.1}$$

siendo

$$|\psi\rangle = \begin{pmatrix}\alpha\\ \beta\end{pmatrix} \qquad |\psi^\perp\rangle = \begin{pmatrix}-\beta^*\\ \alpha^*\end{pmatrix}. \tag{8.2}$$

Idealmente Carlos querría producir el estado

$$\rho_{12}^{\text{clon}} = \rho_\psi \otimes \rho_\psi \qquad \text{(clonación perfecta)}\,, \tag{8.3}$$

pero lo prohíbe el teorema de no clonación. De hecho se puede ver fácilmente que no existe un hipotético canal cuántico tal que

$$\forall|\psi\rangle \qquad \rho_\psi \to T(\rho_\psi) = \rho_\psi \otimes \rho_\psi. \tag{8.4}$$

Basta considerar la mezcla

$$\frac{1}{2}(\rho_\psi + \rho_{\psi^\perp}) = \frac{1}{2} I\,. \tag{8.5}$$

Después de mezclar la dependencia en $|\psi\rangle$ desaparece. Sin embargo su imagen

$$\frac{1}{2}(\rho_\psi \otimes \rho_\psi + \rho_{\psi^\perp} \otimes \rho_{\psi^\perp}) \tag{8.6}$$

sí dependería de $|\psi\rangle$. [8.1] Por tanto tal canal de clonación no existe.

Carlos tiene que producir una copia imperfecta. Un primer protocolo, extremadamente pobre, es producir el estado mezcla

$$\rho''_{12} = \rho_{\text{azar}} \otimes \rho_{\text{azar}} \qquad \rho_{\text{azar}} \equiv \frac{1}{2} I\,, \tag{8.7}$$

y enviar los registros 1 y 2 a Andrea y Benito, respectivamente. Este protocolo descarta la información sobre $|\psi\rangle$ de entrada, aunque tiene la virtud de que Carlos se puede quedar con $|\psi\rangle$. Aun así, Andrea y Benito reciben las matrices reducidas

$$\rho''_1 = \rho''_2 = \rho_{\text{azar}} = \frac{1}{2}\rho_\psi + \frac{1}{2}\rho_{\psi^\perp}\,, \tag{8.8}$$

(las matrices ρ''_1 y ρ''_2 son iguales, pero cada una en su espacio) siendo

$$\rho''_1 = \text{Tr}_2(\rho''_{12}), \qquad \rho''_2 = \text{Tr}_1(\rho''_{12})\,. \tag{8.9}$$

A la vista de (8.8) se puede decir que cada parte tiene un 50 % de probabilidad de obtener el estado correcto (y otro tanto de recibir el estado ortogonal). Igualmente correcto es decir que el qubit que reciben es completamente aleatorio, todo es ruido y nada señal.

Respecto a las fidelidades $F(\rho''_{12}, \rho^{\text{clon}}_{12}) = \frac{1}{4}$ y $F''_1 = F(\rho''_1, \rho_\psi) = \frac{1}{2}$. Podemos medir la *eficiencia* del método mediante la **fidelidad media**; la media de las fidelidades de los qubits recibidos por Andrea y Benito respecto del ideal $|\psi\rangle$,

$$F'' = \frac{1}{2}(F''_1 + F''_2) = \frac{1}{2}\,. \tag{8.10}$$

[8.1] Sí hay cancelación de la dependencia en $\frac{1}{4}(\rho_\psi \otimes \rho_\psi + \rho_\psi \otimes \rho_{\psi^\perp} + \rho_{\psi^\perp} \otimes \rho_\psi + \rho_{\psi^\perp} \otimes \rho_{\psi^\perp}) = \frac{1}{4} I \otimes I$.

Si no se requiere que Carlos conserve $|\psi\rangle$, un protocolo más lógico, aunque no sofisticado, es que con probabilidad 1/2 Carlos envíe $|\psi\rangle$ a Andrea y un qubit al azar a Benito, o viceversa, es decir:

$$\begin{aligned}\rho'_{12} &= \frac{1}{2}\rho_\psi \otimes \rho_{\text{azar}} + \frac{1}{2}\rho_{\text{azar}} \otimes \rho_\psi \qquad \text{(protocolo no sofisticado)} \\ &= \frac{1}{2}\rho_\psi \otimes \rho_\psi + \frac{1}{4}\rho_\psi \otimes \rho_{\psi^\perp} + \frac{1}{4}\rho_{\psi^\perp} \otimes \rho_\psi \,.\end{aligned} \tag{8.11}$$

Para esta mezcla ρ'_{12}, lo que reciben Andrea y Benito es

$$\rho'_1 = \rho'_2 = \frac{1}{2}\,\rho_\psi + \frac{1}{2}\,\rho_{\text{azar}} = \frac{3}{4}\,\rho_\psi + \frac{1}{4}\,\rho_{\psi^\perp}\,. \tag{8.12}$$

Esto es un 50 % de probabilidad de obtener $|\psi\rangle$ y otro tanto de un estado al azar, o equivalentemente, un 75 % de probabilidad de obtener el estado correcto y 25 % el incorrecto. $F(\rho'_{12}, \rho_{12}^{\text{clon}}) = \frac{1}{2}$ y $F(\rho'_1, \rho_\psi) = \frac{3}{4}$. [8.2] Es decir, para este protocolo

$$F' = \frac{3}{4}\,. \tag{8.13}$$

El protocolo ρ'_{12} es esencialmente clásico. Es el mismo que se usaría con bits si por algún motivo Carlos recibe un bit y no puede clonarlo (por ejemplo, no lo puede ver) y quiere una "clonación aproximada". El protocolo no utiliza superposición cuántica. Veamos que usando superposición se puede mejorar el resultado (análogo a lo que ocurre con la violación cuántica de las desigualdades de Bell).

Querríamos producir un estado $|\psi,\psi\rangle_{12}$ a partir de $|\psi\rangle$, pero esto no es posible ya que tal aplicación no es lineal. Aunque los estados estén normalizados, la linealidad se viola porque si se cambia la fase de $|\psi\rangle$ por un factor ω ($|\omega|^2 = 1$) el cambio en $|\psi,\psi\rangle_{12}$ es un factor ω^2. Este problema se puede corregir considerando un qubit auxiliar, de modo que el estado producido sea del tipo $|\psi,\psi,\psi^\perp\rangle_{123}$ (el estado $|\psi^\perp\rangle$ se transforma con ω^*). Teniendo en cuenta la identidad

$$|\Phi_-\rangle = \frac{1}{\sqrt{2}}(|\psi,\psi^\perp\rangle - |\psi^\perp,\psi\rangle) \qquad \forall|\psi\rangle \tag{8.14}$$

el problema de linealidad se resuelve completamente produciendo un estado de tipo $|\psi\rangle_1 \otimes |\Phi_-\rangle_{23}$: es lineal porque $|\Phi_-\rangle_{23}$ no depende de $|\psi\rangle$ y contiene el estado clonado $|\psi,\psi\rangle_{12}$. [8.3] Igualmente

[8.2] Por supuesto, las expresiones de ρ'_{12} y ρ'_1 cumplen que la dependencia en $|\psi\rangle$ se cancela si se suma la misma expresión con $|\psi\rangle \to |\psi^\perp\rangle$.

[8.3] Obviamente, un estado del tipo $\langle\psi|0\rangle|\psi,\psi\rangle$ también produce una fase ω en vez de ω^2 pero estamos buscando una clonación aproximada con fidelidad uniforme, lo cual requiere coeficientes constantes, independientes de $|\psi\rangle$.

podemos considerar el estado en el que el qubit sale con seguridad por el canal 2,

$$\begin{aligned} |\chi_1\rangle_{123} &= |\psi\rangle_1 \otimes |\Phi_-\rangle_{23}, \\ |\chi_2\rangle_{123} &= |\psi\rangle_2 \otimes |\Phi_-\rangle_{13}. \end{aligned} \tag{8.15}$$

En el primer estado Andrea recibe el qubit correcto probabilidad (fidelidad) 1, y Benito lo recibe con probabilidad $1/2$, y viceversa para el segundo estado. Si se hace una **mezcla** de estos dos estados al 50% (o de hecho con cualquier otro peso) la fidelidad media es como antes, $3/4$. Pero se puede obtener una mejora haciendo una **superposición** en lugar de una mezcla:

$$\begin{aligned} |\chi\rangle_{123} &= c_1|\chi_1\rangle + c_2|\chi_2\rangle \\ &= \frac{c_1+c_2}{\sqrt{2}}|\psi,\psi,\psi^\perp\rangle - \frac{c_1}{\sqrt{2}}|\psi,\psi^\perp,\psi\rangle - \frac{c_2}{\sqrt{2}}|\psi^\perp,\psi,\psi\rangle \end{aligned} \tag{8.16}$$

y la condición de normalización es $\frac{1}{2}|c_1+c_2|^2 + \frac{1}{2}|c_1|^2 + \frac{1}{2}|c_2|^2 = 1$.

En $|\chi\rangle_{123}$ no hay contribuciones de tipo $|\psi^\perp\psi^\perp\rangle_{12}$. La probabilidad del estado ideal $|\psi,\psi\rangle_{12}$, en el que Andrea y Benito obtienen el qubit correcto es $\frac{1}{2}|c_1+c_2|^2$. Para Andrea la probabilidad de obtener el qubit correcto se aumenta con la probabilidad del estado $|\psi,\psi^\perp\rangle_{12}$, a saber, $|c_1|^2/2$,

$$F_1 = \frac{1}{2}(|c_1+c_2|^2) + \frac{1}{2}|c_1|^2 = 1 - \frac{1}{2}|c_2|^2, \tag{8.17}$$

y análogamente para Benito. El mismo resultado se obtiene con las matrices reducidas

$$\begin{aligned} \rho_1 &= (1-\frac{1}{2}|c_2|^2)\,\rho_\psi + \frac{1}{2}|c_2|^2\,\rho_{\psi^\perp} \\ \rho_2 &= (1-\frac{1}{2}|c_1^2|)\,\rho_\psi + \frac{1}{2}|c_1^2|\,\rho_{\psi^\perp}. \end{aligned} \tag{8.18}$$

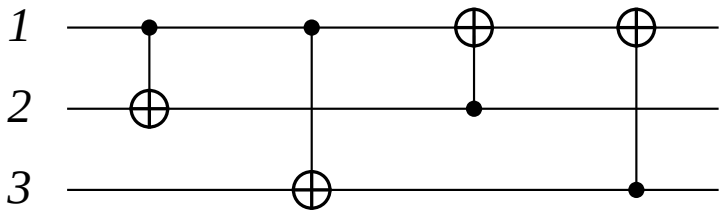


Figura 8.1: Circuito cuántico para una máquina clonadora.

Para construir el estado (8.16) se puede usar el circuito de la Fig. 8.1, con tres registros de un qubit cada uno. Este circuito es clásico (no incluye puertas con fases, no se sale de la base computacional) y es fácil ver que su efecto es

$$|x_1,x_2,x_3\rangle \longrightarrow |x_1\oplus x_2\oplus x_3, x_1\oplus x_2, x_1\oplus x_3\rangle\,. \tag{8.19}$$

Consideremos dos posibles inputs para los registros 2 y 3,

$$\begin{aligned} |\Psi_+\rangle_{23} &= \frac{1}{\sqrt{2}}(|0\rangle_2|0\rangle_3 + |1\rangle_2|1\rangle_3), \\ |0\rangle_2|+\rangle_3 &= \frac{1}{\sqrt{2}}(|0\rangle_2|0\rangle_3 + |0\rangle_2|1\rangle_3), \end{aligned} \tag{8.20}$$

junto con el qubit $|\psi\rangle$ en el registro 1.

Al aplicar el circuito con cada una de las dos entradas se obtiene: [8.4]

$$\begin{aligned} |\psi\rangle_1\otimes|\Psi_+\rangle_{23} &\longrightarrow |\psi\rangle_1\otimes|\Psi_+\rangle_{23}\,, \\ |\psi\rangle_1\otimes|0,+\rangle_{23} &\longrightarrow |\psi\rangle_2\otimes|\Psi_+\rangle_{13}\,. \end{aligned} \tag{8.21}$$

En el primer caso $|\psi\rangle$ permanece en el registro 1 y en el segundo la información contenida en $|\psi\rangle$ se ha desplazado del registro 1 al 2. El primer caso es como un circuito $|\,\rangle_{12}\to|\,\rangle_{12}$ que no hace nada (la identidad) y el segundo es un SWAP. Los circuitos son objetos clásicos y no se pueden superponer pero aquí se consigue hacerlo introduciendo un qubit auxiliar. Consideramos un estado del tipo

$$c_1|\Psi_+\rangle_{23} + c_2|0,+\rangle_{23}, \qquad \frac{1}{2}(|c_1+c_2|^2+|c_1|^2+|c_2|^2) = 1\,. \tag{8.22}$$

Este estado está normalizado, por $\langle\Psi_+|0,+\rangle = \frac{1}{2}$. Después de aplicar el circuito se obtiene

$$|F\rangle_{123} \equiv c_1|\psi\rangle_1\otimes|\Psi_+\rangle_{23} + c_2|\psi\rangle_2\otimes|\Psi_+\rangle_{13}, \tag{8.23}$$

[8.4] En detalle,

$$\begin{aligned} |\psi\rangle_1\otimes|\Psi_+\rangle_{23} &= (\alpha|0\rangle+\beta|1\rangle)\frac{1}{\sqrt{2}}(|00\rangle+|11\rangle) \longrightarrow \alpha|0\rangle|\Psi_+\rangle+\beta|1\rangle|\Phi_+\rangle \longrightarrow \alpha|0\rangle|\Psi_+\rangle+\beta|1\rangle|\Psi_+\rangle = |\psi\rangle|\Psi_+\rangle \\ &\longrightarrow |\psi\rangle\frac{1}{\sqrt{2}}|00\rangle + (\alpha|1\rangle+\beta|0\rangle)\frac{1}{\sqrt{2}}|11\rangle \longrightarrow |\psi\rangle\frac{1}{\sqrt{2}}|00\rangle + |\psi\rangle\frac{1}{\sqrt{2}}|11\rangle = |\psi\rangle_1|\Psi_+\rangle_{23} \\ |\psi\rangle_1\otimes|0\rangle_2|+\rangle_3 &\longrightarrow \alpha|00+\rangle+\beta|11+\rangle \longrightarrow \alpha|00+\rangle+\beta|11+\rangle \longrightarrow \alpha|00+\rangle+\beta|01+\rangle \\ &\longrightarrow \frac{1}{\sqrt{2}}\alpha(|000\rangle+|101\rangle)+\frac{1}{\sqrt{2}}\beta(|010\rangle+|111\rangle) = |\psi\rangle_2|\Psi_+\rangle_{13} \end{aligned}$$

que es simétrico respecto del intercambio de etiquetas 1 y 2 (una simetría útil para comprobaciones). Este estado no es exactamente el de (8.16) pero este último se obtiene a partir de $|F\rangle$ aplicando una puerta $-i\sigma_y$ al tercer qubit, por $I\otimes(-i\sigma_y)|\Psi_+\rangle = |\Phi_-\rangle$.

$$(-i\sigma_y)_3|F\rangle = |\chi\rangle, \tag{8.24}$$

por tanto $|F\rangle$ es equivalente a $|\chi\rangle$ a efectos de los registros 1 y 2 (o simplemente se añade la puerta $-i\sigma_y$ al final del circuito en el tercer registro).

Volviendo a (8.18), si se elige $c_2 = 0$ (entonces $c_1^2 = 1$) el estado $|\psi\rangle$ queda copiado totalmente en 1 y 2 será aleatorio, y viceversa. Dado que $c_1 = c_2 = 0$ no es posible, no se puede copiar $|\psi\rangle$ en 1 y 2 a la vez con probabilidad 1. La fidelidad media en una clonación perfecta sería

$$F^{\text{clon}} = 1 \qquad \text{(clonación perfecta)} \tag{8.25}$$

En lugar de eso, con este protocolo se obtiene

$$F = \frac{1}{2}(1-\frac{1}{2}|c_2|^2 + 1 - \frac{1}{2}|c_2|^2) = \frac{1}{2} + \frac{1}{4}|c_1+c_2|^2 \tag{8.26}$$

y su máximo corresponde a la elección simétrica

$$c_1 = c_2, \quad |c_1| = \frac{1}{\sqrt{3}} \tag{8.27}$$

que produce

$$F = \frac{5}{6} = 0.83 \tag{8.28}$$

y para las matrices reducidas

$$\rho_1 = \rho_2 = \frac{5}{6}\rho_\psi + \frac{1}{6}\rho_{\psi^\perp} = \frac{2}{3}\rho_\psi + \frac{1}{3}\rho_{\text{azar}}\,. \tag{8.29}$$

El resultado se puede expresar diciendo que cuando Andrea mida su qubit tiene un 83 % de probabilidad de recibir $|\psi\rangle$ frente a un 17 % de recibir $|\psi^\perp\rangle$, y lo mismo Benito. O equivalentemente, Andrea tiene una probabilidad $2/3$ de obtener el qubit en el estado correcto frente a $1/3$ de un estado al azar.

Se obtiene cierta mejora al seguir el protocolo de la máquina clonadora aproximada (83 % frente a 75 %) aunque sin llegar al 100 % correspondiente a la clonación perfecta.

Nótese también que Andrea y Benito reciben un qubit con cierta probabilidad conocida de ser $|\psi\rangle$ y no $|\psi^\perp\rangle$, pero no saben cuándo es un caso y cuándo otro, es decir, no reciben una versión marcada del tipo

$$\frac{2}{3}\rho_\psi \otimes |0\rangle\langle 0| + \frac{1}{3}\rho_{\text{azar}} \otimes |1\rangle\langle 1|\,. \tag{8.30}$$

Una versión marcada en ρ_{12} del tipo

$$\frac{1}{3}\rho_\psi \otimes \rho_\psi \otimes |0\rangle\langle 0| + \frac{1}{3}\rho_\psi \otimes \rho_{\text{azar}} \otimes |1\rangle\langle 1| + \frac{1}{3}\rho_{\text{azar}} \otimes \rho_\psi \otimes |2\rangle\langle 2| \tag{8.31}$$

es inviable: si al medir la etiqueta sale $|1\rangle$ se toma el primer registro (que contiene $|\psi\rangle$) y se repite el proceso, y análogamente si sale la etiqueta $|2\rangle$, eventualmente saldrá la etiqueta $|0\rangle$ y se habrá clonado el estado, y el proceso se podría iterar para conseguir tantas copias como se desee. Esto no puede ocurrir porque viola el teorema de no clonación. [8.5]

8.2.2. Puerta UNOT

También es instructivo ver la información del qubit 3. Como ya se ha comentado anteriormente, la correcta transformación bajo un cambio de fase $|\psi\rangle \to \omega|\psi\rangle$ requiere que cada $|\psi\rangle$ que se añada en otro registro debe ir acompañado de un $|\psi^\perp\rangle$ en un tercer registro; sólo se pueden añadir *pares* $|\psi, \psi^\perp\rangle$, no $|\psi\rangle$ sueltos. Al usar superposición se ha conseguido incrementar el peso de la componente $|\psi\psi\rangle_{12}$, más allá de lo permitido clásicamente. Y eso va necesariamente acompañado de un aumento del peso de $|\psi^\perp\rangle_3$.

Este resultado es de interés por su relación con el problema de la puerta UNOT (*Universal* NOT). Esta es una puerta hipotética que transformaría un qubit en su ortogonal,

$$\text{UNOT}|\psi\rangle = |\psi^\perp\rangle, \qquad \begin{pmatrix}\alpha\\ \beta\end{pmatrix} \longrightarrow \begin{pmatrix}-\beta^*\\ \alpha^*\end{pmatrix}. \tag{8.32}$$

Obviamente esta transformación no es lineal (aunque sí es isométrica) y por tanto no es físicamente realizable de manera *automática o universal*, como también era el caso de la clonación. (Como para la clonación, sí es posible realizar UNOT si se tiene suficiente información sobre el estado o muchas copias, etc. Por ejemplo, X funciona perfectamente como negación para la base computacional pero su calidad como negación se deteriora para estados en el ecuador de la esfera de Bloch.) Pero sí es posible hacerla de manera aproximada o **probabilística**, y aquí se va a utilizar el tercer registro para ello.

Como se ha dicho, si al final del circuito de la Fig. 8.1 aplicamos la puerta

$$U_0 = -i\sigma_y, \qquad U_0|0\rangle = |1\rangle, \qquad U_0|1\rangle = -|0\rangle \tag{8.33}$$

[8.5] De hecho, la expresión en (8.31) para ρ_{12} no es consistente, si se suma con la misma expresión cambiando $|\psi\rangle$ por $|\psi^\perp\rangle$ no se cancela la dependencia en el estado, como debería.

sobre el registro 3, se produce el estado $|\chi\rangle$. Para el tercer registro se obtiene

$$\rho_3 = \mathrm{Tr}_{12}(|\chi\rangle\langle\chi|) = \frac{2}{3}\rho_{\psi^\perp} + \frac{1}{3}\rho_\psi = \frac{1}{3}\rho_{\psi^\perp} + \frac{2}{3}\rho_{\mathrm{azar}}. \tag{8.34}$$

La probabilidad de que el estado resultante en el registro 3 sea $|\psi^\perp\rangle$ es un 67 % (a comparar con el azar, que da un 50 %). De hecho se puede probar que este resultado es óptimo: para protocolos que funcionen uniformemente sobre estados genéricos no se puede imitar UNOT con más eficiencia que ésta.

Nótese que la parte 3 no sabe cuándo su qubit es realmente $|\psi^\perp\rangle$ ya que el resultado correcto no está marcado. Una matriz tal y como

$$\frac{2}{3}\rho_{\psi^\perp}|0\rangle\langle 0| + \frac{1}{3}\rho_\psi|1\rangle\langle 1| \tag{8.35}$$

no se puede conseguir. El motivo es que midiendo la etiqueta, si sale $|1\rangle$ se repite el proceso hasta obtener $|0\rangle$, eso equivale a obtener $\rho_{\psi^\perp}$ con 100 % de probabilidad, lo cual viola la fidelidad máxima $2/3$.

Veamos otro método para implementar la operación UNOT de manera aproximada y que también produce la misma solución óptima. La entrada es $|\psi\rangle$ y denominemos $|\phi\rangle$ la salida. El protocolo es

i) Se genera un punto $\hat{n}$ al azar (es decir, equidistribuido) en la esfera de Bloch, correspondiente a un estado $|\eta\rangle$. [8.6]

ii) Se mide el estado $|\psi\rangle$ con el PVM $\{|\eta\rangle\langle\eta|, |\eta^\perp\rangle\langle\eta^\perp|\}$.

iii) Si el resultado de la medida es $|\eta\rangle$ la salida $|\phi\rangle$ (nuestro $|\psi^\perp\rangle$ aproximado) es $|\eta^\perp\rangle$, y si el resultado es de la medida es $|\eta^\perp\rangle$ la salida es $|\phi\rangle = |\eta\rangle$. [8.7]

En la práctica la forma de hacer la medida será rotar $|\psi\rangle$ por una rotación U que lleve $|\eta\rangle$ a $|0\rangle$ y se medirá $U|\psi\rangle$ en la base estándar. Pero lo más lógico es directamente generar una rotación U al azar uniformemente sobre $\mathrm{SU}(2)$ (usando la medida de Haar del grupo) y medir $U|\psi\rangle$ en la base estándar. Entonces, según salga $|0\rangle$ o $|1\rangle$ el resultado será $|\phi\rangle = U^{-1}|1\rangle$ o $U^{-1}|0\rangle$. [8.8]

[8.6] Obviamente, si se tiene información sobre $|\psi\rangle$ tal que su distribución no es uniforme también se puede aumentar la eficiencia usando una distribución no uniforme adecuada para $|\eta\rangle$.

[8.7] Si al hacer la medida el resultado es $|\eta\rangle$, el estado de nuestro qubit pasará de $|\psi\rangle$ a $|\eta\rangle$ (suponiendo que sobreviva a la medida) entonces lo cambiamos a mano a $|\eta^\perp\rangle$ (o creamos otro qubit en estado $|\eta^\perp\rangle$). No hay problema en hacer esto ya que el estado $|\eta\rangle$ es conocido, a diferencia de $|\psi\rangle$. Y análogamente si el resultado de la medida es $|\eta^\perp\rangle$.

[8.8] En ningún caso se genera un $|\eta\rangle$ físico al azar (sin saber qué estado es) ya que no podríamos medirlo con $|\psi\rangle$ ni construir $|\eta^\perp\rangle$.

Para un $|\eta\rangle$ dado, la matriz densidad del estado así generado será

$$\rho(\eta) = |\langle\psi|\eta\rangle|^2 |\eta^\perp\rangle\langle\eta^\perp| + |\langle\psi|\eta^\perp\rangle|^2 |\eta\rangle\langle\eta|. \tag{8.36}$$

Falta promediar sobre $\hat{n}$ (para obtener un resultado uniforme con respecto de $|\psi\rangle$) lo cual produce el resultado

$$\rho = \frac{2}{3}|\psi^\perp\rangle\langle\psi^\perp| + \frac{1}{3}|\psi\rangle\langle\psi|, \tag{8.37}$$

que es el mismo resultado hallado en (8.34).

8.2.2.1. Cálculo del promedio de $\rho(\eta)$

Queremos calcular el promedio sobre $|\eta\rangle$, esto es sobre $\hat{\boldsymbol{n}}$ con la medida uniforme sobre la esfera $d^2\Omega_{\hat{\boldsymbol{n}}}$ (ángulo sólido)

$$\rho = \langle\rho(\eta)\rangle = \frac{1}{4\pi}\int d^2\Omega_{\hat{\boldsymbol{n}}}\,\rho(\eta) = \frac{1}{4\pi}\int_{-1}^{+1} d\cos(\theta)\int_0^{2\pi} d\phi\,\rho(\eta)\,. \tag{8.38}$$

La distribución uniforme sobre la esfera de Bloch es invariante bajo rotaciones, por tanto podemos trabajar en cualquier base y elegimos la base ortonormal $\{|\psi\rangle, |\psi^\perp\rangle\}$:

$$\begin{aligned} |\eta\rangle &= e^{-i\phi/2}\cos(\theta/2)|\psi\rangle + e^{+i\phi/2}\,\mathrm{sen}(\theta/2)|\psi^\perp\rangle \\ |\eta^\perp\rangle &= -e^{-i\phi/2}\,\mathrm{sen}(\theta/2)|\psi\rangle + e^{+i\phi/2}\cos(\theta/2)|\psi^\perp\rangle \end{aligned} \tag{8.39}$$

$$|\langle\psi|\eta\rangle|^2 = \cos^2(\theta/2), \qquad |\langle\psi|\eta^\perp\rangle|^2 = \mathrm{sen}^2(\theta/2), \tag{8.40}$$

$$\begin{aligned} |\eta\rangle\langle\eta| &= \begin{pmatrix} \cos^2(\theta/2) & e^{-i\phi}\cos(\theta/2)\,\mathrm{sen}(\theta/2) \\ e^{i\phi}\cos(\theta/2)\,\mathrm{sen}(\theta/2) & \mathrm{sen}^2(\theta/2) \end{pmatrix}, \\ |\eta^\perp\rangle\langle\eta^\perp| = I - |\eta\rangle\langle\eta| &= \begin{pmatrix} \mathrm{sen}^2(\theta/2) & -e^{-i\phi}\cos(\theta/2)\,\mathrm{sen}(\theta/2) \\ -e^{i\phi}\cos(\theta/2)\,\mathrm{sen}(\theta/2) & \cos^2(\theta/2) \end{pmatrix}. \end{aligned} \tag{8.41}$$

Para los elementos de matriz se obtiene

$$\begin{aligned} \langle\psi|\rho(\eta)|\psi\rangle &= 2\cos^2(\theta/2)\,\mathrm{sen}^2(\theta/2) = \frac{1}{2}(1-\cos^2(\theta)) \\ \langle\psi^\perp|\rho(\eta)|\psi^\perp\rangle &= \cos^4(\theta/2) + \mathrm{sen}^4(\theta/2) = \frac{1}{2}(1+\cos^2(\theta)) \\ \langle\psi^\perp|\rho(\eta)|\psi\rangle &= \langle\psi|\rho(\eta)|\psi^\perp\rangle^* = -e^{i\phi}\cos^3(\theta/2)\,\mathrm{sen}(\theta/2) + e^{i\phi}\cos(\theta/2)\,\mathrm{sen}^3(\theta/2) = -\frac{1}{2}e^{i\phi}\cos(\theta)\,\mathrm{sen}(\theta). \end{aligned} \tag{8.42}$$

$$\rho(\eta) = \frac{1}{2}\begin{pmatrix} 1-\cos^2(\theta) & -e^{-i\phi}\cos(\theta)\,\mathrm{sen}(\theta) \\ -e^{i\phi}\cos(\theta)\,\mathrm{sen}(\theta) & 1+\cos^2(\theta) \end{pmatrix} \tag{8.43}$$

Al promediar sobre ϕ sólo quedan los términos diagonales. Por otro lado el promedio sobre θ es

$$\langle \cos^2(\theta)\rangle = \frac{1}{2}\int_{-1}^{+1} \cos^2(\theta)\, d\cos(\theta) = \frac{1}{3}, \qquad \rho = \langle \rho(\eta)\rangle = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{2}{3} \end{pmatrix}, \tag{8.44}$$

que reproduce (8.37).

Una estrategia similar para clonación (generar $|\eta\rangle$ al azar, medir $|\psi\rangle$ y producir $|\eta\rangle \otimes |\eta\rangle$ o $|\eta^\perp\rangle \otimes |\eta^\perp\rangle$) no funciona bien. Produce una fidelidad $2/3$ frente a $5/6$ de la máquina de clonación aproximada o $3/4$ del envío de $|\psi\rangle$ al azar.

8.3. Máquinas programables: un resultado general

Examinamos ahora máquinas programables, también llamados procesadores cuánticos. Por una parte está el espacio $\mathscr{H}_d$ de los datos $|\psi\rangle_d$ que se van a procesar y por otra el programa que también se codifica mediante un estado cuántico $|\Xi\rangle_p$ en un espacio $\mathscr{H}_p$. Cambiando el programa el procesador puede realizar distintas tareas sin cambiar el circuito. Además enviando una superposición de estados programa se pueden realizar en paralelo.

Sin embargo hay limitaciones para construir un **procesador universal**, es decir, tal que con un input $|\Xi\rangle_p$ adecuado aplique cualquier operador unitario U sobre el estado $|\psi\rangle_d$ de los datos: el número de operadores implementable es a lo sumo la dimensión del espacio $\mathscr{H}_p$.

Para obtener este resultado, sea G el procesador, un operador unitario en el espacio producto $\mathscr{H}_d \otimes \mathscr{H}_p$, tal que $|\Xi\rangle_p$ implementa el operador U (que actúa en $\mathscr{H}_d$):

$$G|\psi\rangle_d \otimes |\Xi\rangle_p = U|\psi\rangle_d \otimes |\Xi'\rangle_p\,. \tag{8.45}$$

Se supone un estado final separable para poder extraer el estado saliente $U|\psi\rangle_d$, que es el resultado que se quiere obtener usando el procesador.

Veamos primero que los estados programa salientes $|\Xi'\rangle_p$ no dependen del estado dato $|\psi\rangle_d$ (lo cual es casi obvio por linealidad): [8.9]

[8.9] Esencialmente es el mismo principio que impide la clonación: la información cuántica contenida en $|\psi\rangle$ puede fluir y

Consideremos dos estados de datos $|\psi_1\rangle_d$ y $|\psi_2\rangle_d$ arbitrarios y dos programas $|\Xi_1\rangle_p$ y $|\Xi_2\rangle_p$, que implementan los operadores unitarios U_1 y U_2,

$$\begin{aligned} G|\psi_1\rangle_d \otimes |\Xi_1\rangle_p &= U_1|\psi_1\rangle_d \otimes |\Xi_1'\rangle_p\,, \\ G|\psi_2\rangle_d \otimes |\Xi_2\rangle_p &= U_2|\psi_2\rangle_d \otimes |\Xi_2'\rangle_p\,. \end{aligned} \tag{8.46}$$

Por ser G unitario se deduce

$$\langle\psi_1|\psi_2\rangle\langle\Xi_1|\Xi_2\rangle = \langle\psi_1|U_1^{-1}U_2|\psi_2\rangle\langle\Xi_1'|\Xi_2'\rangle\,. \tag{8.47}$$

Si aquí se elige $|\Xi_1\rangle_p = |\Xi_2\rangle_p$, y por tanto $U_1 = U_2$, y $|\psi_1\rangle_d$ y $|\psi_2\rangle_d$ son arbitrarios, se sigue que $1 = \langle\Xi_1'|\Xi_2'\rangle$, es decir $|\Xi_1'\rangle_p = |\Xi_2'\rangle_p$. Esto implica que $|\Xi\rangle_p$ determina $|\Xi'\rangle_p$, y no depende de $|\psi\rangle_d$.

Se acaba de establecer que en (8.47), los coeficientes $\langle\Xi_1|\Xi_2\rangle$ y $\langle\Xi_1'|\Xi_2'\rangle$ no dependen de los estados $|\psi_1\rangle_d$ y $|\psi_2\rangle_d$, que son arbitrarios. Entonces ambos lados de la ecuación se pueden ver como elementos de matriz $\langle\psi_1|\cdot|\psi_2\rangle$ de operadores en $\mathscr{H}_d$, y se obtiene la siguiente relación operatorial:

$$\langle\Xi_1|\Xi_2\rangle\, U_1 = \langle\Xi_1'|\Xi_2'\rangle\, U_2 \tag{8.48}$$

Hay dos posibilidades:

i) $\langle\Xi_1|\Xi_2\rangle = 0$, los programas $|\Xi_1\rangle_p$ y $|\Xi_2\rangle_p$ son ortogonales.

ii) $\langle\Xi_1|\Xi_2\rangle \neq 0$, entonces U_1 y U_2 tienen que ser proporcionales, $U_1 = e^{i\alpha}U_2$.

Se concluye entonces que cada operador U distinto requiere una dirección nueva en $\mathscr{H}_p$ y en consecuencia sólo se pueden codificar tantos operadores unitarios como sea la dimensión de $\mathscr{H}_p$.

Un corolario es que mientras nos mantengamos en espacios de dimensión finita, no puede haber un procesador universal que implemente cualquier operador. Ni siquiera una familia particular pero continua de operadores (digamos un grupo uniparamétrico) ya que tal conjunto contiene un número infinito de operadores distintos.

Esta limitación era de esperar ya que hay un conflicto (o al menos tensión) entre la linealidad en $|\Xi\rangle_p$ en la parte izquierda de (8.45) y la unitaridad de U en la parte derecha (la ligadura $U^\dagger U = I$ es no lineal, los operadores unitarios no forman un espacio vectorial). La tensión se resuelve de modo que

repartirse por varios espacios (como por ejemplo en Fig. 5.2) pero cada vez sólo puede llenar un subespacio de dimensión igual a la que originalmente ocupaba $|\psi\rangle$. Si en la parte derecha de la ec. (8.45) la información de $|\psi\rangle$ está totalmente en un factor no puede estar también en el otro.

los programas tienen que ser ortogonales, lo cual equivale a decir clásicos (una computación en la que sólo intervienen estados de la base computacional, sin superposiciones, equivale a una computación clásica con bits).

El problema persiste en elaboraciones más sofisticadas. Consideremos por ejemplo un procesador G que actuara según

$$G|\psi\rangle \otimes |\Xi\rangle \otimes |0\rangle = \sum_{j=1}^{n} V_j U|\psi\rangle \otimes |\Xi'_j\rangle \otimes |j\rangle \tag{8.49}$$

siendo $\{|j\rangle\}$ una base ortonormal de un espacio auxiliar. Se supone que los V_j son unos operadores fijos conocidos que actúan en $\mathscr{H}_d$. Entonces, haciendo una medida del registro auxiliar se conoce j y también $V_j U|\psi\rangle$, basta entonces aplicar V_j^{-1} para obtener $U|\psi\rangle$. Sin embargo esta mayor flexibilidad no permite sortear la limitación anterior. Si consideramos dos datos arbitrarios y dos programas, por unitaridad de G se sigue obteniendo

$$\langle\psi_1|\psi_2\rangle\langle\Xi_1|\Xi_2\rangle = \langle\psi_1|U_1^{-1}U_2|\psi_2\rangle \sum_j \langle\Xi'_{1j}|\Xi'_{2j}\rangle \tag{8.50}$$

con la misma conclusión de que operadores U distintos requieren programas ortogonales.

8.4. Procesadores cuánticos estocásticos

8.4.1. Implementación de un grupo uniparamétrico

El teorema que prohíbe procesadores cuánticos que implementen más operadores unitarios que la dimensión de $\mathscr{H}_p$ no se aplica cuando el protocolo es probabilístico, esto es, sólo se obtiene el resultado deseado con cierta probabilidad (pero se sabe cuándo esto ha ocurrido).

Veamos un ejemplo muy simple. El dato y el programa son un qubit cada uno y el operador depende de un parámetro

$$U_\gamma = e^{i\gamma\sigma_z} = \begin{pmatrix} e^{i\gamma} & 0 \\ 0 & e^{-i\gamma} \end{pmatrix} = \begin{pmatrix} \omega & 0 \\ 0 & \omega^{-1} \end{pmatrix} \qquad 0 \le \gamma < \pi \qquad \omega \equiv e^{i\gamma}. \tag{8.51}$$

Consideremos un circuito (Fig. 8.2) formado por una puerta CNOT en la que el qubit de control es el dato y el controlado el programa. El programa correspondiente al parámetro γ es

$$|\Xi_\gamma\rangle = \frac{1}{\sqrt{2}}(\omega|0\rangle + \omega^{-1}|1\rangle) = U_\gamma|+\rangle. \tag{8.52}$$

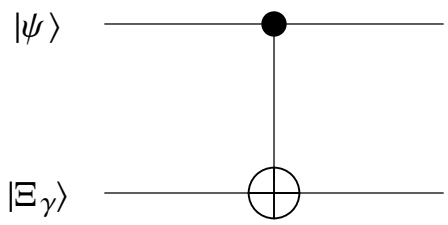


Figura 8.2: Circuito para el procesador probabilístico en (8.53).

Después de pasar por el circuito se tiene

$$\begin{aligned} |\psi\rangle \otimes |\Xi_\gamma\rangle &= (\alpha|0\rangle + \beta|1\rangle) \otimes \frac{1}{\sqrt{2}}(\omega|0\rangle + \omega^{-1}|1\rangle) \\ &\longrightarrow \frac{1}{\sqrt{2}}(\alpha\omega|00\rangle + \alpha\omega^{-1}|01\rangle + \beta\omega|11\rangle + \beta\omega^{-1}|10\rangle) \\ &= \frac{1}{\sqrt{2}}(U_\gamma|\psi\rangle \otimes |0\rangle + U_\gamma^{-1}|\psi\rangle \otimes |1\rangle) \end{aligned} \tag{8.53}$$

Si ahora se mide el estado del programa (el segundo qubit) se tiene una probabilidad $1/2$ de encontrar 0 y obtener el resultado deseado en el primer qubit: $U_\gamma|\psi\rangle$.

Se tiene un procesador que implementa una familia uniparamétrica de operadores, aunque sólo de forma probabilística, pero cuando ocurre sí se sabe que es el resultado correcto. El estado $|\Xi_\gamma\rangle$ requiere aplicar U_γ pero sólo sobre el estado $|+\rangle$, a cambio se obtiene (con cierta portabilidad) $U_\gamma|\psi\rangle$ para un estado cualquiera (que puede no conocerse y sin perder coherencia cuántica).

Si tenemos copias de $|\psi\rangle$ y no se obtiene el resultado la primera vez se puede simplemente repetir. Pero lo más usual es que $|\psi\rangle$ sea único (y de hecho no se sepa qué estado es). En ese caso, si se tiene $U_\gamma^{-1}|\psi\rangle$ aún se puede intentar obtener $U_\gamma|\psi\rangle$ a base de aplicar de nuevo el procesador pero con programa $|\Xi_{2\gamma}\rangle$. Esto produce

$$U_\gamma^{-1}|\psi\rangle \otimes |\Xi_{2\gamma}\rangle \longrightarrow \frac{1}{\sqrt{2}}(U_\gamma|\psi\rangle \otimes |0\rangle + U_{3\gamma}^{-1}|\psi\rangle \otimes |1\rangle)\,. \tag{8.54}$$

Si al medir el segundo qubit resulta 0 se tiene el resultado deseado. Con dos intentos la probabilidad sube a $\frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{3}{4}$. El proceso se puede iterar para mejorar la probabilidad.

Hay un teorema de teoría de circuitos (**principio de las medidas diferidas**) que afirma que las

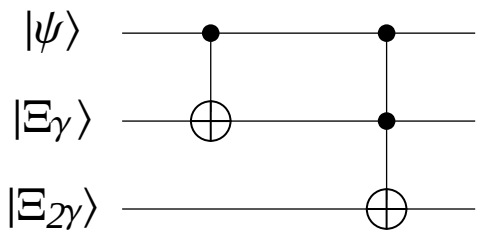


Figura 8.3: Circuito para el procesador probabilístico con dos intentos.

medidas se pueden cambiar por puertas condicionales cuánticas y posponerlas. [8.10] Se puede aplicar este principio para hacer los dos intentos con un solo circuito. Basta volver a procesar el segundo qubit condicionalmente, como se muestra en la Fig. 8.3. Haciendo esto se obtiene

$$\frac{1}{2}U_\gamma|\psi\rangle_1 \otimes (\omega^2|00\rangle + \omega^{-2}|01\rangle + |10\rangle)_{23} + \frac{1}{2}U_{3\gamma}^{-1}|\psi\rangle_1 \otimes |11\rangle_{23} \tag{8.55}$$

Ahora se miden los registros 2 y 3 en la base computacional. Con probabilidad $3/4$ se obtiene uno de los resultados 00, 01 o 10 y eso nos proporciona el resultado deseado. Con probabilidad $1/4$ se obtiene $U_{3\gamma}^{-1}|\psi\rangle$, igual que midiendo después de cada intento.

[8.10] Nielsen y Chuang, pág. 186.

9. Corrección de errores

9.1. Corrección de errores clásicos y redundancia

En computación o comunicación clásicas se producen errores en forma de mutación de bits. La forma usual de protegerse frente a estos errores es mediante redundancia. Por ejemplo se puede cambiar cada 0 o 1 de un mensaje (o registro en una computación) por $0_L = 000$ y $1_L = 111$, respectivamente. En caso de un error que produzca la mutación de un bit $0 \leftrightarrow 1$, digamos $000 \to 010$, el error se detecta y se corrige usando el criterio de la mayoría. En el ejemplo hay mayoría de ceros y a la vista de 010 se concluye que lo correcto es 000, suponiendo que no haya más de un error.

Si suponemos que los errores en los bits son independientes y que cada bit tiene una probabilidad p de error, es fácil ver que, desde el punto de vista de corrección de errores, es beneficioso usar redundancia cuando $p < 1/2$. En efecto, supongamos que cada bit lógico x_L equivale a n bits físicos: $x_L = x \cdots x$ (longitud n, y n impar). La probabilidad de exactamente m errores será $\binom{n}{m} p^m (1-p)^{n-m}$. Entonces el criterio de la mayoría restaurará el resultado correcto si y sólo si $m < n/2$, es decir, la nueva probabilidad de error usando redundancia será $p_n = \sum_{m>n/2} \binom{n}{m} p^m (1-p)^{n-m}$. Se puede probar que p_n es una función estrictamente decreciente de n cuando $0 < p < 1/2$ y por tanto $p_n < p_1 = p$ si $n > 1$. [9.1]

Cambiar cada bit por su versión codificada no es la única posibilidad y puede ser mejor codificar cadenas de k bits. Esencialmente se tiene el espacio $\mathbb{Z}_2^k$ de palabras (mensajes, datos a manipular, etc) de longitud k que se quieren codificar para protegerlas frente a errores, y un espacio $\mathbb{Z}_2^n$ que contiene las palabras codificadas, de longitud $n \geq k$. Se suele decir que es una codificación de tipo $[n,k]$. El ejemplo anterior es del tipo $[3,1]$.

La codificación es una aplicación inyectiva $f_c : \mathbb{Z}_2^k \to \mathbb{Z}_2^n$, que hace corresponder a cada palabra de $\mathbb{Z}_2^k$ una versión codificada en $\mathbb{Z}_2^n$, y que llena un conjunto $V_c = f_c(\mathbb{Z}_2^k) \subset \mathbb{Z}_2^n$. Si no hubiera mutaciones las palabras en V_c se mantendrían ahí todo el tiempo, pero los errores van a cambiar bits de modo que cada $x \in V_c$ puede cambiar hasta t bits después de t errores. Se dice que dos palabras están a distancia d cuando difieren en d bits. Si se estima un número t de errores como máximo, se querrá que las palabras en V_c estén a distancia $d \geq 2t+1$ unas de otras para que distintas palabras mutadas no se

[9.1] Al poner más bits la probabilidad de que alguno falle (es decir de $m \geq 1$) aumenta, pero lo que importa es que la probabilidad de que la mayoría falle disminuye.

mezclen y se puedan corregir unívocamente. Un código $[n,k]$ con distancia mínima d entre puntos de V_c se suele denotar $[n,k,d]$.

A menudo la función de codificación f_c es lineal (sobre el cuerpo $\mathbb{Z}_2$) ya que es así es más fácil diseñar sus propiedades óptimamente (se quiere n lo más pequeño posible). Los códigos clásicos se extienden trivialmente a cuánticos en los relativo a bits, pero no a errores en fases de los qubits. Notablemente el método CSS (Calderbank, Steane, Shor) permite extender de manera sistemática los códigos clásicos lineales a cuánticos, esencialmente usando dos códigos clásicos, uno para corregir los bits y otro para corregir las fases.

9.2. Codificación de Shor

Cuando se va del caso clásico al cuántico hay que tener en cuenta:

i) Como se ha dicho, cuánticamente puede haber errores no sólo en el bit $|0\rangle \leftrightarrow |1\rangle$ sino también en fases $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \leftrightarrow \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = |-\rangle$.

ii) No se puede simplemente medir el estado para detectar los errores ya que eso produciría un cambio en el estado.

iii) No se puede copiar el estado para corregirlo ya que lo prohíbe el teorema de no clonación.

Supongamos que de momento nos preocupamos sólo de errores en bits y no en fases y usamos el esquema [9.2]

$$|0\rangle \to |0\rangle_L = |000\rangle, \qquad |1\rangle \to |1\rangle_L = |111\rangle\,. \tag{9.1}$$

Esta codificación se puede implementar sin dificultad usando por ejemplo el circuito de la Fig. 9.1.

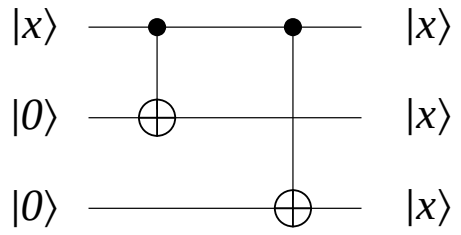


Figura 9.1: Circuito para codificación $|x\rangle \to |x\rangle_L = |xxx\rangle$.

[9.2] El espacio subtendido por $|0\rangle_L$, $|1\rangle_L$ es un qubit en todos los sentidos, aunque sean estados "compuestos". En realidad no se ha especificado tampoco la naturaleza elemental o compuesta de $|0\rangle$ y $|1\rangle$, y no es relevante.

Circuitos igualmente simples se usan para otras codificaciones. Lo importante en todo caso es que se haga preservando el paralelismo cuántico, ya que en aplicaciones el input será una superposición de estados de la base computacional.

Para esta codificación, los operadores $Z_L := Z_1Z_2Z_3$ y $X_L := X_1X_2X_3$ sirven como Z y X lógicos, es decir, $Z_L|1\rangle_L = -|1\rangle_L$, $X_L|0\rangle_L = +|1\rangle_L$, etc. Aquí Z_j es $Z = \sigma_z$ sobre el qubit j, y X_j es $X = \sigma_x$ sobre el qubit j.

Puesto que no permitimos mutaciones en fases, los operadores que pueden actuar son de tipo X, así un error X_2 produce

$$X_2(\alpha|000\rangle + \beta|111\rangle) = \alpha|010\rangle + \beta|101\rangle. \tag{9.2}$$

Es importante notar que las mutaciones también son acciones físicas y por tanto se implementan a través de operadores unitarios. También nótese que una mutación que afecte sólo a un qubit no puede ser por ejemplo del tipo

$$\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \longrightarrow \frac{1}{\sqrt{2}}(|010\rangle + |111\rangle)\,. \tag{9.3}$$

El qubit (en este caso el 2) es un solo sistema físico y no puede ser mutado en una componente y no en la otra. [9.3]

Si suponemos que a lo sumo hay *un* error en un qubit ($t = 1$ en la discusión de más arriba) ese error (que suponemos de tipo X) hará que dos bits contiguos difieran y eso se puede detectar aplicando operadores del tipo Z_iZ_{i+1}:

$$\begin{aligned} Z_1Z_2(\alpha|010\rangle + \beta|101\rangle) &= -(\alpha|010\rangle + \beta|101\rangle) \\ Z_2Z_3(\alpha|010\rangle + \beta|101\rangle) &= -(\alpha|010\rangle + \beta|101\rangle) \end{aligned} \tag{9.4}$$

Los operadores Z_1Z_2 y Z_2Z_3 se denominan **estabilizadores** ya que dejan invariantes los estados de qubits lógicos $|000\rangle$ y $|111\rangle$. Los autovalores de los estabilizadores o **síndrome** caracterizan el error. Así en el ejemplo $(Z_1Z_2, Z_2Z_3) = (-,-)$ caracteriza a la mutación X_2. En última instancia los dos signos menos corresponden a que X_2 *anticonmuta* con Z_1Z_2 y con Z_2Z_3. La Tabla 2 muestra los posibles síndromes y los errores asociados.

Una vez identificado el error se puede corregir aplicando el operador X_i correspondiente (por ejemplo X_2 para $(Z_1Z_2, Z_2Z_3) = (-,-)$). Lo importante es que la detección e identificación se pueden hacer de manera automática, por el *principio de medidas diferidas* no es necesario medir los observables Z_iZ_{i+j}, identificar el error y corregirlo.

[9.3] Esta transformación conserva la norma y por tanto es realizable unitariamente, pero no con operadores que actúen sólo sobre el qubit 2.

			$Z_1Z_2I_3$	$I_1Z_2Z_3$
000	111	I	$+$	$+$
100	011	X_1	$-$	$+$
010	101	X_2	$-$	$-$
001	110	X_3	$+$	$-$

Tabla 2: Síndromes para la codificación $|x\rangle_L = |xxx\rangle$ (se supone $t = 1$ y errores de tipo X_j). El signo es $+$ cuando mutación y estabilizador conmutan y $-$ cuando anticonmutan.

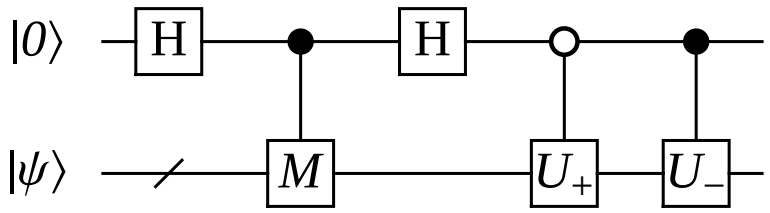


Figura 9.2: Circuito para aplicar $U_\pm$ sobre $|\psi\rangle$ según $M = \pm 1$.

Veamos como funciona esto: Tenemos un estado $|\psi\rangle$ en un estado $\mathscr{H}$. En ese espacio, consideremos un operador unitario M con autovalores ± 1 (por tanto también es un observable hermítico). Queremos que cuando $M = +1$ actúe un operador unitario U_+ y cuando $M = -1$ actúe otro operador U_-. Eso se puede hacer con el circuito de la Fig. 9.2, que produce

$$|0\rangle \otimes |\psi\rangle \longrightarrow |0\rangle \otimes U_+ P_+ |\psi\rangle + |1\rangle \otimes U_- P_- |\psi\rangle\,, \qquad P_\pm = \frac{1}{2}(I \pm M)\,. \tag{9.5}$$

Esta construcción (extendida con más operadores M y $U_\pm$) se aplica automáticamente a la corrección de errores, siendo M los operadores de tipo Z_iZ_{i+1} (o I) que detectan los errores y $U_\pm$ los operadores de tipo X_j (o I) que los corrigen sobre la marcha. [9.4] El mismo método sirve para errores más generales.

Si en lugar de en el *bit*, el error es en el *signo*, la mutación es $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \leftrightarrow \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, es decir, $|+\rangle \leftrightarrow |-\rangle$. Esto indica que se puede proceder esencialmente como antes pero intercambiando las bases $\{|0\rangle, |1\rangle\} \leftrightarrow \{|+\rangle, |-\rangle\}$, y al mismo tiempo $Z \leftrightarrow X$. (Equivale a aplicar una transformación unitaria, a saber, H). Así la base computacional lógica es ahora

$$|0\rangle \to |0\rangle_L = |+++\rangle\,, \qquad |1\rangle \to |1\rangle_L = |---\rangle\,, \tag{9.6}$$

[9.4] Por ejemplo, si el estado debería ser $|0\rangle$ pero puede entrar $|1\rangle$, $M = Z$ detecta el error y $U_- = X$ ($U_+ = I$) lo corrige. Concretamente $|0\rangle \otimes |\psi\rangle \to |\psi\rangle \otimes |0\rangle$.

junto con $Z_L = X_1X_2X_3$ y $X_L = Z_1Z_2Z_3$. [9.5]

Ahora los errores son de tipo Z_j y los estabilizadores son X_1X_2 y X_2X_3. Así, una mutación tal como $|+-+\rangle$ (producida por Z_2) es detectada por el síndrome $(X_1X_2, X_2X_3) = (-,-)$ y se corrige volviendo a aplicar Z_2.

La **codificación de Shor**, con nueve qubits, combina los dos esquemas (bit y signo) mediante

$$\begin{aligned} &|0\rangle \longrightarrow |0\rangle_L = |\bar{+}\bar{+}\bar{+}\rangle, \qquad |1\rangle \longrightarrow |1\rangle_L = |\bar{-}\bar{-}\bar{-}\rangle, \\ &|\bar{+}\rangle = \frac{1}{\sqrt{2}}(|\bar{0}\rangle + |\bar{1}\rangle), \qquad |\bar{-}\rangle = \frac{1}{\sqrt{2}}(|\bar{0}\rangle - |\bar{1}\rangle), \qquad |\bar{0}\rangle \equiv |000\rangle, \qquad |\bar{1}\rangle \equiv |111\rangle\,. \end{aligned} \tag{9.7}$$

Los 9 qubits se agrupan en tres bloques de tres, y se pueden etiquetar con $11, 12, \ldots, 33$. La redundancia en signo se aplica a nivel de bloques y la de bits a nivel de qubits dentro de cada bloque.

Los operadores $\bar{X}_i := X_{i1}X_{i2}X_{i3}$ y $\bar{Z}_i := Z_{i1}Z_{i2}Z_{i3}$, satisfacen

$$\bar{X}|\bar{+}\rangle = |\bar{+}\rangle, \qquad \bar{X}|\bar{-}\rangle = -|\bar{-}\rangle, \qquad \bar{Z}|\bar{+}\rangle = |\bar{-}\rangle, \qquad \bar{Z}|\bar{-}\rangle = |\bar{+}\rangle. \tag{9.8}$$

Entonces los operadores Z y X lógicos se pueden representar con $Z_L = \bar{X}_1\bar{X}_2\bar{X}_3$ y $X_L = \bar{Z}_1\bar{Z}_2\bar{Z}_3$. (Es como la codificación (9.6) pero con barras encima.)

De nuevo suponemos que a lo sumo ocurre un error (de tipos X_{ij} o Z_{ij}) en uno de los 9 qubits.

Un cambio en un **signo** en uno de los 9 qubits producirá una mutación $|\bar{+}\rangle \leftrightarrow |\bar{-}\rangle$ en uno de los bloques. Ese error se produce con Z_{ij} y es exactamente el mismo error que produciría $\bar{Z}_i$, por tanto se detecta con $\bar{X}_1\bar{X}_2$ y $\bar{X}_2\bar{X}_3$.

Análogamente, una mutación en el **bit** en uno de los 9 qubits, se produce con X_{ij}, y lo detectan los operadores $Z_{i1}Z_{i2}$ y $Z_{i2}Z_{i3}$.

Es decir, los signos se corrigen a escala de bloques y los bits a escala de qubits dentro de un bloque. Aunque no sea obvio, la codificación de Shor no sólo corrige errores de bit o signo sino toda clase de errores que afecten a un solo qubit ($t = 1$).

[9.5] Recordemos que

$$Z\begin{pmatrix}|0\rangle\\|1\rangle\end{pmatrix} = \begin{pmatrix}+|0\rangle\\-|1\rangle\end{pmatrix}, \qquad Z\begin{pmatrix}|+\rangle\\|-\rangle\end{pmatrix} = \begin{pmatrix}|-\rangle\\|+\rangle\end{pmatrix}, \qquad X\begin{pmatrix}|0\rangle\\|1\rangle\end{pmatrix} = \begin{pmatrix}|1\rangle\\|0\rangle\end{pmatrix}, \qquad X\begin{pmatrix}|+\rangle\\|-\rangle\end{pmatrix} = \begin{pmatrix}+|+\rangle\\-|-\rangle\end{pmatrix}.$$

9.3. Generalidades en el caso cuántico

La situación general es que se tiene un espacio de Hilbert $\mathscr{H}_L$ de k qubits (denominados qubits lógicos) con los estados que se quieren codificar. Se codifican en un espacio mayor $\mathscr{H}_W$ de n qubits (denominados qubits físicos) mediante un operador isométrico de modo que $\mathscr{H}_L \longrightarrow \mathscr{H}_c \subset \mathscr{H}_W$, siendo $\mathscr{H}_L \longrightarrow \mathscr{H}_c$ una biyección. El subespacio $\mathscr{H}_c$ contiene los estados codificados (sin mutaciones, tales como $|000\rangle$) y $\mathscr{H}_W$ contiene todos los estados físicos (con sus mutaciones, tales como $|010\rangle$). En ausencia de errores los estados permanecerían en $\mathscr{H}_c$ durante su procesamiento, hasta ser decodificados al final. Los errores posibles se introducen a través de operadores que forman un espacio lineal $\mathscr{E}_L$. Este conjunto de operadores va a depender de qué tipo de errores se incluyan (digamos, a cuántos qubits pueden llegar a afectar las mutaciones). Por efecto de los errores se pasará de $\mathscr{H}_c$ a un subespacio mayor $\mathscr{H}_M = \mathscr{E}_L\mathscr{H}_c \subset \mathscr{H}_W$. [9.6]

Cuando se habla de una **codificación**, uno se refiere no al diccionario $\mathscr{H}_L \leftrightarrow \mathscr{H}_c$ concreto que es lo de menos, sino a la elección de $\mathscr{H}_c$ como subespacio de $\mathscr{H}_W$. Aunque todos los espacios de igual dimensión son isomorfos, distintas elecciones de $\mathscr{H}_c \subset \mathscr{H}_W$ son inequivalentes por la estructura de producto tensorial existente en $\mathscr{H}_W$ y por la naturaleza de los errores permitidos $\mathscr{E}_L$. De lo que se trata es de elegir $\mathscr{H}_c$ de modo que el espacio $\mathscr{H}_M$ de estados con errores aún permita su recuperación. Es decir, si $|\psi\rangle$ es cualquier estado de $\mathscr{H}_c$ y M cualquier operador de $\mathscr{E}_L$, el estado $|\psi'\rangle = M|\psi\rangle$ debe determinar $|\psi\rangle$ unívocamente. (Si además M también queda unívocamente identificado se dice que la codificación es **no degenerada**.) Como en el caso clásico se necesita que los estados en $\mathscr{H}_c$ estén suficientemente separados dentro de $\mathscr{H}_W$ de modo que al pasar de $\mathscr{H}_c$ a $\mathscr{H}_M$ los estados sigan siendo distinguibles. Al mismo tiempo se quiere minimizar la cantidad de recursos dedicados a corrección de errores, en particular minimizar n (el número de qubits físicos). Eso lleva a intentar optimizar el empaquetamiento de $\mathscr{H}_c$ dentro de $\mathscr{H}_W$ (hasta donde sea práctico hacerlo).

Para un qubit, el conjunto $\{I,X,Y,Z\}$ (siendo $Y = \sigma_y = iXZ$) forma una base lineal de operadores. Cada uno de estos operadores tiene su versión $\{I_j,X_j,Y_j,Z_j\}$ actuando sobre el qubit j de $\mathscr{H}_W$. Sus productos tensoriales (n factores) forman una base del espacio de operadores en $\mathscr{H}_W$, con 4^n elementos. Los operadores de tipo X,Y,Z son los que introducen errores. Se dice que un operador de $\mathscr{H}_W$ tiene peso t cuando al expresarlo en esa base hay componentes no nulas con exactamente t factores de tipo X, Y o Z (y por tanto $n-t$ de tipo I), pero ninguna componente con más de t tales factores. [9.7] Para cada t, los operadores de peso menor o igual a t forman un espacio vectorial $\hat{\mathscr{H}}_t$. Decimos que se han producido t errores cuando sobre $\mathscr{H}_c$ ha actuado un operador de peso t. Típicamente $\mathscr{E}_L$ es o

[9.6] También se incluye el operador identidad en $\mathscr{E}_L$ de modo que $\mathscr{H}_c \subset \mathscr{H}_M$.

[9.7] Dicho de otro modo, el operador $\mathscr{O}$ tiene peso t cuando el cambio $(I,X,Y,Z) \to (I,\lambda X,\lambda Y,\lambda Z)$ transforma $\mathscr{O}$ en un polinomio de grado t en el parámetro λ.

está contenido en $\hat{\mathscr{H}_t}$ para algún nivel de errores t.

Hay que notar que $\mathscr{H}_W$ es solamente el espacio de codificación $\mathscr{H}_c$ extendido con mutaciones arbitrarias. El espacio completo será $\mathscr{H}_W \otimes \mathscr{H}_E$ que tiene en cuenta los estados del entorno. La interacción con el ambiente es la que va a producir los errores. Es cierto que los errores actúan a través de operadores unitarios pero son operadores unitarios definidos en el espacio completo $\mathscr{H}_W \otimes \mathscr{H}_E$. En general no son producto de unitarios en cada espacio por separado. Entonces desde el punto de vista de $\mathscr{H}_W$ la evolución se verá como un canal cuántico, T_E,

$$\rho \to T_E(\rho) = \sum_\mu M_\mu \rho M_\mu^\dagger, \qquad \sum_\mu M_\mu^\dagger M_\mu = I \qquad \text{en } \mathscr{H}_W\,. \tag{9.9}$$

Los operadores de Kraus M_μ están en el espacio $\mathscr{E}_L$ de operadores que implementan las mutaciones.

Corregir los errores no es más que invertir el efecto del canal cuántico T_E. Como ya se vio, un canal cuántico no es invertible a menos que sea unitario, pero sí puede serlo cuando ρ se restringe a un subespacio $\mathscr{H}_c$ como es nuestro caso. Por tanto se requiere que exista otro canal R tal que $R \circ T_E = \hat{I}$ sobre $\mathscr{H}_c$. Cuando los M_μ actúan sobre $\mathscr{H}_c$ producen el espacio $\mathscr{H}_M$ que incluye los estados mutados. A efectos prácticos T_E va de $\mathscr{H}_c$ a $\mathscr{H}_M$ y R de $\mathscr{H}_M$ a $\mathscr{H}_c$.

Teorema La condición necesaria y suficiente para que un canal cuántico T_E se pueda invertir cuando se restringe a un subespacio $\mathscr{H}_c \subset \mathscr{H}_W$, es que los operadores de Kraus satisfagan

$$P_c M_\mu^\dagger M_\nu P_c = m_{\mu\nu} P_c\,, \tag{9.10}$$

para ciertos coeficientes $m_{\mu\nu}$. P_c denota el proyector ortogonal sobre $\mathscr{H}_c$.

Equivale a decir que $M_\mu^\dagger M_\nu$ es un múltiplo de la identidad cuando se restringe a $\mathscr{H}_c$.

Demostración: De la expresión se deduce que $m_{\mu\nu}$ es una matriz positiva, entonces se puede hacer una rotación unitaria entre los operadores de Kraus para diagonalizarla (y los seguimos denotando M_μ)

$$P_c M_\mu^\dagger M_\nu P_c = m_\mu \delta_{\mu\nu} P_c, \qquad m_\mu > 0 \qquad \sum_\mu m_\mu = 1\,. \tag{9.11}$$

Los operadores M_μ con $m_\mu = 0$ pueden excluirse ya que se anulan sobre $\mathscr{H}_c$. Sea $\mathscr{H}_M$ el espacio subtendido por los estados $M_\mu \mathscr{H}_c$ (los estados mutados). Entonces se puede definir el canal R en $\mathscr{H}_M$ con operadores de Kraus

$$R_\nu = \frac{1}{\sqrt{m_\nu}} P_c M_\nu^\dagger\,. \tag{9.12}$$

Comprobamos la normalización:

$$N \equiv \sum_\nu R_\nu^\dagger R_\nu = \sum_\nu \frac{1}{m_\nu} M_\nu P_c M_\nu^\dagger \tag{9.13}$$

Entonces

$$N M_\mu P_c = \sum_\nu \frac{1}{m_\nu} M_\nu P_c M_\nu^\dagger M_\mu P_c = M_\mu P_c \tag{9.14}$$

Esto implica que $N = P_M$, el proyector ortogonal sobre el espacio $\mathscr{H}_M$. Entonces el superoperador R está bien normalizado en $\mathscr{H}_M$ que es todo lo que hace falta. Si se quiere que R actúe en todo $\mathscr{H}_W$ se puede completar añadiendo otro operador de Kraus $R_\nu = P_M^\perp$ que no tiene ningún efecto.

Podemos ahora comprobar que R invierte el efecto de T_E cuando éste último actúa sobre un ρ en $\mathscr{H}_c$:

$$\begin{aligned} R \circ T_E(\rho) &= \sum_{\mu,\nu} R_\nu M_\mu \rho M_\mu^\dagger R_\nu^\dagger = \sum_{\mu,\nu} R_\nu M_\mu P_c \rho P_c M_\mu^\dagger R_\nu^\dagger \\ &= \sum_{\mu,\nu} \frac{1}{\sqrt{m_\nu}} P_c M_\nu^\dagger M_\mu P_c \rho P_c M_\mu^\dagger M_\nu P_c \frac{1}{\sqrt{m_\nu}} \\ &= \sum_{\mu,\nu} m_\mu \delta_{\mu\nu} P_c \rho P_c = \sum_\mu m_\mu \rho = \rho\,, \qquad \rho \in \mathscr{H}_c\,. \end{aligned} \tag{9.15}$$

Esto prueba que (9.10) es suficiente para que T_E se pueda invertir. Veamos que la condición también es necesaria.

Suponemos que $R \circ T_E = \hat{I}$ sobre $\mathscr{H}_c$. Esto implica que

$$\forall |\psi\rangle \in \mathscr{H}_c \qquad |\psi\rangle\langle\psi| = R \circ T_E(|\psi\rangle\langle\psi|) = \sum_\alpha \sum_\nu R_\alpha M_\mu |\psi\rangle\langle\psi| M_\mu^\dagger R_\alpha^\dagger\,. \tag{9.16}$$

Como ya se vio (pág. 18) un estado puro es extremal: no se puede escribir como una mezcla de otros estados. Eso implica

$$R_\alpha M_\mu |\psi\rangle = \lambda_{\alpha\mu} |\psi\rangle. \tag{9.17}$$

Es decir todos los vectores de $\mathscr{H}_c$ son propios de $R_\alpha M_\mu$. Esto implica que todos los valores propios (al variar $|\psi\rangle$) son iguales y $\lambda_{\alpha\mu}$ no depende de $|\psi\rangle$. Es decir

$$\forall \alpha, \mu \qquad R_\alpha M_\mu P_c = \lambda_{\alpha\mu} P_c\,. \tag{9.18}$$

Entonces

$$P_c M_\mu^\dagger R_\alpha^\dagger R_\beta M_\nu P_c = \lambda_{\alpha\mu}^* \lambda_{\beta\nu} P_c \tag{9.19}$$

Tomando $\alpha = \beta$ y sumando se deduce

$$P_c M_\mu^\dagger M_\nu P_c = m_{\mu\nu} P_c \,. \tag{9.20}$$

□

Como se ha dicho, suponemos $M_\mu \in \mathscr{E}_L$. Entonces

Corolario Una condición *suficiente* para que T_E sea invertible cuando actúa en $\mathscr{H}_c$ es

$$\forall M_a, M_b \in \mathscr{E}_L \qquad P_c M_a^\dagger M_b P_c = C_{ab} P_c \,, \tag{9.21}$$

para ciertos coeficientes C_{ab} (que se deduce que forman una matriz positiva). Obviamente basta garantizar la condición para los estados de una base de $\mathscr{H}_c$

$$\langle u_j | M_a^\dagger M_b | u_k \rangle = C_{ab} \delta_{jk} \,, \tag{9.22}$$

siendo $\{|u_j\rangle\}$ una base ortonormal de $\mathscr{H}_c$. También basta que se cumpla para los operadores de una base de $\mathscr{E}_L$.

Usando sendas bases de $\mathscr{E}_L$ y $\mathscr{H}_c$ se puede comprobar que la codificación de Shor corrige cualquier error que afecte a un qubit, es decir para $\mathscr{E}_L = \hat{\mathscr{H}}_{t=1}$.

9.4. Método de estabilizadores

Una forma práctica de plantear la corrección de errores es mediante el formalismo de **estabilizadores**. Veamos la idea con un ejemplo.

La **cota de Singleton** establece que una codificación de k qubits usando n qubits y que corrija errores en hasta t qubits debe satisfacer la desigualdad $n \geq 4t + k$. [9.8] Para $k = t = 1$ el mínimo es $n = 5$ y de hecho hay una codificación cuántica (y esencialmente única) de tipo $[5, 1]$. En $\mathscr{H}_W = (\mathbb{C}^2)^{\otimes 5}$ (dimensión 32) definimos 4 operadores, los denominados **estabilizadores**:

$$\begin{aligned} A_1 &= X_1 Z_2 Z_3 X_4 I_5, \\ A_2 &= I_1 X_2 Z_3 Z_4 X_5, \\ A_3 &= X_1 I_2 X_3 Z_4 Z_5, \\ A_4 &= Z_1 X_2 I_3 X_4 Z_5, \end{aligned} \tag{9.23}$$

[9.8] Nielsen y Chuang, pág. 568

Teniendo en cuenta que operadores de distinto qubit conmutan y para un mismo qubit anticonmutan,

$$XY = -YX, \qquad XZ = -ZX, \qquad YZ = -ZY, \tag{9.24}$$

se deduce que los cuatro estabilizadores A_α conmutan entre sí, y también con $Z_L \equiv Z_1Z_2Z_3Z_4Z_5$. Los cinco operadores forman un CCOC. Todos tienen autovalores ± 1 y cada autovalor corresponde a un espacio de dimensión $2^4 = 16$.

El espacio de codificación $\mathscr{H}_c$ va a ser el subespacio bidimensional caracterizado por $A_1 = A_2 = A_3 = A_4 = +1$ (de ahí el nombre de estabilizadores). Los dos estados en $\mathscr{H}_c$ se distinguen mediante $Z_L = \pm 1$,

$$Z_L|0\rangle_L = +|0\rangle_L, \qquad Z_L|1\rangle_L = -|1\rangle_L. \tag{9.25}$$

Es decir, en el formalismo de estabilizadores el espacio de codificación $\mathscr{H}_c$ y la codificación concreta $|x\rangle_L$ se fijan a través de la elección de operadores.

Los estados con errores tendrán $A_\alpha = -1$ para al menos un α. Si por ejemplo a un estado $|\psi\rangle \in \mathscr{H}_c$ se le aplica X_1 (que lo saca fuera de $\mathscr{H}_c$) se tendrá $|\psi'\rangle = X_1|\psi\rangle$, y entonces

$$A_\alpha|\psi'\rangle = A_\alpha X_1|\psi\rangle = \begin{cases} +X_1A_\alpha|\psi\rangle = +X_1|\psi\rangle = +|\psi'\rangle & (\alpha = 1,2,3) \\ -X_1A_\alpha|\psi\rangle = -|\psi'\rangle & (\alpha = 4) \end{cases} \tag{9.26}$$

dado que X_1 conmuta con $A_{1,2,3}$ y anticonmuta con A_4. El patrón de autovalores o **síndrome** $(+++-)$ identifica unívocamente al error X_1 lo cual permite corregirlo.

	$A_1A_2A_3A_4$		$A_1A_2A_3A_4$		$A_1A_2A_3A_4$
X_1	$+++-$	Z_1	$-+-+$	Y_1	$-+--$
X_2	$-+++$	Z_2	$+-+-$	Y_2	$--+-$
X_3	$--++$	Z_3	$++-+$	Y_3	$---+$
X_4	$+--+$	Z_4	$-++-$	Y_4	$----$
X_5	$++--$	Z_5	$+-++$	Y_5	$+---$

Tabla 3: Tabla de síndromes para la codificación $[5,1]$ y errores de un qubit ($t = 1$).

Los síndromes se muestran en la Tabla 3. [9.9] Los $3 \times 5 = 15$ síndromes correspondientes a $t = 1$ son todos distintos, eso implica que a nivel $t = 1$ el error se puede identificar unívocamente. Se dice entonces que es un código **no degenerado**. Añadiendo el caso $(++++)$ salen los 2^4 espacios

[9.9] La columna Y es el producto de las columnas X y Z dado que Y es proporcional a XZ entonces, si para un error M, $XM = \sigma MX$ y $ZM = \sigma' MZ$, se tendrá $YM = \sigma\sigma' MY$ $(\sigma, \sigma' = \pm 1)$.

bidimensionales $(A_1A_2A_3A_4) = (\pm\pm\pm\pm)$. Se deduce entonces que la codificación no corrige ningún error de tipo $t \geq 2$ porque ya no quedan síndromes sin usar.

Como se ha dicho, en el espacio $\mathscr{H}_c$ el operador lógico Z es $Z_L = Z^{\otimes 5}$, que por construcción actúa correctamente sobre $|0\rangle_L$ y $|1\rangle_L$. El operador X lógico se puede elegir como $X_L = X^{\otimes 5}$. En efecto, esta elección satisface los requerimientos básicos: i) deja $\mathscr{H}_c$ invariante ya que conmuta con los estabilizadores, ii) $X_L^2 = I$, y iii) anticonmuta con Z_L. Esto garantiza que $X_L|0\rangle_L = \omega|1\rangle_L$ y $X_L|1\rangle_L = \omega^*|0\rangle_L$ con $|\omega| = 1$. Las condiciones (9.25) no fijan las fases de los estados $|0\rangle_L$ y $|1\rangle_L$. La elección de X_L fija la fase de $|1\rangle_L$ relativa a $|0\rangle_L$ de modo que $\omega = 1$, dicho de otro modo $|1\rangle_L := X_L|0\rangle_L$. (La fase global queda sin fijar y no es relevante.) El operador Y_L se define a partir de X_L y Z_L mediante $Y_L = iX_LZ_L$, y por construcción satisface las relaciones de conmutación, matriz en la base de computación lógica, etc, apropiadas.

De hecho, una de las virtudes de la codificación basada en estabilizadores es que se basa en operadores y no estados lo cual evita entrar en detalles de elección de convenios de fases, hasta donde sea posible.

La codificación $|x\rangle_L = |xxx\rangle$ (que sólo protege frente a cambios en bits, X_1, X_2, X_3) tiene $Z_1Z_2I_3$ y $I_1Z_2Z_3$ como estabilizadores y los dos estados $(++)$ se distinguen mediante $Z_L = Z_1Z_2Z_3$.

	1	2	3	4	5	6	7
A_1	I	I	I	X	X	X	X
A_2	I	X	X	I	I	X	X
A_3	X	I	X	I	X	I	X
A_4	I	I	I	Z	Z	Z	Z
A_5	I	Z	Z	I	I	Z	Z
A_6	Z	I	Z	I	Z	I	Z

Tabla 4: Estabilizadores para la codificación de Steane $[7,1]$.

La codificación con $n = 5$ es la más compacta para $k = 1$ pero a menudo se utiliza la **codificación de Steane**, con $n = 7$ porque es más regular y es de la familia de codificaciones CSS. Los 6 estabilizadores de la codificación de Steane se muestran en la Tabla 4. El espacio $\mathscr{H}_c$ es $(+++++++)$ y es de dimensión 2. En ese espacio los operadores lógicos actúan como $X_L = X^{\otimes 7}$ y $Z_L = Z^{\otimes 7}$. Los 3×7 síndromes son todos distintos, es un código no degenerado ya que todos los errores (operadores) a nivel $t = 1$ se pueden identificar unívocamente.

Igualmente la codificación de Shor, de tipo $[9,1]$, se puede expresar mediante 8 estabilizadores y corrige cualquier error a nivel $t = 1$. Este código es degenerado ya que hay errores distintos (tales

como Z_{11} y Z_{12}) que producen la misma mutación. Aun así la corrección funciona ya que identifica la mutación (mutaciones distintas tienen síndromes distintos) para corregirla.

10. Bibliografía

Referencias

[1] R. P. Feynman, *Simulating physics with computers,* Int. J. Theor. Phys. **21** (1982), 467-488 doi:10.1007/BF02650179

[2] A. Peres, *Separability criterion for density matrices,* Phys. Rev. Lett. **77** (1996), 1413-1415 doi:10.1103/PhysRevLett.77.1413 [arXiv:quant-ph/9604005 [quant-ph]].

[3] L-M. Duan, G. Giedke, J. I. Cirac y P. Zoller, *Inseparability criterion for continuous variable systems,* Phys. Rev. Lett. **84** (2000), 2722-2725 doi:10.1103/PhysRevLett.84.2722

[4] C. H. Bennett, H. J. Bernstein, S. Popescu, y B. Schumacher, *Concentrating partial entanglement by local operations,* Phys. Rev. A **53** (1996), 2046 doi:10.1103/PhysRevA.53.2046

[5] W. K. Wootters, *Entanglement of formation of an arbitrary state of two qubits,* Phys. Rev. Lett. **80** (1998), 2245 doi:10.1103/PhysRevLett.80.2245

[6] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa y B. Schumacher, *Noncommuting mixed states cannot be broadcast,* Phys. Rev. Lett. **76** (1996), 2818-2821 doi:10.1103/PhysRevLett.76.2818

[7] P. O. Boykin, V. Roychowdhury, *Optimal encryption of quantum bits,* Phys. Rev. A **67** (2003), 042317 https://journals.aps.org/pra/pdf/10.1103/PhysRevA.67.042317