Una breve introducción al procesado cuántico de la información

Versión 1.1 (Curso 2003/2004)

David J. Santos
Departamento de Teoría de la Señal y Comunicaciones
Universidad de Vigo

Prólogo

Pese a tratarse de una disciplina reciente, puede decirse que el Procesado Cuántico de la Información (PCI) se halla en estos momentos en fase de consolidación formal. En numerosas universidades en todo el mundo se imparten hoy en día cursos de doctorado sobre esta disciplina.

El texto que el lector tiene en sus manos constituye la base del curso de Doctorado "Introducción al Procesado Cuántico de la Información", del Programa de Doctorado no presencial "Tecnologías de la Información", que se imparte actualmente en el Departamento de Teoría de la Señal y Comunicaciones de la Universidad de Vigo. El objetivo principal del curso es introducir gradualmente al alumno, típicamente Ingeniero de Telecomunicación, en los conceptos y técnicas fundamentales del PCI.

El curso pretende ser autocontenido, por lo que no son necesarios conocimientos especiales. No obstante, aquellos alumnos que ya tengan conocimientos, aunque sean muy someros, de Mecánica Cuántica encontrarán menos dificultades en los primeros temas del curso.

Se ha optado por dividir el temario del programa en cuatro temas que se corresponden con los cuatro capítulos de la presente obra. El primero de los capítulos revisa la base matemática de la Mecánica Cuántica, que no es otra que la teoría del espacio de Hilbert. Advertimos ya desde aquí que la sofisticación matemática en el tratamiento es baja. El segundo de los capítulos lo dedicamos a establecer la conexión entre el formalismo introducido en el capítulo precedente y los principios físicos de la Mecánica Cuántica. El tratamiento, aunque sigue en gran medida el enfoque "tradicional" en los libros de texto sobre Mecánica Cuántica, se desvía de éste en varios aspectos considerados normalmente "avanzados": la introducción de medidas generalizadas (las llamadas POVMs), la descripción del fenómeno del "entanglement", la introducción del concepto de superoperador, y un primer tratamiento cuantitativo del fenómeno de la decoherencia. El tercer capítulo se consigna por entero a reunir los resultados fundamentales de la teoría cuántica de de la información. Finalmente, el cuarto capítulo proporciona una breve panorámica de aplicaciones del procesado cuántico de la información.

Además de la bibliografía que se cita al final de este documento, existe una gran cantidad de información de muy alta calidad diseminada por Internet. Un buen lugar de partida en busca de estas fuentes es www.qubit.org. Otra gran fuente de bibliografía es el servidor de pre-prints ArXiV (xxx.arxiv.org), bajo el epígrafe quant-ph. De este servidor existe una copia actualizada en la Universidad de Zaragoza (xxx.unizar.es).

Finalmente, me gustaría agradecer la colaboración prestada por Marcos Pérez Suárez y Marcos Curty Alonso, dos excelentes alumnos de doctorado, en la elaboración del material reunido en este documento.

David J. Santos, noviembre de 2003.

Índice general

CAPÍTULO 1

El formalismo matemático de la Mecánica Cuántica

CONTENIDOS

La Física Cuántica constituye en la actualidad la base del conocimiento del mundo que nos rodea. A escala atómica y subatómica cualquier fenómeno precisa para su explicación de esta teoría; incluso fenómenos catalogables como macroscópicos requieren de argumentos cuánticos para su correcta asimilación.

Este primer capítulo está dirigido a aquellas personas que no estén familiarizadas con el formalismo hilbertiano de la Mecánica Cuántica. Su propósito, por tanto, no es otro que presentar de manera clara los aspectos de dicho formalismo matemático necesarios para comprender los capítulos posteriores.

1.1. Descripción cuántica de un sistema

Para poder entender someramente la diferencia entre la descripción clásica y cuántica de un sistema físico comenzaremos por el estudio de un caso muy sencillo: el sistema elemental formado por una sola partícula.

Desde el punto de vista clásico, la dinámica de este sistema viene dada por el conocimiento de su posición, $\mathbf{r}(t)$, y su velocidad, $\mathbf{v}(t)$. El estado del mismo está determinado, por tanto, para cada t, por seis parámetros, correspondientes a las distintas componentes de los campos posición y velocidad de la partícula. Por lo que respecta a la evolución del sistema, conocido su estado en

un instante inicial, es posible precisarlo en cualquier instante posterior mediante las leyes de la mecánica newtoniana. Por tanto, el estado es determinista[1], esto es, descrito con incertidumbre nula.

Por otra parte, desde el punto de vista cuántico, la descripción del sistema presenta notables diferencias con respecto a lo anteriormente expuesto para el caso clásico:

- **Estado:** El estado del sistema viene dado por el conocimiento de un campo escalar, denominado función de onda, $\psi(\mathbf{r}, t)$, en todos los puntos del espacio. Es decir, ahora el estado, para cada t, viene dado por un número infinito de parámetros, en lugar de los seis del caso clásico.

- **Evolución:** La evolución del estado del sistema está determinada por la ecuación de Schrödinger para la función de onda:

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r},t) + V(\mathbf{r},t)\psi(\mathbf{r},t) = i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t), \tag{1.1}$$

  donde m es la masa de la partícula, $V(\mathbf{r}, t)$ es el potencial al que ésta se ve sometida, y $\hbar$ es la constante de Planck normalizada: $\hbar = h/(2\pi)$, con $h \approx 6{,}6210^{-34}$ J·s. Así pues, conocida $\psi(\mathbf{r}, 0)$, es posible, mediante la resolución de (1.1), determinar $\psi(\mathbf{r}, t)$ para cualquier instante de tiempo posterior.

- **Interpretación probabilística:** Otra de las notables diferencias del formalismo cuántico con respecto al clásico es el grado de determinismo que ofrecen uno y otro. En el caso clásico se ha visto que la incertidumbre en la descripción del estado del sistema es nula; en el caso cuántico, la teoría es intrínsecamente probabilística. Así, para el caso sencillo que nos ocupa, la probabilidad de encontrar a nuestra partícula en el instante t en el volumen diferencial $\mathrm{d}^3r = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ alrededor del punto del espacio $\mathbf{r}$ viene dada por $|\psi(\mathbf{r}, t)|^2\,\mathrm{d}^3r$. Por tanto, la función de onda actúa en la teoría como función de densidad de probabilidad, y, como tal, ha de verificar:

$$\int |\psi(\mathbf{r},t)|^2 \mathrm{d}^3r = 1, \tag{1.2}$$

  donde la integración está extendida a todo el espacio[2].

1.2. El espacio de las funciones de onda $\mathcal{F}$

De acuerdo con la interpretación probabilística de la Mecánica Cuántica, la función de onda ha de verificar (1.2). Parece, pues, aconsejable restringir el estudio de las funciones de onda a aquellas tales que su módulo al cuadrado sea integrable. El espacio general de funciones de onda que verifican tal propiedad suele denominarse $\mathcal{L}^2$. Sin embargo, dado que estas funciones de onda satisfacen también ciertos criterios de regularidad (están acotadas, están definidas para cualquier $\mathbf{r}$, y son continuas e infinitamente diferenciables), se supondrá en lo sucesivo que pertenecen a un subespacio de $\mathcal{L}^2$ al que se denominará $\mathcal{F}$. Es posible demostrar[3] que dicho espacio es de Hilbert; es decir, es un espacio vectorial lineal, hay en él definido un producto escalar, y es completo. Con respecto a la segunda de estas características, y dado el uso que en posteriores ocasiones se

[1] Fue el científico francés marqués de Laplace el primero en argumentar, a principios del s.XIX, el carácter completamente determinista del universo. Laplace sugirió que debía existir un conjunto de leyes científicas que permitieran predecir absolutamente todos los eventos, con tal de conocer el estado completo del mismo en un instante de tiempo.

[2] Salvo que se indique explícitamente lo contrario, siempre será así.

[3] Véase el excelente texto de Abellanas y Galindo [1].

hará de ella, a continuación se proponen, sin demostración, las principales propiedades del producto escalar[4]:

- **Definición:** Dadas dos funciones de onda cualesquiera, $\psi(\mathbf{r})$ y $\phi(\mathbf{r})$, pertenecientes a $\mathcal{F}$, es posible definir su producto escalar, que se denotará mediante $(\psi(\mathbf{r}), \phi(\mathbf{r}))$, como el escalar, en general complejo:

$$(\psi, \phi) = \int \psi^*(\mathbf{r})\phi(\mathbf{r})\mathrm{d}^3 r, \tag{1.3}$$

  donde * denota complejo conjugado.

- **Propiedades:** Su verificación es inmediata a partir de (1.3):

  1. Simetría:

$$(\psi, \phi) = (\phi, \psi)^*. \tag{1.4}$$

  2. Linealidad con respecto a la segunda componente:

$$(\psi, \lambda_1\phi_1 + \lambda_2\phi_2) = \lambda_1(\psi, \phi_1) + \lambda_2(\psi, \phi_2). \tag{1.5}$$

  3. Antilinealidad con respecto a la primera componente:

$$(\lambda_1\psi_1 + \lambda_2\psi_2, \phi) = \lambda_1^*(\psi_1, \phi) + \lambda_2^*(\psi_2, \phi). \tag{1.6}$$

  4. Ortogonalidad: Si el producto escalar de dos funciones de onda es cero, entonces dichas funciones de onda se dicen ortogonales.
  5. Autoproducto escalar: El producto escalar de una función de onda consigo misma es siempre un escalar real positivo. El autoproducto es cero sólo si la función de onda es idénticamente cero ($\psi(\mathbf{r}) = 0$).
  6. Norma: A partir del producto escalar definido puede obtenerse una norma sobre el espacio $\mathcal{F}$: $(\psi, \psi)^{\frac{1}{2}}$.
  7. Desigualdad de Cauchy-Schwarz: Para cualquier par de funciones de onda ψ_1, ψ_2 se verifica la desigualdad:

$$|(\psi_1, \psi_2)| \leq (\psi_1, \psi_1)^{\frac{1}{2}}(\psi_2, \psi_2)^{\frac{1}{2}}. \tag{1.7}$$

     La igualdad sólo es cierta si ψ_1 y ψ_2 son proporcionales.

Sobre las funciones de onda de $\mathcal{F}$ pueden actuar operadores lineales[5]. Un operador lineal A es una entidad matemática que, al aplicarse sobre una función de onda $\psi(\mathbf{r}) \in \mathcal{F}$, la transforma, mediante una correspondencia lineal, en otra $\psi'(\mathbf{r})$ no necesariamente perteneciente a $\mathcal{F}$:

$$\psi'(\mathbf{r}) = A\psi(\mathbf{r}). \tag{1.8}$$

Se define el producto de dos operadores lineales A y B como

$$AB\psi(\mathbf{r}) = A(B\psi(\mathbf{r})). \tag{1.9}$$

Esta operación no es, en general, conmutativa; es decir, $AB\psi(\mathbf{r}) \neq BA\psi(\mathbf{r})$. De hecho, sólo si el llamado conmutador de ambos operadores, definido como $[A, B] = AB - BA$, se anula, es el producto conmutativo.

[4] Nótese que se ha omitido, para aliviar la notación, el parámetro t de la dependencia funcional de las funciones de onda.

[5] Los operadores lineales serán tratados con mayor profundidad en la Sección 1.6.

1.3. Bases del espacio $\mathcal{F}$

Se ha visto que el espacio $\mathcal{F}$ tiene la estructura de un espacio de Hilbert; se trata, por tanto, de un espacio vectorial de dimensión infinita. Dicho espacio puede ser expandido de forma única por ciertos conjuntos de funciones (no necesariamente pertenecientes a $\mathcal{F}$), cada uno de los cuales recibe el nombre de base del espacio $\mathcal{F}$. A continuación se presentan ambas posibilidades.

1.3.1 Bases pertenecientes a $\mathcal{F}$ Antes de abordar de manera rigurosa el concepto de base, se considerará la generalización de la propiedad de ortogonalidad del producto escalar, ya que será de utilidad en su definición. Así, dado un conjunto $\{u_i(\mathbf{r})\}_i$ de funciones pertenecientes a $\mathcal{F}$, dicho conjunto se dice ortonormal si sus elementos son mutuamente ortogonales y de norma unidad:

$$(u_i, u_j) = \int u_i^*(\mathbf{r}) u_j(\mathbf{r}) = \delta_{ij}, \tag{1.10}$$

donde δ_{ij} es la función de Kronecker (igual a 1 para $i = j$, y 0 para $i \neq j$).

Pues bien, un conjunto ortonormal de funciones se dice base de $\mathcal{F}$ si cualquier función de onda de $\mathcal{F}$, $\psi(\mathbf{r})$, puede expresarse de forma unívoca como combinación lineal de las funciones que constituyen el conjunto:

$$\psi(\mathbf{r}) = \sum_i c_i u_i(\mathbf{r}), \tag{1.11}$$

donde $c_i = (u_i, \psi)$. De esta manera, el conjunto de números complejos c_i representará a $\psi(\mathbf{r})$ en la base $\{u_i(\mathbf{r})\}_i$.

Una vez establecida una base, se puede expresar de manera muy sencilla el producto escalar de dos funciones de onda en función de sus componentes en dicha base. Es decir, dada $\{u_i(\mathbf{r})\}_i$ base de $\mathcal{F}$, y $\psi(\mathbf{r})$, $\phi(\mathbf{r}) \in \mathcal{F}$ tales que

$$\psi(\mathbf{r}) = \sum_i a_i u_i(\mathbf{r}), \tag{1.12}$$

$$\phi(\mathbf{r}) = \sum_j b_j u_j(\mathbf{r}), \tag{1.13}$$

se tiene que

$$(\psi, \phi) = \sum_i a_i^* b_i. \tag{1.14}$$

Puede demostrarse [28] que la condición para que el conjunto $\{u_i(\mathbf{r})\}_i$ sea una base se puede expresar como:

$$\sum_i u_i^*(\mathbf{r}') u_i(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}'), \tag{1.15}$$

donde $\delta(x)$ es la función Delta de Dirac. Este tipo de relación se denomina de clausura.

1.3.2 Bases no pertenecientes a $\mathcal{F}$

Sea ahora el conjunto de funciones $\{\omega_\alpha(\mathbf{r})\}_\alpha$ con α un parámetro de enumeración real. Este conjunto se dice ortonormal si

$$\int \omega_\alpha^*(\mathbf{r})\omega_{\alpha'}(\mathbf{r})\mathrm{d}^3r = \delta(\alpha - \alpha'). \tag{1.16}$$

Nótese que la norma de estas funciones diverge, por lo que no pertenecen a $\mathcal{L}^2$. El conjunto es, no obstante, una base si es posible una descomposición lineal unívoca de cualquier función de onda:

$$\psi(\mathbf{r}) = \int c_\alpha \omega_\alpha(\mathbf{r})\mathrm{d}\alpha, \tag{1.17}$$

con $c_\alpha = \int \omega_\alpha^*(\mathbf{r})\psi(\mathbf{r})\mathrm{d}^3r$.

Al igual que en el caso de bases pertenecientes a $\mathcal{F}$, dadas dos funciones de onda $\psi(\mathbf{r})$, $\phi(\mathbf{r}) \in \mathcal{F}$ cuyos componentes en una base concreta $\{\omega_\alpha(\mathbf{r})\}_\alpha$ son conocidos,

$$\psi(\mathbf{r}) = \int a_\alpha \omega_\alpha(\mathbf{r})\mathrm{d}\alpha, \tag{1.18}$$

$$\phi(\mathbf{r}) = \int b_{\alpha'} \omega_{\alpha'}(\mathbf{r})\mathrm{d}\alpha', \tag{1.19}$$

su producto escalar se puede calcular como

$$(\psi, \phi) = \int a_\alpha^* b_\alpha \mathrm{d}\alpha. \tag{1.20}$$

De forma similar a (1.15), es posible también en este caso definir una relación de clausura [28]:

$$\int \omega_\alpha^*(\mathbf{r}')\omega_\alpha(\mathbf{r})\mathrm{d}\alpha = \delta(\mathbf{r} - \mathbf{r}'). \tag{1.21}$$

Para finalizar con el concepto de bases del espacio $\mathcal{F}$, conviene realizar un breve comentario. De lo expuesto hasta al momento cabría deducir que todas las bases de $\mathcal{F}$ han de ser conjuntos ortonormales de funciones, y esto no es así. No existe razón alguna por la cual un conjunto no ortogonal, pero completo, no pueda ser una base; basta con que verifique la relación de clausura adecuada.

1.4. Notación de Dirac. Espacio de estados $\mathcal{E}$

En las secciones precedentes hemos visto que el estado cuántico de una partícula viene dado, en un instante determinado, por una función de onda $\psi(\mathbf{r}, t)$ perteneciente a un espacio vectorial $\mathcal{F}$ en el que se definen unas bases que permiten expresar de forma alternativa la información contenida en la misma. Tanto para el caso de una base discreta como continua, esta expresión alternativa requiere, además del conjunto de funciones que actúa de base, del conocimiento de los coeficientes de expansión (véanse las ecuaciones (1.11) y (1.17)). Por tanto, dada una cierta base, son necesarios únicamente dichos coeficientes para definir unívocamente el estado del sistema. Esto sugiere caracterizar éste mediante la secuencia de los c_i o c_α en lugar de mediante la función de onda original. Cabe así definir el vector de estado, que se denotará mediante $|\psi\rangle$, como el vector compuesto por todos los coeficientes de expansión. Sin embargo, está introducción del concepto de vector de estado no debe ser contemplada únicamente desde este punto de vista, es decir, como una simplificación del formalismo, sino como una generalización, ya que existen sistemas cuya descripción cuántica no puede realizarse mediante una función de onda y sí mediante un vector de estado.

Por tanto, para cada instante de tiempo, el estado cuántico de una partícula se caracterizará mediante un vector $|\psi\rangle$ perteneciente a un cierto espacio vectorial abstracto, $\mathcal{E}$, llamado espacio de estados de la partícula. Esta notación fue introducida por el físico Paul A. M. Dirac, de ahí su nombre.

En esta sección se reescribirán, bajo la notación de Dirac, los principales resultados obtenidos hasta el momento.

1.4.1 Kets y bras Es un resultado conocido de Álgebra Lineal que a cada espacio vectorial se le puede asociar un espacio, también vectorial, dual. En efecto, considérese un funcional lineal arbitrario χ de $\mathcal{E}$ que a cada ket $|\psi\rangle$ de $\mathcal{E}$ le asocia un número complejo:

$$|\psi\rangle \in \mathcal{E} \xrightarrow{\chi} \chi(|\psi\rangle) \in \mathbb{C}. \tag{1.22}$$

Puede demostrarse que el conjunto de todos los funcionales lineales definibles sobre $\mathcal{E}$ posee estructura de espacio de Hilbert. Denominaremos a dicho espacio $\mathcal{E}^*$. Dirac denominó originalmente a los vectores de $\mathcal{E}$ "kets", y a los funcionales χ de $\mathcal{E}^*$ "bras". La razón de esta denominación es aparente cuando se escribe la acción del funcional χ sobre el vector $|\psi\rangle$ de la forma $\langle\chi|\psi\rangle$. Los dos elementos de esta operación, χ y ψ, se hallan entre "brackets" (en inglés, los símbolos '<'y '>').

Una de las principales ventajas de la notación de Dirac es que permite escribir el producto escalar definido sobre $\mathcal{E}$ —véase la ecuación (1.3)— de una forma más compacta basándose en la correspondencia existente entre kets y bras. En efecto, dado un ket $|\psi\rangle$ perteneciente a $\mathcal{E}$, siempre se le puede asociar un bra $\langle\psi|$ perteneciente a su espacio dual $\mathcal{E}^*$ cuya acción sobre cualquier $|\phi\rangle \in \mathcal{E}$ sea asignarle el número complejo $(|\psi\rangle, |\phi\rangle)$, es decir, el producto escalar entre el ket al que está asociado $\langle\psi|$ y $|\phi\rangle$. Utilizando, por tanto, esta notación se tiene que el producto escalar entre dos vectores del espacio $\mathcal{E}$ se puede expresar como:

$$\langle\psi|\phi\rangle = (|\psi\rangle, |\phi\rangle). \tag{1.23}$$

A continuación se repiten, de manera concisa, las principales propiedades del producto escalar en la notación de Dirac, ya que será la empleada casi con exclusividad a partir de este momento:

$$\langle\psi|\phi\rangle = \langle\phi|\psi\rangle^*, \tag{1.24}$$

$$\langle\psi|\lambda_1\phi_1 + \lambda_2\phi_2\rangle = \lambda_1\langle\psi|\phi_1\rangle + \lambda_2\langle\psi|\phi_2\rangle, \tag{1.25}$$

$$\langle\lambda_1\psi_1 + \lambda_2\psi_2|\phi\rangle = \lambda_1^*\langle\psi_1|\phi\rangle + \lambda_2^*\langle\psi_2|\phi\rangle, \tag{1.26}$$

$$\langle\psi|\psi\rangle \quad \text{es real y positivo; sólo es cero si } |\psi\rangle = 0. \tag{1.27}$$

Se ha mostrado cómo a cada ket se le puede asociar un bra, pero, ¿es posible esta misma correspondencia a la inversa? La respuesta es no. En general, el espacio dual $\mathcal{E}^*$ y el espacio de estados $\mathcal{E}$ no son isomorfos, salvo, por supuesto, que $\mathcal{E}$ sea de dimensión finita. Para paliar esta deficiencia, se utilizan los llamados kets generalizados que presentan norma infinita, pero cuyo producto escalar con cualquier ket de $\mathcal{E}$ es finito.

1.4.2 Bases del espacio de estados $\mathcal{E}$ De manera similar a como analizamos la existencia de bases del espacio de funciones de onda $\mathcal{F}$ en la Sección 1.3, trataremos ahora las bases del espacio de estados $\mathcal{E}$.

Dado un conjunto de kets $\{|u_i\rangle\}_i \in \mathcal{E}$ o un conjunto de kets generalizados $\{|\omega_\alpha\rangle\}_\alpha \notin \mathcal{E}$, éste se dice ortonormal si cumple la siguiente condición:

$$\langle u_i|u_j\rangle = \delta_{ij}, \tag{1.28}$$

o

$$\langle \omega_\alpha | \omega_{\alpha'} \rangle = \delta(\alpha - \alpha'). \tag{1.29}$$

Pues bien, un conjunto ortonormal[6] de kets o de kets generalizados se dice base de $\mathcal{E}$, discreta o continua, respectivamente, si sólo hay una única forma de expresar cualquier $|\psi\rangle \in \mathcal{E}$ como combinación lineal de los elementos que constituyen el conjunto:

$$|\psi\rangle = \sum_i c_i |u_i\rangle, \tag{1.30}$$

o

$$|\psi\rangle = \int c_\alpha |\omega_\alpha\rangle \mathrm{d}\alpha, \tag{1.31}$$

con $c_i = \langle u_i | \psi \rangle$ y $c_\alpha = \langle \omega_\alpha | \psi \rangle$. Esta condición también puede expresarse mediante una relación de clausura [28]:

$$\sum_i |u_i\rangle\langle u_i| = 1, \tag{1.32}$$

o

$$\int |\omega_\alpha\rangle\langle\omega_\alpha| \mathrm{d}\alpha = 1, \tag{1.33}$$

donde 1 denota el operador identidad en $\mathcal{E}$.

1.5. Representaciones en el espacio de estados $\mathcal{E}$

Una representación no es más que la elección de una base en el espacio de estados $\mathcal{E}$, bien sea continua o discreta, con la intención de conseguir que el cálculo con kets, bras y operadores se convierta en un cálculo matricial. Pese a que la decisión en un principio es arbitraria, parece obvio escoger aquella base que facilite el cálculo lo más posible.

1.5.1 Representación de kets Dada una base discreta $\{|u_i\rangle\}_i$ o continua $\{|\omega_\alpha\rangle\}_\alpha$, un ket $|\psi\rangle$ cualquiera perteneciente a $\mathcal{E}$ se representa mediante un vector columna compuesto por sus componentes en la base:

$$\begin{pmatrix} \langle u_1|\psi\rangle \\ \langle u_2|\psi\rangle \\ \vdots \\ \langle u_i|\psi\rangle \\ \vdots \end{pmatrix}, \tag{1.34}$$

para el caso de una base discreta, y

$$\begin{pmatrix} \vdots \\ \langle \omega_\alpha|\psi\rangle \\ \vdots \end{pmatrix}, \tag{1.35}$$

si ésta es continua.

[6] La condición de ortonormalidad, al igual que en el espacio $\mathcal{F}$, no es necesaria, pero sí muy útil.

1.5.2 Representación de bras Sea $\langle\psi|$ un bra arbitrario perteneciente a $\mathcal{E}^*$ y sean las bases discreta y continua $\{|u_i\rangle\}_i$ y $\{|\omega_\alpha\rangle\}_\alpha$, respectivamente. Siempre se puede expresar $\langle\psi|$ de forma única en función de los bras $\langle u_i|$ o $\langle\omega_\alpha|$:

$$\langle\psi| = \sum_i \langle\psi|u_i\rangle\langle u_i|, \tag{1.36}$$

o

$$\langle\psi| = \int \langle\psi|\omega_\alpha\rangle\langle\omega_\alpha|\mathrm{d}\alpha, \tag{1.37}$$

donde los componentes de $\langle\psi|$, tanto en la base discreta como continua, $\langle\psi|u_i\rangle$ y $\langle\psi|\omega_\alpha\rangle$, respectivamente, no son más que los complejos conjugados de los componentes de su ket asociado en esas mismas bases. De esta manera, su representación se realiza mediante un vector fila:

$$(\langle\psi|u_1\rangle\langle\psi|u_2\rangle\cdots\langle\psi|u_i\rangle\cdots), \tag{1.38}$$

$$(\cdots\langle\psi|\omega_\alpha\rangle\cdots). \tag{1.39}$$

Resulta útil observar que, mediante este convenio, $\langle\psi|\phi\rangle$ no es más que el producto de un vector fila por un vector columna, obteniéndose, como era de esperar, un número, en general complejo, como resultado.

1.5.3 Representación de operadores Al introducir, en la Sección 1.2, el espacio $\mathcal{F}$ de las funciones de onda, vimos que sobre éstas podían actuar operadores lineales. De igual manera, éstos también pueden actuar sobre $\mathcal{E}$. Pese a que la teoría correspondiente se presentará más detenidamente en la siguiente sección, mencionaremos aquí brevemente que su representación se corresponde con una matriz cuyos elementos se calculan, según la base sea discreta o continua, como

$$A_{ij} = \langle u_i|A|u_j\rangle, \tag{1.40}$$

o

$$A(\alpha,\alpha') = \langle\omega_\alpha|A|\omega_{\alpha'}\rangle. \tag{1.41}$$

1.5.4 Cambio de representación Hasta el momento se ha mostrado cómo, para una base concreta, kets, bras y operadores pueden representarse de manera matricial. Se verá a continuación cómo se relacionan dos representaciones arbitrarias cualesquiera[7].

Sean dos bases discretas de $\mathcal{E}$, $\{|u_i\rangle\}_i$ y $\{|v_k\rangle\}_k$. Se define la matriz de cambio de base B como

$$B_{ik} = \langle u_i|v_k\rangle. \tag{1.42}$$

Pues bien, dado un ket arbitrario $|\psi\rangle \in \mathcal{E}$, es inmediato comprobar que la relación entre sus componentes en cada

$$\langle v_k|\psi\rangle = \sum_i B^\dagger_{ki}\langle u_i|\psi\rangle. \tag{1.43}$$

Por otro lado, si se toma su bra asociado, $\langle\psi|$, se tiene

[7]Por simplicidad se considera tan solo el caso de bases discretas en $\mathcal{E}$. La generalización al caso continuo no resulta difícil a partir de lo aquí expuesto.

$$\langle\psi|v_k\rangle = \sum_i B_{ik}\langle\psi|u_i\rangle. \tag{1.44}$$

De manera similar, y considerando un operador lineal A cualquiera con componentes A_{ij} en la base $\{|u_i\rangle\}_i$, su representación en $\{|v_k\rangle\}_k$ tiene por elementos de la matriz (en función de los A_{ij}):

$$A_{kl} = \langle v_k|A|v_l\rangle = \sum_{ij} B^\dagger_{ki}A_{ij}B_{jl}. \tag{1.45}$$

1.6. Operadores lineales

Un operador lineal A asocia, mediante una correspondencia lineal, a cada ket $|\psi\rangle \in \mathcal{E}$ otro ket $|\psi'\rangle \in \mathcal{E}$:[8]

$$|\psi'\rangle = A|\psi\rangle. \tag{1.46}$$

Pero A también puede actuar sobre el espacio dual $\mathcal{E}^*$, es decir, sobre los bras. Considérese un bra cualquiera $\langle\psi|$ y el conjunto de todos los kets $|\phi\rangle \in \mathcal{E}$. A cada uno de estos kets se le puede hacer corresponder el número complejo $\langle\psi|(A|\phi\rangle)$. Dado que A es lineal, y teniendo en cuenta la propiedad de linealidad con respecto a la segunda componente del producto escalar, se tiene que, una vez fijado el operador A y el bra $\langle\psi|$, se ha definido un nuevo funcional lineal que actúa sobre el espacio de kets $\mathcal{E}$. Este nuevo bra se denotará mediante $\langle\psi|A$. Así pues, la acción de un operador lineal sobre el espacio dual $\mathcal{E}^*$ se puede expresar como:

$$\langle\psi| \in \mathcal{E}^* \xrightarrow{A} \langle\psi|A \in \mathcal{E}^*. \tag{1.47}$$

Sea ahora $|\psi\rangle$ el ket dual del bra $\langle\phi|A$. Dado que la correspondencia ket-bra es antilineal, $|\psi\rangle$ puede interpretarse como la acción de un operador lineal, que se denotará mediante $A^\dagger$, sobre $|\phi\rangle$, el ket asociado al bra $\langle\phi|$. El operador $A^\dagger$, relacionado con A de la manera que se acaba de mostrar, se denomina el hermítico conjugado de A.

Puesto que $|\psi\rangle = A^\dagger|\phi\rangle$ es el ket asociado al bra $\langle\phi|A$, el producto escalar de $|\psi\rangle$ con un ket arbitrario $|\alpha\rangle$ ha de ser, en virtud de la propiedad de simetría del producto escalar, el conjugado del producto escalar entre $|\alpha\rangle$ y $|\psi\rangle$, es decir:

$$\langle\alpha|A^\dagger|\psi\rangle = \langle\psi|A|\alpha\rangle^*, \tag{1.48}$$

de donde se deduce que $(A^\dagger)^\dagger = A$; o lo que es lo mismo: la conjugación es una operación hermítica.

De forma similar pueden deducirse las siguientes relaciones:

$$(\lambda A)^\dagger = \lambda^* A^\dagger, \tag{1.49}$$

$$(A + B)^\dagger = A^\dagger + B^\dagger, \tag{1.50}$$

$$(AB)^\dagger = B^\dagger A^\dagger, \tag{1.51}$$

$$(|\alpha\rangle\langle\beta|)^\dagger = |\beta\rangle\langle\alpha|, \tag{1.52}$$

con A y B dos operadores lineales cualesquiera y λ un escalar arbitrario.

[8] Aunque, como ya avanzamos, $|\psi'\rangle$ puede no pertenecer a $\mathcal{E}$, en principio ese caso no es de interés.

Antes de continuar, conviene realizar un pequeño inciso acerca de la expresión (1.52). Se trata de una igualdad entre dos operadores lineales; para comprobarlo, no hay más que considerar un ket arbitrario $|\psi\rangle$ al que aplicarle $|\alpha\rangle\langle\beta|$ para obtener el ket $|\alpha\rangle\langle\beta|\psi\rangle$. Luego $|\alpha\rangle\langle\beta|$ no es más que un operador lineal.

A la vista de los resultados anteriores, se da la siguiente regla para obtener el resultado de la conjugación sobre una expresión algebraica: sustituir todos los escalares y operadores por sus complejos conjugados, todos los kets por sus bras y viceversa, y, finalmente, invertir el orden de los diversos símbolos.

1.6.1 Autovalores y autovectores de operadores lineales Como se verá más adelante, la medida de una magnitud física lleva asociado el cálculo del espectro o conjunto de autovalores de un operador lineal asociado a dicha magnitud. Este hecho justifica que se dedique atención a las peculiaridades algebraicas de este cálculo.

Sea A un operador lineal. Por definición, el escalar, en general complejo, a es autovalor de A, y el ket $|\psi\rangle$ autoestado asociado a a s:

$$A|\psi\rangle = a|\psi\rangle. \tag{1.53}$$

Evidentemente, si $|\psi\rangle$ es autoestado de A asociado a a, el ket $c|\psi\rangle$, con c un escalar arbitrario, también lo será. Resulta inmediato verificar que si existen varios autoestados de A asociados a a linealmente independientes, cualquier combinación lineal de estos kets es autoestado de A asociado a a. Además, estos autoestados linealmente independientes determinan un espacio vectorial. Si este espacio —un cierto subespacio de $\mathcal{E}$— es unidimensional, el autovalor a se dice sencillo; en caso contrario, degenerado.

De forma más general, teniendo en cuenta la posible degeneración, el problema de autovalores de un operador lineal A puede escribirse de la forma

$$A|\psi_n^m\rangle = a_n|\psi_n^m\rangle. \tag{1.54}$$

El significado de la notación es el siguiente: n enumera los distintos autovalores de A (los a_n), para cada uno de los cuales existe un conjunto de kets linealmente independientes, $\{|\psi_n^m\rangle, m = 1, \ldots, g_n\}$, donde g_n es el índice de degeneración del autovalor. Como se ha visto, g_n es a su vez la dimensión del subespacio generado por el conjunto $\{|\psi_n^m\rangle\}_m$.

1.6.2 Traza de un operador lineal Dado un operador lineal A y una representación del mismo en la base $\{|u_i\rangle\}_i$, se denomina traza de A, y se denota mediante $\mathrm{tr}\{A\}$, a la suma de los componentes de la diagonal principal[9] de su representación matricial, es decir,

$$\mathrm{tr}\{A\} = \sum_i \langle u_i|A|u_i\rangle = \sum_i A_{ii}. \tag{1.55}$$

Si la base utilizada fuese continua, $\{|\omega_\alpha\rangle\}_\alpha$, se tendría que

$$\mathrm{tr}\{A\} = \int \langle\omega_\alpha|A|\omega_\alpha\rangle \mathrm{d}\alpha = \int A(\alpha,\alpha)\mathrm{d}\alpha. \tag{1.56}$$

A partir de lo expuesto en la Subsección 1.5.4, es claro que la traza de un operador no depende de la representación elegida.

[9] En el caso de un espacio de estados $\mathcal{E}$ de dimensión infinita, la traza de un operador A sólo se define si las expresiones (1.55) y (1.56) convergen.

1.6.3 Restricción de un operador lineal a un subespacio

Sea un operador lineal arbitrario A que actúa en $\mathcal{E}$. Se define la restricción de A en un subespacio $\mathcal{E}_q$ (q-dimensional) de $\mathcal{E}$, y se denota por A_q, al operador:

$$A_q = P_q A P_q, \tag{1.57}$$

donde

$$P_q = \sum_{i=1}^{q} |u_i\rangle\langle u_i|, \tag{1.58}$$

con $\{|u_i\rangle\}_i$, $i = 1, \ldots, q$, una base ortonormal de $\mathcal{E}_q$. Posteriormente, en la Subsección 1.6.6, se verá que P_q no es más que el operador proyector en el subespacio $\mathcal{E}_q$.

1.6.4 Teoremas básicos de álgebra de operadores

Se muestran a continuación, sin demostración[10], algunos resultados fundamentales para dos operadores A y B cuyo conmutador no es cero. En algún caso se considerarán funciones de A y B sobre el espacio de estados $\mathcal{E}$; supondremos siempre que éstas pueden ser desarrolladas en serie de potencias:

$$F(B) = \sum_{n=0}^{\infty} c_n B^n, \tag{1.59}$$

con c_n los coeficientes del desarrollo.

Teorema 1 *Dado un operador cualquiera A y una función arbitraria de este operador, $F(A)$, $[F(A), A] = 0$.*

Teorema 2 *Si A y B son dos operadores no conmutables, y si ξ es un escalar arbitrario, entonces, si n es entero:*

$$e^{\xi A} B^n e^{-\xi A} = (e^{\xi A} B e^{-\xi A})^n, \tag{1.60}$$

y

$$e^{\xi A} F(B) e^{-\xi A} = F(e^{\xi A} B e^{-\xi A}). \tag{1.61}$$

Teorema 3 *Si A y B son dos operadores no conmutables, y el operador inverso de A, A^{-1}, existe, se tiene que:*

$$A B^n A^{-1} = (A B A^{-1})^n, \tag{1.62}$$

con n entero, y

$$A F(B) A^{-1} = F(A B A^{-1}). \tag{1.63}$$

Teorema 4 *Si A y B son dos operadores no conmutables, y ξ es un escalar arbitrario, entonces:*

$$e^{\xi A} B e^{-\xi A} = B + \xi[A, B] + \frac{\xi^2}{2!}[A, [A, B]] + \frac{\xi^3}{3!}[A, [A, [A, B]]] + \cdots \tag{1.64}$$

[10] Al lector interesado se le remite a [28].

Teorema 5 *Si A y B son dos operadores no conmutables que satisfacen las condiciones*

$$[A,[A,B]] = [B,[A,B]] = 0, \tag{1.65}$$

se cumple que

$$e^{A+B} = e^A e^B e^{-\frac{1}{2}[A,B]} = e^B e^A e^{\frac{1}{2}[A,B]}. \tag{1.66}$$

1.6.5 Operadores lineales más relevantes Hasta el momento, se han caracterizado de forma genérica los operadores lineales que actúan sobre los elementos del espacio $\mathcal{E}$. A partir de esta sección se presentarán algunos que, por sus peculiaridades, resultan, como se tendrá ocasión de apreciar más adelante, especialmente interesantes en Mecánica Cuántica.

- Un operador lineal H es **hermítico** si es su propio conjugado: $H = H^\dagger$.
- Un operador lineal I es **antihermítico** si es el opuesto de su conjugado: $I = -I^\dagger$.
- Cualquier operador lineal A puede escribirse de forma única como suma de dos operadores, uno hermítico y otro antihermítico: $A = H_A + I_A$, con $H_A = (A + A^\dagger)/2$ y $I_A = (A - A^\dagger)/2$.
- Se llama **inverso** de un operador A al operador A^{-1} que verifica $AA^{-1} = A^{-1}A = 1$.
- Un operador U es **unitario** si es el inverso de su propio hermítico conjugado: $UU^\dagger = U^\dagger U = 1$, donde con 1 se denota, tanto en este párrafo como en el anterior, el operador identidad. La propiedad característica de este tipo de operadores es que su actuación sobre los kets de un espacio de estados $\mathcal{E}$ de dimensión finita conserva el producto escalar. De esta manera, la acción de un operador unitario U sobre una base ortonormal de $\mathcal{E}$ da como resultado otra base también ortonormal.

1.6.6 Operadores proyección Un tipo de operador hermítico muy versátil en el álgebra de operadores es el operador proyección o proyector. Se denomina así a cualquier operador hermítico P que verifica $P^2 = P$. Es inmediato comprobar que todo operador que cumple esta condición lleva asociado un subespacio de $\mathcal{E}$ sobre el que "proyecta" el ket genérico sobre el que actúa.

Considérese el ejemplo del proyector elemental. Sea el ket normalizado $|\psi_1\rangle$ tal que $\langle\psi_1|\psi_1\rangle = 1$. Este ket subtiende un espacio de dimensión 1 al que se denominará $\mathcal{E}_1$. Sea ahora el operador P_{ψ_1} definido como $P_{\psi_1} = |\psi_1\rangle\langle\psi_1|$, y sea un ket arbitrario $|\phi\rangle$. Si se llama $|\phi_1\rangle$ al ket resultante de aplicar el operador P_{ψ_1} sobre $|\phi\rangle$, es decir, $|\phi_1\rangle = P_{\psi_1}|\phi\rangle = |\psi_1\rangle\langle\psi_1|\phi\rangle$. No hay duda de que $|\phi_1\rangle$ es la proyección sobre $\mathcal{E}_1$ de $|\phi\rangle$. El coeficiente de proporcionalidad, obviamente, es $\langle\psi_1|\phi\rangle$. El significado geométrico del operador P_{ψ_1} está por tanto claro: no es más que la proyección ortogonal en el espacio $\mathcal{E}_1$.

Es fácil verificar que los autovalores de P_{ψ_1} son 1, sencillo, y 0, con índice de degeneración infinito. Además, todo ket perteneciente a $\mathcal{E}$ puede ser descompuesto de forma única como la suma de su proyección sobre un espacio $\mathcal{E}_a$ y su proyección sobre un espacio $\mathcal{E}_b$, complementario al primero.

La generalización del concepto de operador proyección sobre espacios de dimensión mayor que uno es muy sencilla a partir de lo anterior. Así, sea $|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_q\rangle$, un conjunto ortonormal de vectores que subtienden un espacio $\mathcal{E}_q$ de dimensión q. Pues bien, el operador proyección sobre $\mathcal{E}_q$ no es más que:

$$P_q = \sum_{i=1}^{q} |\psi_i\rangle\langle\psi_i|. \tag{1.67}$$

De esta manera, y recordando la acción de la restricción A_q al subespacio $\mathcal{E}_q$ sobre un ket cualquiera $|\psi\rangle \in \mathcal{E}$, puede entenderse ésta como la proyección en $\mathcal{E}_q$ del ket resultado de aplicar A a la proyección de $|\psi\rangle$ en dicho subespacio.

1.6.7 Operadores observables Se trata de una clase de operadores hermíticos que, como se verá posteriormente, desempeñan un importante papel en la interpretación física de la Teoría Cuántica. Pero antes de abordar su definición, veremos dos propiedades importantes de los autovalores y autovectores de un operador hermítico.

Considérese un operador A hermítico al que, por simplicidad, se le supondrá un espectro discreto de autovalores $\{a_n; n = 1, 2, \ldots\}$ con índice de degeneración g_n. De acuerdo con la notación introducida, $|\psi_n^m\rangle$ $(m = 1, 2, \ldots, g_n)$ representará el conjunto de kets linealmente independientes asociados al autovalor a_n que determinan el subespacio $\mathcal{E}_n$. Obviamente, este conjunto siempre puede ser escogido ortonormal. Se puede demostrar que [28]

- Los autovalores de un operador hermítico son reales.
- Los autoestados de un operador hermítico asociados a diferentes autovalores son ortogonales.

Es inmediato comprobar, por tanto, en base a la segunda propiedad, y al hecho de que el conjunto de autoestados de un autovalor a_n cualquiera pueda ser elegido ortonormal, que todo operador hermítico posee un conjunto ortonormal de autoestados. Pues bien, si dicho conjunto constituye una base de $\mathcal{E}$, es decir, verifica una relación de clausura del tipo:

$$\sum_{n=1}^{\infty} \sum_{m=1}^{g_n} |\psi_n^m\rangle\langle\psi_n^m| = 1, \tag{1.68}$$

se dice que el operador hermítico es un observable. A partir de esta definición es claro que A se puede expresar como

$$A = \sum_n a_n P_n, \tag{1.69}$$

donde P_n es el proyector sobre el espacio determinado por los autovectores asociados al autovalor a_n. Además, su traza se puede obtener como:

$$\mathrm{tr}\{A\} = \sum_n a_n. \tag{1.70}$$

Para comprobarlo no hay más que calcularla utilizando sus autoestados como base para la representación matricial.

1.7. Producto tensorial de espacios de estados

En lo que va de capítulo, por simplicidad en el formalismo, se ha supuesto que nuestro sistema cuántico elemental estaba formado por una partícula de masa m. Se ha reunido el conocimiento del estado dinámico del sistema, bajo la notación de Dirac, en el ket $|\psi\rangle$ perteneciente a un cierto espacio de estados $\mathcal{E}$. En esta sección complicaremos ligeramente nuestro sistema: ahora éste estará compuesto por dos partículas lo suficientemente separadas como para que podamos despreciar su interacción.

De acuerdo con el formalismo introducido en secciones anteriores, podemos asignar a cada una de las dos partículas de nuestro sistema un estado $|\psi_i\rangle$ y un espacio de estados $\mathcal{E}_i$ (con $i = 1, 2$). Sin embargo, cabe preguntarse si existirá alguna forma de describir conjuntamente el sistema mediante un único estado $|\psi\rangle$ de un cierto espacio global de estados $\mathcal{E}_{\mathrm{Total}}$. Pues bien, tal descripción es

posible en el espacio de estados resultante del producto tensorial de los espacios de estados $\mathcal{E}_1$ y $\mathcal{E}_2$: $\mathcal{E}_{\text{Total}} = \mathcal{E}_1 \otimes \mathcal{E}_2$. En este espacio, $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$ suele expresarse, de forma simplificada, como $|\psi_1, \psi_2\rangle$ o $|\psi_1\rangle|\psi_2\rangle$. A continuación se presentan algunos resultados fundamentales de álgebra sobre el espacio producto tensorial.

- **Linealidad con respecto al producto por escalares:**

$$(\lambda|\psi_1\rangle) \otimes |\psi_2\rangle = \lambda(|\psi_1\rangle \otimes |\psi_2\rangle), \tag{1.71}$$

$$|\psi_1\rangle \otimes (\lambda|\psi_2) = \lambda(|\psi_1\rangle \otimes |\psi_2\rangle). \tag{1.72}$$

- **Distributividad con respecto a la suma de kets:**

$$|\psi_1\rangle \otimes (|\psi_2\rangle + |\phi_2\rangle) = |\psi_1\rangle \otimes |\psi_2\rangle + |\psi_1\rangle \otimes |\phi_2\rangle, \tag{1.73}$$

$$(|\psi_1\rangle + |\phi_1\rangle) \otimes |\psi_2\rangle = |\psi_1\rangle \otimes |\psi_2\rangle + |\phi_1\rangle \otimes |\psi_2\rangle. \tag{1.74}$$

- **Bases del espacio de estados total:** Dada la base $\{|u_i\rangle\}_i$ de $\mathcal{E}_1$ y $\{|v_j\rangle\}_j$ de $\mathcal{E}_2$, el conjunto de kets $\{|u_i\rangle \otimes |v_j\rangle\}_{ij}$ es base del espacio total $\mathcal{E}_{\text{Total}}$. Lógicamente, si la dimensión de $\mathcal{E}_1$ fuera N_1, y la de $\mathcal{E}_2$ N_2, el espacio $\mathcal{E}_{\text{Total}}$ sería de dimensión $N_1 \cdot N_2$. Además, si $|\psi_1\rangle = \sum_i a_i|u_i\rangle$ y $|\psi_2\rangle = \sum_j b_j|v_j\rangle$, entonces

$$|\psi_1\rangle \otimes |\psi_2\rangle = \sum_{i,j} a_i b_j |u_i\rangle \otimes |v_j\rangle. \tag{1.75}$$

  Finalmente, un estado arbitrario $|\psi\rangle$ del espacio $\mathcal{E}_{\text{Total}}$ se escribe, en función de la base $\{u_i\rangle \otimes |v_j\rangle\}_{ij}$, como

$$|\psi\rangle = \sum_{i,j} c_{ij} |u_i\rangle \otimes |v_j\rangle. \tag{1.76}$$

- **Producto escalar:** Se define el producto escalar entre los kets $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$ y $|\phi\rangle = |\phi_1\rangle \otimes |\phi_2\rangle$ del espacio total $\mathcal{E}_{\text{Total}}$ como

$$\langle\phi|\psi\rangle = \langle\phi_1, \phi_2|\psi_1, \psi_2\rangle = \langle\phi_1|\psi_1\rangle\langle\phi_2|\psi_2\rangle. \tag{1.77}$$

- **Operadores:** Dados los operadores A_1 y A_2 pertenecientes, respectivamente, a los espacios $\mathcal{E}_1$ y $\mathcal{E}_2$, se define la acción del operador $A_1 \otimes A_2$ sobre el ket $|\psi_1\rangle \otimes |\psi_2\rangle$ de la forma

$$(A_1 \otimes A_2)(|\psi_1\rangle \otimes |\psi_2\rangle) = (A_1|\psi_1\rangle) \otimes (A_2|\psi_2\rangle). \tag{1.78}$$

  En algunas ocasiones puede ser necesario determinar el estado resultante de la operación del operador A_1, por ejemplo, sobre $|\psi_1\rangle \otimes |\psi_2\rangle$. Dado que A_1 no está definido sobre $\mathcal{E}_{\text{Total}}$, se recurre a hacer operar sobre $\mathcal{E}_{\text{Total}}$ a la extensión sobre este espacio de A_1. Este operador extendido, representado por $\widetilde{A}_1$, actúa de la forma siguiente:

$$\widetilde{A}_1(|\psi_1\rangle \otimes |\psi_2\rangle) = (A_1|\psi_1\rangle) \otimes |\psi_2\rangle. \tag{1.79}$$

  Se puede así interpretar $\widetilde{A}_1$ como $A_1 \otimes 1_2$, donde 1_2 es el operador identidad en el espacio $\mathcal{E}_2$.

- **Autovalores y autovectores de operadores extendidos:** Es inmediato comprobar a partir de (1.79) que si un operador A_1 perteneciente al espacio $\mathcal{E}_1$ tiene, por ejemplo, un espectro discreto de autovalores,

$$A_1|\psi_{1n}^m\rangle = a_{1n}|\psi_{1n}^m\rangle \quad ; \quad m = 1, \ldots, g_n, \tag{1.80}$$

el operador extendido $\widetilde{A}_1$ perteneciente al espacio $\mathcal{E}_{\text{Total}} = \mathcal{E}_1 \otimes \mathcal{E}_2$ presenta el problema de autovalores

$$\widetilde{A}_1(|\psi_{1n}^m\rangle \otimes |\phi_2\rangle) = a_{1n}(|\psi_{1n}^m\rangle \otimes |\phi_2\rangle) \quad ; m = 1, \ldots, g_n, \tag{1.81}$$

donde $|\phi_2\rangle$ representa a un ket cualquiera perteneciente a $\mathcal{E}_2$.

A partir de estos resultados se puede concluir lo siguiente:

- Si A_1 es un observable en $\mathcal{E}_1$, entonces el operador extendido $\widetilde{A}_1$ también será un observable en $\mathcal{E}_{\text{Total}}$.
- El espectro de $\widetilde{A}_1$ en $\mathcal{E}_{\text{Total}}$ es el mismo que el de A_1 en $\mathcal{E}$.
- Un autovalor a_{1n} con índice de degeneración g_n en $\mathcal{E}_1$ tendrá en $\mathcal{E}_{\text{Total}}$ un índice de degeneración $N_2 \cdot g_n$, donde N_2 es la dimensión del espacio $\mathcal{E}_2$.

Por simplicidad, el formalismo descrito en esta sección se ha limitado al producto tensorial de dos espacios de estados. En cualquier caso, no ofrece la más mínima dificultad, a partir de lo aquí expuesto, la generalización de los resultados obtenidos a un producto tensorial de un número arbitrario de espacios de estados.

CAPÍTULO 2

Descripción cuántica de fenómenos físicos

CONTENIDOS

El objetivo principal del capítulo anterior era introducir el lenguaje matemático asociado a la Física Cuántica. En el capítulo que se inicia ahora se emplearán estos primeros conocimientos, junto con otros que se irán introduciendo de forma paulatina, para describir la evolución del estado y la medida en sistemas cuánticos[1].

2.1. Postulados de la Mecánica Cuántica

Tal y como vimos en el capítulo anterior, el fin último de la Física Cuántica es describir con la mayor precisión y consistencia fenómenos físicos. La caracterización de estos fenómenos pasa por determinar en qué grado ciertas magnitudes físicas son medibles sobre el sistema físico asociado al fenómeno. Se hace necesario, por tanto, establecer con exactitud cómo la realización de medidas puede describirse cuánticamente. A ello, y a la evolución del estado, se consignan los siguientes postulados[2].

[1] En este capítulo y en los posteriores se empleará la notación caligráfica para referirse a las magnitudes físicas y a los espacios de estados. Para diferenciar a los escalares de los operadores, emplearemos en éstos la notación tradicional del circunflejo.

[2] En la exposición de los postulados se consideran únicamente observables con espectro discreto; obviamente, los postulados son extensibles al caso continuo.

2.1.1 Primer postulado Como ya se ha dicho, el estado cuántico de una partícula puede ser descrito mediante una función de onda perteneciente a un espacio $\mathcal{F}$. Posteriormente se asigna a cada una de las funciones pertenecientes a $\mathcal{F}$ un ket perteneciente a un espacio de estados $\mathcal{E}$. Pues bien, el primer postulado no es más que una generalización de la descripción del estado cuántico de una partícula al de un sistema físico cualquiera:

> *Para un instante cualquiera t_0, el estado cuántico de un sistema físico está perfectamente determinado por un ket $|\psi(t_0)\rangle$ perteneciente a $\mathcal{E}$.*

2.1.2 Segundo postulado El segundo postulado relaciona los operadores observables con las magnitudes físicas, haciendo aún más notorias las diferencias existentes entre la descripción clásica y cuántica de un sistema:

> *A toda magnitud física $\mathcal{A}$ medible sobre un sistema físico se le puede asociar un operador $\hat{A}$ que actúa sobre el espacio de estados $\mathcal{E}$. Además, $\hat{A}$ ha de ser un observable.*

2.1.3 Tercer postulado En el postulado precedente simplemente se asigna a cada magnitud física un operador observable, pero sin entrar en absoluto en los posibles resultados de la medida . Será este tercer postulado el que lo haga:

> *Los únicos resultados posibles de la medida de la magnitud $\mathcal{A}$ sobre el sistema físico son los autovalores del observable $\hat{A}$ asociado a dicha magnitud. Así, si expresamos la ecuación de autovalores del observable $\hat{A}$ de la forma:*
>
> $$\hat{A}|\psi_n^m\rangle = a_n|\psi_n^m\rangle, \tag{2.1}$$
>
> *los posibles valores de la medida de $\mathcal{A}$ son los a_n. Nótese que, como todo observable es un operador hermítico, los resultados de la medida han de ser siempre reales.*

2.1.4 Cuarto postulado Este postulado es, realmente, la continuación inmediata del tercero, pues proporciona información sobre los valores concretos que resultarán de realizar medidas en un sistema físico. El postulado anterior dejaba completamente indeterminado este extremo, ya que reconocía que los únicos resultados posibles de la medida eran los autovalores del observable asociado, pero no desvelaba cuál de ellos se obtendría. Esto se aclarará ahora, aunque la información que se proporciona es, salvo en ciertos casos, puramente probabilística:

> *Cuando la magnitud $\mathcal{A}$ se mide sobre un sistema físico caracterizado por el estado $|\psi\rangle$, la probabilidad de obtener como resultado de la medida a_n es*
>
> $$\Pr(a_n) = \sum_{m=1}^{g_n} |\langle\psi_n^m|\psi\rangle|^2, \tag{2.2}$$
>
> *donde recuérdese que g_n es el índice de degeneración del autovalor a_n. Además, una vez que se ha realizado la medida y se ha obtenido como resultado un autovalor a_n cualquiera, el estado del sistema, a partir del instante de medida, pasa a ser*
>
> $$|\psi_n\rangle = \frac{\hat{P}_n|\psi\rangle}{\langle\psi|\hat{P}_n|\psi\rangle^{1/2}}, \tag{2.3}$$
>
> *donde $\hat{P}_n$ es el proyector sobre el espacio subtendido por los autoestados de $\hat{A}$ asociados al autovalor a_n ($\hat{P}_n\hat{P}_m = \delta_{nm}\hat{P}_n$). Se produce, por tanto, un colapso de la función de onda en el estado $|\psi_n\rangle$. Así pues, la medida determina una evolución del sistema que cabría calificar como estrictamente probabilística; no hay más que tener en cuenta el carácter aleatorio del resultado de la medición.*

Una importante consecuencia de este postulado es el hecho de que dos kets que difieran únicamente en un factor de fase representan el mismo estado "físico", pues un factor de fase no afecta a la predicción de una medida. Sin embargo, como se tendrá ocasión de comprobar en la Sección 2.2, las fases de expansión de un vector de estado como combinación lineal de otros sí son significativas.

2.1.4.1 Valor medio y momentos estadísticos Puesto que el resultado de medir la magnitud $\mathcal{A}$ es de naturaleza probabilística, es posible caracterizar estadísticamente el acto de realizar la medida en un cierto estado normalizado $|\psi\rangle$ definiendo, respectivamente, el valor medio y la desviación cuadrática media:

$$\langle \hat{A} \rangle_\psi = \langle \psi | \hat{A} | \psi \rangle, \tag{2.4}$$

$$\Delta \hat{A}_\psi = \langle \psi | (\hat{A} - \langle \hat{A} \rangle_\psi)^2 | \psi \rangle^{1/2}. \tag{2.5}$$

Como resultado más general, es fácil comprobar que el momento estadístico n-ésimo de la medida es:

$$\langle \hat{A}^n \rangle = \langle \psi | \hat{A}^n | \psi \rangle. \tag{2.6}$$

2.1.4.2 Medidas con incertidumbre nula El carácter netamente probabilístico de la medición de una magnitud no impide que, bajo ciertas circunstancias, se pueda predecir el resultado con certeza absoluta. De hecho, cuando el estado $|\psi\rangle$ en que se encuentra el sistema es precisamente uno de los autovectores del observable asociado a la magnitud física a medir, se tiene asegurado que se obtendrá el autovalor asociado al autoestado en cuestión. Obviamente, el valor medio obtenido de la expresión (2.5) coincide con dicho autovalor. Además, todos los momentos estadísticos centrados superiores son cero, como corresponde a la inexistencia de incertidumbre.

2.1.4.3 Magnitudes físicas compatibles y no compatibles Cabe preguntarse si existen estados que sean simultáneamente autoestados de más de un operador observable, es decir, si existen estados en los que se conozcan con certeza absoluta los valores de más de una magnitud física. Nótese que lo que se pretende indagar es una cuestión que no tiene igual en la Física Clásica, donde la precisión en la medida no se ve enturbiada en absoluto por la realización de otras mediciones. Como ya se ha comentado, los únicos límites en la exactitud son en este caso puramente tecnológicos, determinados por los aparatos utilizados.

Sean $\hat{A}$ y $\hat{B}$ dos observables que tienen los mismos autovectores, lo cual no implica que posean necesariamente los mismos autovalores. Según el cuarto postulado, si al medir la magnitud asociada a $\hat{A}$ obtenemos el autovalor a_n, el sistema queda colapsado en el estado $|\psi_n\rangle$. Ahora bien, como este autoestado también lo es de $\hat{B}$, con, por ejemplo, el autovalor asociado b_n, se sabe con toda seguridad que el valor de $\mathcal{B}$ es b_n. La medida de $\mathcal{A}$ no ha afectado negativamente a la precisión con la que se conoce $\mathcal{B}$. Si se hubiese medido primero $\mathcal{B}$, ocurriría exactamente lo mismo con $\mathcal{A}$. La condición necesaria para que esto se produzca es, por tanto, que $\hat{A}$ y $\hat{B}$ conmuten[3]. Se dice entonces que ambas magnitudes son compatibles porque pueden ser medidas simultáneamente sin perturbarse.

Por el contrario, supóngase que $|\psi_n\rangle$ es autovector de $\hat{A}$ pero no de $\hat{B}$. Con toda generalidad, $|\psi_n\rangle$, como cualquier ket perteneciente a $\mathcal{E}$, se puede escribir como combinación lineal de los autovectores del observable $\hat{B}$:

$$|\psi_n\rangle = \sum_i c_i |b_i\rangle. \tag{2.7}$$

[3]Se insta al lector a que, efectivamente, se convenza de ello.

En este caso sólo se conocerán las probabilidades $|c_i|^2$ de obtener como resultado de la medida de $\mathcal{B}$ el valor b_i (autovalor del autoestado $|b_i\rangle$). Es decir, el hecho de medir la magnitud $\mathcal{A}$ perturba el conocimiento certero de $\mathcal{B}$. Se dice, por tanto, que ambas magnitudes son no compatibles. Esto muestra que la Física Cuántica, a diferencia de la Física Clásica, establece límites fundamentales a la precisión con la que se pueden conocer magnitudes físicas de manera simultánea.

2.1.5 Quinto postulado Los cuatro postulados anteriores se ocupan fundamentalmente de la descripción del sistema físico cuando sobre él se realizan medidas. Este último postulado se ocupará de la evolución del estado del sistema inmediatamente después de efectuar la medida:

> *La evolución temporal del estado del sistema, $|\psi\rangle$, viene dada por la ecuación de Schrödinger (1.1). Utilizando la notación de Dirac, ésta se puede escribir de la forma:*
>
> $$i\hbar\frac{d}{dt}|\psi\rangle = \hat{H}|\psi\rangle, \tag{2.8}$$
>
> *en donde $\hat{H}$ es el observable asociado a la energía total del sistema, denominado, por razones históricas, hamiltoniano. Se trata, pues, de una evolución determinista del estado.*

La ecuación (2.8) es de primer orden en el tiempo. Esto implica que, una vez integrada, será preciso el uso de una condición inicial para encontrar el valor de una constante indeterminada. Esta condición es, precisamente, el estado del sistema en algún instante concreto de tiempo t_0, $|\psi(t_0)\rangle$. Este estado se puede conocer, por ejemplo, porque en $t = t_0$ se haya realizado una medida de alguna magnitud, resultando un autovalor cuyo autovector asociado es $|\psi(t_0)\rangle$.

Además se puede comprobar que, dado que la ecuación (2.8) es lineal y $\hat{H}$ es hermítico, la evolución de un estado cuántico se puede describir mediante

$$|\psi(t)\rangle = \hat{U}(t, t_0)|\psi(t_0)\rangle, \tag{2.9}$$

donde $\hat{U}(t, t_0)$ es un operador unitario denominado, por razones obvias, de evolución. En el caso de sistemas conservativos, es decir, aquellos en los que $\hat{H}$ no depende del tiempo, $\hat{U}(t, t_0) = \exp[-i\hat{H}(t - t_0)/\hbar]$.

2.2. Principio de superposición

Sean dos autoestados ortonormales $|\psi_1\rangle$, $|\psi_2\rangle$, asociados respectivamente a dos autovalores sencillos, b_1, b_2, de un observable $\hat{B}$, y sea $|\psi\rangle = \lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle$ el estado del sistema. Según el cuarto postulado, si se mide la magnitud $\mathcal{B}$, la probabilidad de obtener como resultado el valor b_1 es

$$\Pr(b_1) = |\langle\psi_1|\psi\rangle|^2 = |\lambda_1|^2, \tag{2.10}$$

mientras que la probabilidad de conseguir b_2 es

$$\Pr(b_2) = |\langle\psi_2|\psi\rangle|^2 = |\lambda_2|^2. \tag{2.11}$$

Ahora bien, esto no significa en ningún caso que N sistemas como el anterior sean equivalentes a $N|\lambda_1|^2$ sistemas preparados en un estado $|\psi_1\rangle$ y $N|\lambda_2|^2$ caracterizados por $|\psi_2\rangle$. Para comprobarlo, se recurre a un operador observable cualquiera $\hat{A}$ (que contenga en su espectro, por ejemplo, un autovalor sencillo a_n asociado a un autovector $|\phi_n\rangle$), que verifique:

$$\Pr_1(a_n) = |\langle\phi_n|\psi_1\rangle|^2, \tag{2.12}$$

si el sistema estuviese descrito por el ket $|\psi_1\rangle$, y

$$\Pr_2(a_n) = |\langle\phi_n|\psi_2\rangle|^2, \tag{2.13}$$

si estuviese caracterizado por $|\psi_2\rangle$. De este modo, si la interpretación de mezcla estadística fuese correcta, y el estado del sistema fuese $|\psi\rangle$, la probabilidad de obtener a_n al medir la magnitud $\mathcal{A}$ sería

$$\Pr(a_n) = |\lambda_1|^2\Pr_1(a_n) + |\lambda_2|^2\Pr_2(a_n). \tag{2.14}$$

Si recurrimos de nuevo a lo estipulado en el cuarto postulado, es inmediato comprobar que este resultado no concuerda en absoluto con $|\langle\phi_n|\psi\rangle|^2$. Operando de la manera adecuada se obtiene

$$\Pr(a_n) = |\langle\phi_n|\psi\rangle|^2 = |\lambda_1|^2\Pr_1(a_n) + |\lambda_2|^2\Pr_2(a_n) + 2\Re\{\lambda_1\lambda_2^*\langle\phi_n|\psi_1\rangle\langle\phi_n|\psi_2\rangle^*\}. \tag{2.15}$$

Se observa, por tanto, que no sólo los módulos de λ_1 y λ_2 son significativos, sino que aparecen también unos efectos de interferencia en los que la fase relativa de ambas constantes es relevante (a través del término $\lambda_1\lambda_2^*$). A este tipo de superposición se la denomina coherente.

La verdadera importancia de esta interferencia será clara cuando tratemos, al final del capítulo, el fenómeno de la decoherencia. Entonces se verá cómo la interacción entre dos sistemas (los cuales pasan a estar correlados o "entangled") impide acceder de manera aislada a las fases relativas de la superposición que describe el estado de cualquiera de ellos. Su caracterización individual se convierte entonces en una superposición incoherente.

2.3. Principio de incertidumbre de Heisenberg

Tal y como se ha adelantado ya, en la Física Cuántica existe una limitación sobre la certeza con la que se pueden medir dos o más magnitudes no compatibles simultáneamente. Este hecho fue formulado por el científico alemán Werner Heisenberg, quien demostró que la incertidumbre en el conocimiento de la posición de una partícula, multiplicada por la incertidumbre en su cantidad de movimiento (o "momentum"), nunca puede ser más pequeña que una cierta cantidad, la constante de Planck. Además, este límite es una propiedad fundamental e ineludible, que no depende de la forma en que se realicen las medidas, ni del tipo de partícula.

Así, sean dos observables $\hat{A}$ y $\hat{B}$, tales que su conmutador es $[\hat{A}, \hat{B}] = i\hat{C}$, donde $\hat{C}$ es un operador arbitrario. Heisenberg estableció que, para cualquier estado $|\psi\rangle$ del sistema físico sobre el que se realice la medida simultánea de ambos, se verifica

$$(\Delta\hat{A})^2(\Delta\hat{B})^2 \geq \frac{1}{4}|\langle\hat{C}\rangle|^2. \tag{2.16}$$

A fin de demostrar este resultado, considérense los operadores desplazados $\hat{\alpha}$ y $\hat{\beta}$:

$$\hat{\alpha} = \hat{A} - \langle\hat{A}\rangle, \tag{2.17}$$

$$\hat{\beta} = \hat{B} - \langle\hat{B}\rangle. \tag{2.18}$$

Estos operadores verifican las relaciones:

$$(\Delta\hat{\alpha})^2 = (\Delta\hat{A})^2 = \langle\hat{\alpha}\rangle, \tag{2.19}$$

$$(\Delta\hat{\beta})^2 = (\Delta\hat{B})^2 = \langle\hat{\beta}\rangle, \tag{2.20}$$

por lo que, en la demostración, se pueden utilizar estos últimos.

Sea ahora el ket arbitrario $|\phi\rangle = (\hat{\alpha} + i\lambda\hat{\beta})|\psi\rangle$, con λ un escalar real cualquiera. Si se aplica la Propiedad 5 del producto escalar, se tiene que

$$\langle\hat{\alpha}^2\rangle - \lambda\langle\hat{C}\rangle + \lambda^2\langle\hat{\beta}^2\rangle \geq 0. \tag{2.21}$$

Esta desigualdad es cierta para cualquier λ; luego en particular lo será para $\lambda = \langle\hat{C}\rangle/(2\langle\hat{\beta}\rangle^2)$. Pues bien, para obtener (2.16) no hay más que sustituir este valor en (2.21), y deshacer los cambios (2.17) y (2.18).

Una vez comprobada la validez de la ecuación (2.16), cabe ahora preguntarse por aquellos estados que cumplen la igualdad. Esta condición es equivalente a $\langle\phi|\phi\rangle = 0$, lo cual, como es sabido por la Propiedad 5, sólo es posible si $|\phi\rangle = 0$. En definitiva, los estados, llamados de mínima incertidumbre, que dan lugar a la igualdad en (2.16) son aquellos que verifican la condición

$$(\hat{A} - \langle\hat{A}\rangle)|\psi\rangle = -i\lambda(\hat{B} - \langle\hat{B}\rangle)|\psi\rangle. \tag{2.22}$$

Nótese que si los dos observables conmutan, (2.16) se transforma en:

$$(\Delta\hat{A})^2(\Delta\hat{B})^2 \geq 0; \tag{2.23}$$

es decir, para poder realizar la medida simultánea de $\hat{A}$ y $\hat{B}$ con absoluta certeza es necesario que $\hat{A}$ y $\hat{B}$ conmuten.

Este principio posee profundas implicaciones no del todo libres de controversia. En el momento de su formulación marcó el fin del sueño Laplaciano de una teoría determinista de la ciencia; ciertamente, no se pueden predecir los acontecimientos futuros con exactitud si ni siquiera teóricamente se puede determinar el estado presente del universo de forma precisa.

2.4. Imágenes de la Mecánica Cuántica

En todos los capítulos de este libro se entiende la evolución temporal del sistema según la imagen ("picture") de Schrödinger, en la cual el estado del mismo depende del tiempo, mientras que los operadores son invariantes. Sin embargo, existen formulaciones alternativas que dan lugar a diferentes imágenes y que, pese a que no serán utilizadas, sí merece la pena tener una idea general de ellas.

Así, la llamada imagen de Heisenberg posee una estructura dual a la de Schrödinger: los operadores son los que ahora dependen del tiempo, mientras que el estado permanece fijo. La relación entre ambas viene dada por la siguiente ecuación de evolución:

$$|\psi_S(t)\rangle = U(t, t_0)|\psi_H(t_0)\rangle, \tag{2.24}$$

donde los subíndices S y H denotan, respectivamente, la imagen, de Schrödinger o Heisenberg, a la que pertenece el estado. Como $U(t_0, t_0) = 1$, ambos kets coinciden necesariamente en t_0. De esta manera, y teniendo en cuenta que las dos han de predecir idénticos resultados, se tiene que

$$\langle\hat{A}\rangle = \langle\psi_S(t)|\hat{A}_S|\psi_S(t)\rangle = \langle\psi_H(t_0)|\hat{A}_H|\psi_H(t_0)\rangle, \tag{2.25}$$

con $\hat{A}$ un operador cualquiera, y $\hat{A}_S$ y $\hat{A}_H$ la forma de denotarlo en cada una de las imágenes. A partir de lo anterior, es inmediato conseguir una transformación de semejanza entre ambos operadores:

$$\hat{A}_H(t) = U^\dagger(t, t_0)\hat{A}_S U(t, t_0). \tag{2.26}$$

La evolución temporal de los operadores $\hat{A}_H$ viene dada por la ecuación de Heisenberg:

$$\frac{d}{dt}\hat{A}_H(t) = \frac{1}{i\hbar}[\hat{A}_H(t), \hat{H}_H(t)]. \tag{2.27}$$

Esta imagen resulta especialmente útil cuando se cuantifica un sistema descrito clásicamente, ya que su evolución se expresa mediante una serie de ecuaciones operacionales, obtenidas a partir de (2.27), bastante similares a las correspondientes clásicas.

Obviamente, hay muchas más imágenes, tantas como transformaciones de semejanza, del tipo de la ecuación (2.26), se quieran establecer.

2.5. Descripción de un estado cuántico mediante el operador densidad

Hasta ahora se ha supuesto que el estado cuántico de un sistema era perfectamente conocido con anterioridad a la realización de cualquier medida; es lo que se denomina un estado puro. Sin embargo, muy rara vez es así. En la mayoría de las ocasiones la información que se posee acerca del estado del sistema es incompleta, y sólo es posible su descripción mediante una mezcla estadística. La gran ventaja del formalismo del operador densidad es que permite caracterizar la medida y evolución de un sistema físico en ambos casos. Constituye, pues, una descripción alternativa al conocimiento de $|\psi\rangle$.

2.5.1 Caso de estado puro Considérese un sistema físico tal que su estado viene dado por el ket normalizado $|\psi(t)\rangle$, y sea el conjunto $\{|u_i\rangle\}_i$ una base ortonormal de $\mathcal{E}$. El estado $|\psi(t)\rangle$ puede expresarse de la forma:

$$|\psi(t)\rangle = \sum_n c_n(t)|u_n\rangle, \tag{2.28}$$

con $c_n(t) = \langle u_n|\psi(t)\rangle$. Se denomina operador densidad al proyector $\hat{\rho}(t) = |\psi(t)\rangle\langle\psi(t)|$. A partir de esta definición y de (2.28) se obtienen los componentes de su representación matricial en dicha base:

$$\langle u_n|\hat{\rho}(t)|u_m\rangle = c_n(t)c_m^*(t). \tag{2.29}$$

En función de este nuevo operador resulta útil reescribir algunos de los resultados de la Sección 2.1. En efecto, supóngase que se desea calcular el valor medio del observable $\hat{A}$: $\langle\psi(t)|\hat{A}|\psi(t)\rangle$. Sin más que emplear (2.28), se obtiene

$$\langle\hat{A}\rangle = \sum_{n,m} c_n(t)c_m^*(t)\langle u_m|\hat{A}|u_n\rangle. \tag{2.30}$$

Si en esta expresión se usa (2.29), es posible escribir dicho valor medio como

$$\langle\hat{A}\rangle = \sum_n \langle u_n|\hat{\rho}(t)\hat{A}|u_n\rangle = \mathrm{tr}\{\hat{\rho}(t)\hat{A}\}, \tag{2.31}$$

en donde tr denota que ha de calcularse la traza, en cualquier base ortonormal[4], del operador entre llaves.

[4]Recuérdese que la traza es independiente de la representación utilizada, luego la base no ha de ser necesariamente ortonormal; basta con que el conjunto sobre el que se promedia sea base. Sin embargo, la ortonormalidad simplifica el formalismo.

Uno de los resultados más importantes de la primera sección de este capítulo es la primera parte del cuarto postulado, que se puede enunciar como

$$\Pr(a_n) = \langle\psi(t)|\hat{P}_n|\psi(t)\rangle, \tag{2.32}$$

donde $\hat{P}_n$ es el proyector en el subespacio asociado al autovalor a_n. Pues bien, a partir de lo anterior, es sencillo demostrar que

$$\Pr(a_n) = \mathrm{tr}\{\hat{\rho}(t)\hat{P}_n\}. \tag{2.33}$$

El operador densidad no sólo permite la reexpresión de los resultados asociados a la medida de un observable, sino que también, dado que contiene información temporal sobre el estado del sistema, permite caracterizar la dinámica de éste. De hecho, a partir de (2.8) y de la definición de operador densidad, la ecuación de evolución del sistema se puede escribir como

$$\frac{d}{dt}\hat{\rho}(t) = \frac{1}{i\hbar}[\hat{H}(t), \hat{\rho}(t)]. \tag{2.34}$$

2.5.2 Caso de mezcla estadística Como decíamos al principio de esta sección, en el caso más general el estado de sistema no es perfectamente conocido. A lo sumo, se sabe que $|\psi\rangle$ puede ser alguno de los elementos de un conjunto $\{|\psi_i\rangle\}_i$ con probabilidad p_i (evidentemente, $\sum_i p_i = 1$). Se verá a continuación la forma que posee el operador densidad en este caso.

Supóngase inicialmente que el estado del sistema es el k-ésimo de los posibles: $|\psi\rangle = |\psi_k\rangle$. La probabilidad de que al medir con $\hat{A}$ en dicho estado se obtenga a_n es, de acuerdo con (2.33),

$$\Pr_k(a_n) = \langle\psi_k|\hat{P}_n|\psi_k\rangle = \mathrm{tr}\{\hat{\rho}_k\hat{P}_n\}, \tag{2.35}$$

con $\hat{\rho}_k = |\psi_k\rangle\langle\psi_k|$. Al incorporar la incertidumbre en el conocimiento del estado del sistema se tiene que

$$\Pr(a_n) = \sum_k p_k \Pr_k(a_n) = \mathrm{tr}\{\hat{\rho}\hat{P}_n\}, \tag{2.36}$$

en donde ahora el operador densidad adopta la forma

$$\hat{\rho}(t) = \sum_k p_k |\psi_k\rangle\langle\psi_k|. \tag{2.37}$$

Se puede comprobar que todos los resultados obtenidos en la Subsección 2.5.1 son válidos también en este caso, salvo que ahora el operador no es un proyector (no hay más que comprobar que $\hat{\rho}^2(t) \neq \hat{\rho}(t)$). Además, se dice que el estado del sistema $\hat{\rho}(t)$ está formado por una superposición incoherente (ya que las fases relativas de los estados $\{|\psi_i\rangle\}_i$ no son accesibles).

2.5.3 Propiedades generales del operador densidad Acabamos de ver que el operador densidad constituye una interesante alternativa al vector de estado o función de onda en la descripción del estado del sistema. Existe una serie de propiedades generales de este operador cuyo conocimiento resulta imprescindible. Algunas de ellas provienen de su definición formal; otras proceden del significado con el que se ha dotado al formalismo de descripción física basado en la existencia del estado.

1. El operador densidad es hermítico: $\hat{\rho} = \hat{\rho}^\dagger$.

2. La traza del operador densidad es siempre 1.

3. La traza de $\hat{\rho}^2$ es siempre menor o igual que uno. Es uno si y sólo si el sistema se halla en un estado puro.

4. El operador densidad es definido positivo. Es decir:

$$\langle\psi|\hat{\rho}|\psi\rangle \geq 0, \tag{2.38}$$

para cualquier estado $|\psi\rangle$.

5. Todos los autovalores del operador densidad se encuentran entre 0 y 1.

6. El conjunto de todos los posibles operadores densidad es convexo: Dados dos operadores densidad cualesquiera, $\hat{\rho}_1$, $\hat{\rho}_2$, y un escalar real λ ($0 \leq \lambda \leq 1$), el operador

$$\hat{\rho} = \lambda\hat{\rho}_1 + (1-\lambda)\hat{\rho}_2 \tag{2.39}$$

es también un operador densidad. Desde un punto de vista físico, esto supone la posibilidad de cierta mezcla de dos estados para dar lugar a un tercero. Un caso especial lo constituyen los estados puros, los cuales no pueden escribirse según una descomposición del tipo (2.39) o, dicho con otras palabras, no es posible expresarlos como la suma convexa de otros estados, donde por tal suma se entiende

$$\hat{\rho} = \sum_i p_i\hat{\rho}_i, \tag{2.40}$$

con $0 \leq p_i \leq 1$, $\sum_i p_i = 1$, y $\hat{\rho}_i$ operadores densidad. Para comprobar esto último no hay más que considerar un estado puro $\hat{\rho} = |\psi\rangle\langle\psi|$ arbitrario y un ket cualquiera $|\phi\rangle$ ortogonal a $|\psi\rangle$, $\langle\phi|\psi\rangle = 0$. Supóngase ahora que $\hat{\rho}$ puede ser expandido según (2.39),

$$\langle\phi|\hat{\rho}|\phi\rangle = 0 = \lambda\langle\phi|\hat{\rho}_1|\phi\rangle + (1-\lambda)\langle\phi|\hat{\rho}_2|\phi\rangle. \tag{2.41}$$

Dado que la suma de dos términos no negativos se anula, ambos han de ser cero, luego, siempre que λ no sea 0 ó 1 y teniendo en cuenta que $|\phi\rangle$ puede ser cualquier ket ortogonal, se concluye que $\hat{\rho}_1 = \hat{\rho}_2 = \hat{\rho}$. A los elementos de un conjunto convexo que no pueden ser expandidos como combinación lineal de otros elementos del conjunto se les denomina puntos extremos. Obsérvese, por tanto, cómo la convexidad justifica la distinción inicial entre estados puros y mezclas estadísticas. Mientras los primeros pueden ser preparados de una única manera, para los segundos hay infinitas formas de implementarlos como combinación convexa de otros estados. Además, esta gran ambigüedad presente en la mezcla estadística es una característica que contrasta notablemente con las distribuciones de probabilidad clásicas, cuya descomposición en puntos extremos es única.

2.5.4 Descripción de sistemas compuestos

En la Sección 1.7 vimos cómo describir cuánticamente el estado de un sistema compuesto en función del estado de cada uno de sus subsistemas. Dada la equivalencia entre las descripciones facilitadas por el estado y el operador densidad, parece apropiado mostrar aquí cómo este operador permite la descripción de sistemas compuestos.

Considérese, por simplicidad, el sistema cuántico compuesto por los subsistemas A y B. De acuerdo con lo propuesto en la Sección 1.7, el espacio de estados de tal sistema es $\mathcal{E}_{AB} = \mathcal{E}_A \otimes \mathcal{E}_B$, con $\mathcal{E}_A$ y $\mathcal{E}_B$ los espacios de estados de cada uno de los dos subsistemas. Tal y como se acaba de ver, es posible caracterizar de manera alternativa un estado cuántico mediante un operador densidad. De este modo, se puede describir el estado tanto de los subsistemas como del sistema

conjunto mediante los operadores densidad $\hat{\rho}_A$, $\hat{\rho}_B$ y $\hat{\rho}_{AB}$, respectivamente. Se presentarán ahora las relaciones existentes entre estos tres operadores.

Sean los conjuntos $\{|a_i\rangle\}_i$ y $\{|b_j\rangle\}_j$ las bases respectivas de los espacios $\mathcal{E}_A$ y $\mathcal{E}_B$. A partir de ellas se construye una base del espacio total $\mathcal{E}_{AB}$: $\{|a_i, b_j\rangle\}_{i,j}$. Pues bien, las representaciones matriciales de tales operadores densidad son:

$$\rho_{A_{nm}} = \langle a_n|\hat{\rho}_A|a_m\rangle = \sum_j \langle a_n, b_j|\hat{\rho}_{AB}|a_m, b_j\rangle = \langle a_n|\mathrm{tr}_B\{\hat{\rho}_{AB}\}|a_m\rangle, \tag{2.42}$$

$$\rho_{B_{ij}} = \langle b_i|\hat{\rho}_B|b_j\rangle = \sum_n \langle a_n, b_i|\hat{\rho}_{AB}|a_n, b_j\rangle = \langle b_i|\mathrm{tr}_A\{\hat{\rho}_{AB}\}|b_j\rangle; \tag{2.43}$$

en donde, mediante tr_A y tr_B, se designan, respectivamente, las trazas parciales sobre los espacios de estados $\mathcal{E}_A$ y $\mathcal{E}_B$. El concepto de traza parcial es particularmente interesante cuando se pretende realizar algún cálculo sobre sólo uno de los subsistemas constituyentes. Por ejemplo, considérese la obtención del valor medio del observable $\hat{M}_A$ asociado al subsistema A. Puede demostrarse que

$$\langle \hat{M}_A\rangle_A = \mathrm{tr}_A\{\hat{\rho}_A \hat{M}_A\}, \tag{2.44}$$

con $\hat{\rho}_A = \mathrm{tr}_B\{\hat{\rho}_{AB}\}$. Lógicamente, de igual forma se procedería con cualquier observable del subsistema B.

Existen sistemas compuestos por subsistemas denominados incorrelados, en los cuales se tiene que $\hat{\rho}_{AB} = \hat{\rho}_A \otimes \hat{\rho}_B$. Esta incorrelación permite que tanto A como B pueden ser preparados de forma independiente y posteriormente reunidos para constituir el sistema total AB.

2.5.5 La descomposición de Schmidt Se trata de una descomposición sumamente útil, ya que permite expresar cualquier estado de un sistema bipartito (esto es, compuesto por dos únicos subsistemas) como superposición de productos tensoriales de estados pertenecientes a ambos subsistemas.

Sea, como hasta ahora, el sistema compuesto por los subsistemas A y B. Considérese que el estado de B viene dado por el operador densidad $\hat{\rho}_B$, cuyo problema de autovalores se expresa como

$$\hat{\rho}_B|\lambda_k^B\rangle = \lambda_k|\lambda_k^B\rangle; \tag{2.45}$$

lo cual, como se ha visto, permite la representación diagonal del mismo

$$\hat{\rho}_B = \sum_k \lambda_k|\lambda_k^B\rangle\langle\lambda_k^B|. \tag{2.46}$$

Pues bien, se tiene que cualquier estado del sistema conjunto, $|\psi^{AB}\rangle$, puede escribirse de la forma

$$|\psi^{AB}\rangle = \sum_k \sqrt{\lambda_k}|\gamma_k^A\rangle \otimes |\lambda_k^B\rangle, \tag{2.47}$$

en donde la familia $\{|\gamma_k^A\rangle\}_k$ es un conjunto arbitrario de estados ortonormales del espacio $\mathcal{E}_A$. Al número de términos de la expresión (2.47) se le denomina número de Schmidt.

Dado que la demostración de esta posible descomposición es bastante interesante y permite afianzar varios de los conceptos introducidos, se incluye a continuación.

Se parte de una base ortonormal del espacio $\mathcal{E}_A$: $\{|\chi_l^A\rangle\}_l$, la cual, junto con la constituida por los autovectores del operador densidad $\hat{\rho}_B$, determina una base del espacio total $\mathcal{E}_{AB}$, $\{|\chi_l^A\rangle \otimes |\lambda_k^B\rangle\}_{k,l}$. Esto significa que cualquier estado del sistema total puede escribirse de acuerdo con la superposición

$$|\psi^{AB}\rangle = \sum_{k,l} c_{kl}|\chi_l^A\rangle \otimes |\lambda_k^B\rangle, \tag{2.48}$$

en donde, como siempre, los coeficientes de la expansión son la proyección del estado sobre los elementos de la base. Si se define $|\xi_k^A\rangle = \sum_l c_{kl}|\chi_l^A\rangle$, el estado total adquiere la estructura

$$|\psi^{AB}\rangle = \sum_k |\xi_k^A\rangle \otimes |\lambda_k^B\rangle. \tag{2.49}$$

Calculemos ahora, a partir de la expresión (2.49), el operador densidad del subsistema B. Empleando el artificio de la traza parcial, $\hat{\rho}_B = \mathrm{tr}_A\{|\psi^{AB}\rangle\langle\psi^{AB}|\}$, con lo que puede comprobarse sin dificultad que

$$\hat{\rho}_B = \sum_{i,k} \langle\xi_i^A|\xi_k^A\rangle|\lambda_k^B\rangle\langle\lambda_i^B|. \tag{2.50}$$

Si se compara ahora esta expresión con (2.46) se observa que, necesariamente, $\langle\xi_i^A|\xi_k^A\rangle = \lambda_k\delta_{ik}$. Esto sugiere construir el conjunto ortonormal de estados $\{|\gamma_k^A\rangle = 1/\sqrt{\lambda_k}|\xi_k^A\rangle\}_k$ que, al ser sustituido en (2.49) proporciona la descomposición de Schmidt deseada, ecuación (2.47).

2.5.6 El teorema GHJW Se trata de un corolario[5] de la descomposición de Schmidt que refleja el carácter físico de la información. Como veremos inmediatamente, según este teorema la información adquirida al conocer el resultado de medir una magnitud en un sistema B puede llegar a modificar por completo la descripción de A, la cual puede variar desde una superposición incoherente a una coherente. Para constatarlo, se acompaña la demostración.

Sea un sistema A cuya descripción viene dada por el operador densidad $\hat{\rho}_A$, que escribimos de forma general como

$$\hat{\rho}_A = \sum_i p_i|\psi_i^A\rangle\langle\psi_i^A|, \qquad \sum_i p_i = 1; \tag{2.51}$$

donde el conjunto $\{|\psi_i^A\rangle\}_i$ se considera compuesto por kets normalizados, pero no necesariamente ortogonales. Sea ahora una "purificación"[6] $|\Phi_1^{AB}\rangle$ de $\hat{\rho}_A$, es decir, un estado puro del sistema bipartito que satisface

$$|\Phi_1^{AB}\rangle = \sum_i \sqrt{p_i}|\psi_i^A\rangle|\alpha_i^B\rangle, \tag{2.52}$$

y

$$\mathrm{tr}_B(|\Phi_1^{AB}\rangle\langle\Phi_1^{AB}|) = \hat{\rho}_A, \tag{2.53}$$

donde el conjunto de kets $\{|\alpha_i^B\rangle\}_i$ en general se elige ortonormal.

Considérese ahora que escribimos $\hat{\rho}_A$ alternativamente como combinación convexa de otros estados distintos

$$\hat{\rho}_A = \sum_\mu q_\mu|\phi_\mu^A\rangle\langle\phi_\mu^A|, \qquad \sum_\mu q_\mu = 1; \tag{2.54}$$

con $\{|\phi_\mu^A\rangle\}_\mu$ un conjunto normalizado. Al igual que antes, se construye una purificación $|\Phi_2^{AB}\rangle$ tal que

$$|\Phi_2^{AB}\rangle = \sum_\mu \sqrt{q_\mu}|\phi_\mu^A\rangle|\beta_\mu^B\rangle, \tag{2.55}$$

[5] Su nombre se debe a sus autores: Gisin, Hughston, Jozsa y Wootters.

[6] Por purificación de un operador densidad, $\hat{\rho}$, de un espacio de Hilbert $\mathcal{E}_1$, se entiende cualquier estado puro, $|\Phi\rangle$, de un espacio de Hilbert extendido $\mathcal{E}_1 \otimes \mathcal{E}_2$, que verifica: $\hat{\rho} = \mathrm{tr}_2(|\Phi\rangle\langle\Phi|)$.

$$\mathrm{tr}_B(|\Phi_2^{AB}\rangle\langle\Phi_2^{AB}|) = \hat{\rho}_A. \tag{2.56}$$

Es inmediato comprobar que la relación entre $|\Phi_1^{AB}\rangle$ y $|\Phi_2^{AB}\rangle$ es

$$|\Phi_1^{AB}\rangle = (1_A \otimes U_B)|\Phi_2^{AB}\rangle, \tag{2.57}$$

donde U_B es un operador unitario que actúa en el subsistema B. Si se sustituye $|\gamma_\mu^B\rangle = U_B|\beta_\mu^B\rangle$ en (2.57) se obtiene

$$|\Phi_1^{AB}\rangle = \sum_\mu \sqrt{q_\mu}|\phi_\mu^A\rangle|\gamma_\mu^B\rangle. \tag{2.58}$$

Se comprueba así que siempre existe una determinada "purificación" $|\Phi_1^{AB}\rangle$, a partir de la cual se puede preparar $\hat{\rho}_A$ como la combinación de estados (2.51) ó (2.54). Para ello no hay más que proyectar el sistema B sobre el espacio subtendido por los conjuntos $\{|\alpha_i^B\rangle\}_i$ ó $\{|\gamma_\mu^B\rangle\}_\mu$, respectivamente. Obviamente, si se conoce el resultado de la medida, el sistema A pasa a ser descrito por medio de un estado puro, en vez de una mezcla estadística. La demostración se ha realizado únicamente para dos de las posibles combinaciones convexas que pueden preparar una mezcla estadística dada. La generalización del resultado a partir de lo aquí expuesto no ofrece dificultad.

Para confirmar que el teorema GHJW es un corolario de la descomposición de Schmidt, tan solo hay que tener en cuenta que dicha descomposición permite expresar los estados $|\Phi_1^{AB}\rangle$ y $|\Phi_2^{AB}\rangle$ en función de los autovalores, $\{\lambda_k^A\}_k$, y autovectores, $\{|k^A\rangle\}_k$, de $\hat{\rho}_A$:

$$|\Phi_1^{AB}\rangle = \sum_k \sqrt{\lambda_k^A}|k^A\rangle|k_1'^B\rangle, \tag{2.59}$$

$$|\Phi_2^{AB}\rangle = \sum_k \sqrt{\lambda_k^A}|k^A\rangle|k_2'^B\rangle, \tag{2.60}$$

donde $\{|k_1'^B\rangle\}_k$ y $\{|k_2'^B\rangle\}_k$ son dos bases ortonormales del subsistema B. Además, siempre existe una transformación unitaria U_B tal que

$$|k_1'^B\rangle = U_B|k_2'^B\rangle, \tag{2.61}$$

lo cual conduce inmediatamente a (2.57).

2.5.7 Entanglement El "entanglement" o enmarañamiento constituye probablemente el elemento clave que determina la gran diferencia entre las teorías cuántica y clásica de la información; de hecho, como se verá posteriormente, en él están basadas en mayor o menor grado muchas de las más relevantes aplicaciones del procesado cuántico de la información.

Sea el sistema cuántico compuesto por dos subsistemas idénticos A y B, cuyos espacios de estados respectivos, $\mathcal{E}_A$ y $\mathcal{E}_B$, se consideran bidimensionales. Si se supone que las bases de $\mathcal{E}_A$ y $\mathcal{E}_B$ están formadas por los kets $|0\rangle$ y $|1\rangle$, el espacio de estados del sistema compuesto, $\mathcal{E}_{AB} = \mathcal{E}_A \otimes \mathcal{E}_B$, tendrá como base natural el conjunto $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. Sin embargo, no todos los estados pertenecientes a $\mathcal{E}_{AB}$ tienen una interpretación física clara; así, mientras que, por ejemplo, en $|01\rangle$ es evidente que A se encuentra preparado en $|0\rangle$ y B en $|1\rangle$, otras situaciones, como la superposición $1/\sqrt{2}(|00\rangle + |01\rangle)$, son más difíciles de interpretar, pues, aunque en este caso el estado de A es $|0\rangle$, el de B es una superposición de $|0\rangle$ y $|1\rangle$. Finalmente, se pueden considerar ejemplos en los que no es posible asociar un estado cuántico a ninguno de los dos subsistemas de manera independiente: tal circunstancia se da, por ejemplo, en $|\psi\rangle = 1/\sqrt{2}(|00\rangle + |11\rangle)$. Este último tipo de estado, en el que no es posible una factorización de la forma $|\psi_A\rangle \otimes |\psi_B\rangle$, se denomina enmarañado o "entangled".

De hecho, obsérvese que si se realiza la traza parcial sobre el espacio $\mathcal{E}_B$ para conseguir el operador densidad de A, se obtiene un múltiplo del operador identidad en $\mathcal{E}_A$

$$\hat{\rho}_A = \mathrm{tr}_B(|\psi\rangle\langle\psi|) = \frac{1}{2}\hat{1}_A, \tag{2.62}$$

y, de forma similar, $\hat{\rho}_B = \frac{1}{2}\hat{1}_B$. Esto significa que al realizar una medida "local" en cualquiera de los dos subsistemas, no se adquiere ninguna información, es decir, el resultado es totalmente aleatorio (el estado se dice completamente mixto). La consecuencia inmediata de este hecho es que no se puede preparar esta clase de estados mediante transformaciones unitarias del tipo $U_A \otimes U_B$; la única manera de enmarañar dos sistemas es permitir que interactúen entre ellos.

Albert Einstein, Boris Podolsky y Nathan Rosen [35] fueron los primeros en percatarse de las consecuencias de la existencia de estos estados; pero, no pareciéndoles del todo completa esta característica de "no localidad" de la Mecánica Cuántica, propusieron una teoría local de variables ocultas. Tres décadas más tarde, John Bell [11, 12, 13, 44] consiguió probar que, mediante esquemas que supusiesen un comportamiento "local" de ambos subsistemas, es imposible imitar las correlaciones existentes entre los resultados obtenidos en la medida de dos subsistemas "entangled". Los experimentos de Alain Aspect y su grupo [4, 6, 5] confirmaron las predicciones de la Teoría Cuántica y demostraron, por tanto, que el comportamiento de la naturaleza es "no local". Como puede apreciarse, este concepto no tiene igual en la Teoría Clásica, y será imprescindible para entender los capítulos restantes.

2.6. Medidas generalizadas sobre un sistema cuántico

Una vez analizada la utilidad del operador densidad en la descripción de un sistema físico, estudiaremos en esta sección y en la siguiente, utilizando también este operador, la evolución del estado y la medida de magnitudes en dicho sistema. Son cuestiones que ya han sido tratadas de forma breve anteriormente, pero que, debido a su importancia para la completa comprensión de la caracterización cuántica de un sistema, se presentarán ahora con mayor profundidad.

Cuando introdujimos los postulados de la Mecánica Cuántica, aunque no lo indicamos de forma explícita, únicamente consideramos medidas ortogonales o de von Neumann. En este caso, si se mide una magnitud $\mathcal{A}$ en un sistema cuyo estado viene dado por $\hat{\rho}$, los únicos resultados posibles son los autovalores (que se considerarán, por simplicidad, sencillos) del observable $\hat{A}$ asociado. De esta manera, y de acuerdo con el cuarto postulado, el valor a_n se obtendría con probabilidad

$$\Pr(a_n) = \mathrm{tr}\{\hat{P}_n\hat{\rho}\}, \tag{2.63}$$

con lo que, a continuación, el sistema queda descrito por el operador densidad

$$\hat{\rho}' = \frac{\hat{P}_n\hat{\rho}\hat{P}_n}{\mathrm{tr}\{\hat{P}_n\hat{\rho}\}}. \tag{2.64}$$

Obviamente, en caso de no conocer el resultado, en vez de tener un estado puro después de la medición, se dispondría de una mezcla estadística

$$\hat{\rho}' = \sum_n \hat{P}_n\hat{\rho}\hat{P}_n. \tag{2.65}$$

Además, y de acuerdo con la ecuación (1.69), $\hat{A}$ puede expresarse como

$$\hat{A} = \sum_n a_n\hat{P}_n, \tag{2.66}$$

con

$$\hat{P}_n = \hat{P}_n^\dagger, \qquad \hat{P}_n\hat{P}_m = \delta_{nm}\hat{P}_n, \qquad \sum_a \hat{P}_n = \hat{1}. \tag{2.67}$$

Hasta el momento nos hemos limitado a resumir lo ya introducido anteriormente en relación con la medida, pero empleando, en la descripción del estado, el operador densidad. Cabe preguntarse ahora: ¿se pueden realizar medidas sobre un sistema que no sean ortogonales? Aunque tradicionalmente en los textos introductorios de Mecánica Cuántica la medida se introduce asociada a un observable, cuyos autovalores son los resultados posibles de tal medida, es de hecho posible introducir un formalismo de medida mucho más general, del que la medida ortogonal, proyectiva o de von Neumann no sería más que un caso particular.

En efecto, considérese la familia de operadores $\{\hat{M}_m\}_m$ definidos sobre el espacio de estados del sistema. Cada uno de estos operadores lo consideraremos asociado a un posible resultado de la realización de la medida sobre el sistema. Así, para un sistema cuyo estado viene dado por el operador densidad $\hat{\rho}$, en el proceso de medida caracterizado mediante los operadores $\{\hat{M}_m\}_m$ obtendremos el resultado m-ésimo (el asociado a la acción del operador $\hat{M}_m$) con probabilidad

$$p_m = \mathrm{tr}\{\hat{M}_m^\dagger \hat{M}_m \hat{\rho}\}, \tag{2.68}$$

tras lo cual el estado del sistema evolucionará a

$$\hat{\rho}' = \frac{\hat{M}_m \hat{\rho} \hat{M}_m^\dagger}{p_m}. \tag{2.69}$$

Como $\sum_m p_m = 1$, necesariamente $\sum_m \hat{M}_m^\dagger \hat{M}_m = 1$. Esta ecuación suele denominarse de completitud o de resolución de la identidad.

Esta aproximación más general a la medida permite introducir el importante concepto de POVM[7]. A partir de la familia de operadores $\{\hat{M}_m\}_m$ es posible definir otra, dada por los $\hat{F}_m = \hat{M}_m^\dagger \hat{M}_m$. Estos operadores así definidos son, por construcción, hermíticos y positivos, y verifican $\sum_m \hat{E}_m = \hat{1}$. Podemos escribir p_m en función de ellos: $p_m = \mathrm{tr}\{\hat{F}_m \hat{\rho}\}$.

Aunque nosotros aquí hemos obtenido la forma de la POVM a partir de los operadores de medida generalizados, es posible construir POVMs de forma totalmente independiente, sin más que imponer a la familia de operadores su positividad y la resolución de la unidad.

Finalmente, una consideración importante: En la medida proyectiva o de von Neumann el número de resultados posibles es, como máximo, la dimensión del espacio de estados del sistema; sin embargo, esta restricción no existe en el caso de la medida generalizada o en la POVM, al no hallarse vinculados los resultados de la medida con el espectro de ningún observable.

2.7. Superoperadores

El quinto postulado de la Mecánica Cuántica permite determinar la evolución del estado del sistema inmediatamente después de efectuar una medida. En esta sección se recupera esta cuestión, con especial hincapié sobre las diferencias entre los sistemas cuánticos cerrados y abiertos. Recuérdese que estos últimos son aquellos cuya dinámica está influida por su posible interacción con otros sistemas.

[7] Las siglas provienen de la denominación anglosajona: "Positive Operator-Valued Measure".

2.7.1 Evolución en sistemas cuánticos cerrados La evolución en este tipo de sistemas es la reflejada de manera general en la ecuación (2.9), esto es, unitaria. Este hecho puede también ser descrito mediante el operador densidad; así, si el sistema se halla caracterizado por un operador $\hat{\rho}(t)$, es inmediato comprobar que su dinámica puede expresarse de la forma

$$\hat{\rho}(t) = \hat{U}(t, t_0)\hat{\rho}(t_0)\hat{U}^\dagger(t, t_0), \tag{2.70}$$

donde $\hat{\rho}(t_0)$ identifica un estado que se considera inicial (por ejemplo debido al colapso tras la realización de una medida). En el caso particular de que el sistema estuviese descrito por una mezcla estadística, se podría elegir una base en la que $\hat{\rho}(t_0)$ fuese diagonal, de manera que

$$\hat{\rho}(t) = \sum_k p_k \hat{U}(t, t_0)|\psi_k(t_0)\rangle\langle\psi_k(t_0)|\hat{U}^\dagger(t, t_0), \tag{2.71}$$

lo cual concuerda perfectamente con lo esperado. De hecho, se observa fácilmente que cada estado de la mezcla evoluciona según el operador $\hat{U}(t, t_0)$ de manera independiente de los demás, es decir, si en $t = t_0$ el sistema se encuentra en $|\psi_k(t_0)\rangle$ con probabilidad p_k, en t estará en $|\psi_k(t)\rangle = \hat{U}(t, t_0)|\psi_k(t_0)\rangle$ con la misma probabilidad.

Las ecuaciones (2.70) y (2.71) constituyen dos casos particulares de una correspondencia entre operadores densidad denominada superoperador, y cuya acción suele especificarse de forma compacta mediante la notación $\hat{\rho}(t') = \$[\hat{\rho}(t)]$, con $t' > t$.

2.7.2 Evolución en sistemas cuánticos abiertos Hemos visto que, partiendo de un estado puro, una medida ortogonal en un sistema compuesto puede dar lugar a que el estado de uno de los subsistemas pase a ser una mezcla estadística. Además, analizado de manera aislada, puede interpretarse que sobre tal subsistema se ha producido una POVM. De forma similar, en esta sección estudiaremos la dinámica que presenta una parte de dicho sistema cuando éste evoluciona temporalmente de forma unitaria.

Supóngase que, en el instante t_0, los subsistemas A y B se hallan incorrelados, de manera que A puede ser descrito mediante un operador densidad genérico (correspondiente a un estado puro o una mezcla estadística) $\hat{\rho}_A(t_0) = \hat{\rho}_A$, y B utilizando un operador densidad puro $|\psi_B(t_0)\rangle\langle\psi_B(t_0)| = |\psi_B\rangle\langle\psi_B|$. El sistema compuesto estará caracterizado, por tanto, por $\hat{\rho}_{AB} = \hat{\rho}_A \otimes |\psi_B\rangle\langle\psi_B|$. Si la evolución de dicho sistema puede expresarse según la ecuación (2.70), es decir, usando un operador unitario $\hat{U}_{AB}$, entonces el estado en un instante posterior cualquiera es

$$\hat{\rho}'_{AB} = \hat{U}_{AB}(\hat{\rho}_A \otimes |\psi_B\rangle\langle\psi_B|)\hat{U}^\dagger_{AB}. \tag{2.72}$$

A partir de esta ecuación podemos usar el concepto de traza parcial para obtener el estado en t' de, por ejemplo, el subsistema A. Así, empleando la base de B $\{|\mu_B\rangle\}_\mu$, obtenemos

$$\hat{\rho}'_A = \sum_\mu \hat{M}_\mu \hat{\rho}_A \hat{M}^\dagger_\mu, \tag{2.73}$$

donde los miembros de la familia de operadores $\{\hat{M}_\mu\}_\mu$ sobre el espacio $\mathcal{E}_A$ son de la forma

$$\hat{M}_\mu = \langle\mu_B|\hat{U}_{AB}|\psi_B\rangle, \tag{2.74}$$

y verifican

$$\sum_\mu \hat{M}^\dagger_\mu \hat{M}_\mu = \hat{1}_A. \tag{2.75}$$

La expresión (2.73) define una correspondencia entre dos operadores densidad en dos instantes diferentes; se trata, por tanto, de un superoperador. Además, a la suma expresada en dicha ecuación

se la denomina representación de Kraus (los operadores $\hat{M}_\mu$ se denominan de Kraus), y es la forma más general de evolución de un sistema abierto, como es el caso de A (pues interactúa con B), cuando se conoce la evolución unitaria del sistema cerrado conjunto AB. Dado que la representación de Kraus depende de la base elegida para la realización de la traza parcial, dicha representación, lógicamente, no es única. Puede demostrarse que dos descomposiciones de Kraus, $\{\hat{E}_i\}_{i=1}^n$, $\{\hat{F}_j\}_{j=1}^m$, corresponden al mismo superoperador si y sólo si existe una matriz unitaria u_{ij} tal que $\hat{E}_i = \sum_j u_{ij}\hat{F}_j$[8]. Obviamente, también se puede comprobar la implicación inversa, es decir, que siempre es posible que la evolución de A, según un cierto superoperador \$, se realice a través de una transformación unitaria en un sistema compuesto AB:

$$\hat{U}_{AB} : |\psi_A\rangle \otimes |\psi_B\rangle \longrightarrow \sum_\mu \hat{M}_\mu |\psi_A\rangle \otimes |\mu_B\rangle = \sum_\mu |\phi_A\rangle \otimes |\mu_B\rangle, \tag{2.76}$$

donde $\hat{\rho}_A = |\psi_A\rangle\langle\psi_A|$, y seguimos suponiendo que, en el instante inicial, A y B se encuentran incorrelados.

2.7.3 Propiedades de los superoperadores

El formalismo de los superoperadores, como se verá al final del presente capítulo, es sumamente importante en el estudio de la decoherencia (o conversión de estados puros en mezclas estadísticas). Tal y como se ha visto, la evolución unitaria en sistemas cerrados constituye un caso especial en el que la representación de Kraus posee tan solo de un término. Cuando esto no es así —en un sistema abierto—, un estado puro en un subsistema A evoluciona necesariamente a una superposición incoherente. Este hecho justifica que se dedique atención a las propiedades[9] de este tipo de operadores, en cuya exposición se utiliza la notación $\$: \hat{\rho} \longrightarrow \hat{\rho}'$.

1. \$ preserva la hermiticidad, esto es, $\hat{\rho}'$ es hermítico si $\hat{\rho}$ lo es.
2. \$ preserva la traza: $\mathrm{tr}\hat{\rho}' = 1$ si $\mathrm{tr}\hat{\rho} = 1$.
3. \$ es definido positivo, luego si $\hat{\rho}$ lo es, $\hat{\rho}'$ también lo será.
4. \$ es completamente positivo, es decir, dado un subespacio $\mathcal{E}_A$ y una extensión cualquiera del mismo, $\mathcal{E}_A \otimes \mathcal{E}_B$, el operador $\$_A \otimes \hat{1}_B$ es definido positivo. En realidad, viene a ser simplemente una versión más exigente de la propiedad anterior.
5. \$ es lineal[10].
6. Un superoperador \$ es invertible, o lo que es lo mismo, existe su inverso, $\$^{-1}$, si determina una evolución unitaria del operador densidad.

2.8. Decoherencia

Nos hallamos ya en disposición de estudiar un importante fenónemo: la decoherencia. Tal y como avanzamos en la Sección 2.2, cuando un sistema se encuentra caracterizado por un estado puro, por ejemplo mediante la superposición coherente $|\psi\rangle = \lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle$, es sumamente significativa la fase relativa entre λ_1 y λ_2. Mediante unos sencillos ejemplos, y gracias al formalismo del superoperador, vamos a mostrar continuación la conversión, debido a la interacción con otros sistemas, de $|\psi\rangle$ en una mezcla estadística (o superposición incoherente), en la cual ya no es posible

[8] Si $m \neq n$, se complementa la descomposición de Kraus de menor número de operadores con operadores cero hasta lograr $m = n$.

[9] En realidad se trata de las condiciones necesarias que debe cumplir \$ para que permita una representación de Kraus [53].

[10] La posibilidad de una evolución no lineal del estado de un sistema todavía es una cuestión abierta. En cualquier caso, y por simplicidad, esta opción no se tiene en cuenta.

acceder a dichas fases. Este hecho, como tendremos ocasión de comprobar más adelante, cuando se estudie la Teoría de la Información Cuántica, supone la completa pérdida de la información almacenada en el sistema.

Así pues, de forma ilustrativa, abordaremos en las tres subsecciones siguientes el estudio de tres canales cuánticos[11]: el de amortigüamiento de fase, el de amortigüamiento de amplitud y el de depolarización. Se supondrá siempre, por simplicidad, que el espacio de estados del sistema A, $\mathcal{E}_A$, es bidimensional[12], con $\{|0\rangle, |1\rangle\}$ una base ortonormal de dicho espacio. Además, y dado que todo superoperador siempre puede realizarse mediante una transformación unitaria en un espacio extendido, el efecto de cada canal se modelará mediante un operador unitario $\hat{U}_{AE}$, donde el subíndice "E" representa el entorno.

2.8.1 Canal de amortigüamiento de fase

Este canal se emplea en el modelado de gran número de situaciones físicas. La transformación unitaria $\hat{U}_{AE}$ que caracteriza este amortigüamiento es:

$$\hat{U}_{AE} : |0_A\rangle|0_E\rangle \longrightarrow \sqrt{1-p}|0_A\rangle|0_E\rangle + \sqrt{p}|0_A\rangle|1_E\rangle, \tag{2.77}$$

$$\hat{U}_{AE} : |1_A\rangle|0_E\rangle \longrightarrow \sqrt{1-p}|1_A\rangle|0_E\rangle + \sqrt{p}|1_A\rangle|2_E\rangle, \tag{2.78}$$

donde p denota una probabilidad de cambio del estado del entorno $\mathcal{E}_E$, y $\{|0_E\rangle, |1_E\rangle, |2_E\rangle\}$ una base ortonormal del mismo.

Teniendo en cuenta la ecuación (2.74), es inmediato obtener los operadores de Kraus,

$$\hat{M}_0 = \sqrt{1-p}\hat{1}, \qquad \hat{M}_1 = \sqrt{p}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \hat{M}_2 = \sqrt{p}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{2.79}$$

donde es sencillo comprobar la condición de normalización impuesta por la ecuación (2.75): $\hat{M}_0^2 + \hat{M}_1^2 + \hat{M}_2^2 = \hat{1}$. A partir de lo expuesto, y empleando la expresión (2.73), se llega al superoperador que determina la evolución del sistema A:

$$\$(\hat{\rho}) = \hat{M}_0\hat{\rho}\hat{M}_0 + \hat{M}_1\hat{\rho}\hat{M}_1 + \hat{M}_2\hat{\rho}\hat{M}_2 =$$

$$= (1-p)\hat{\rho} + p\begin{pmatrix} \rho_{00} & 0 \\ 0 & \rho_{11} \end{pmatrix} = \begin{pmatrix} \rho_{00} & (1-p)\rho_{01} \\ (1-p)\rho_{10} & \rho_{11} \end{pmatrix}. \tag{2.80}$$

Esto significa que si p depende del tiempo según, por ejemplo, la ecuación $p = \tau\Delta t$, transcurrido un cierto intervalo $t' = n\Delta t$ la dinámica del sistema estará gobernada por $\$^n$. Operando se obtiene trivialmente que $(1-p)^n = (1-\tau\Delta t)^{\frac{t}{\Delta t}} \longrightarrow e^{-\tau t}$ (cuando $\Delta t \longrightarrow 0$). Es decir, si el estado inicial de A es $a|0\rangle + b|1\rangle$, después de un tiempo τ^{-1} se convertirá en la mezcla estadística $\hat{\rho}' = |a|^2|0\rangle\langle 0| + |b|^2|1\rangle\langle 1|$. Se ha producido, por tanto, la desaparición de la superposición coherente inicial (ya no hay efectos de interferencia).

2.8.2 Canal de amortigüamiento de amplitud

Acabamos de ver cómo un superoperador puede provocar la transición de un estado puro a una mezcla estadística. La pregunta ahora es: ¿Podría realizar lo inverso? La respuesta es afirmativa, y para comprobarlo se puede utilizar un modelo que normalmente es empleado en el estudio de la emisión de un fotón por parte de un átomo de dos niveles (debido a su pérdida de energía desde el nivel excitado). El operador de evolución unitario viene dado, en este caso, por

[11] Por analogía con la teoría de la comunicación, en donde la transmisión, a través de medios físicos, de las señales de datos puede provocar una alteración de éstas, se dará a los superoperadores momentáneamente el nombre de canales cuánticos, en tanto en que también describen la evolución temporal del estado del sistema al interaccionar con su entorno.

[12] Como se verá posteriormente, A suele representar un qubit (bit cuántico de información).

$$\hat{U}_{AE} : |0_A\rangle|0_E\rangle \longrightarrow |0_A\rangle|0_E\rangle, \tag{2.81}$$

$$\hat{U}_{AE} : |1_A\rangle|0_E\rangle \longrightarrow \sqrt{1-p}|1_A\rangle|0_E\rangle + \sqrt{p}|0_A\rangle|1_E\rangle, \tag{2.82}$$

donde $|1_A\rangle$ y $|0_A\rangle$ identifican, respectivamente, un estado excitado y uno que no lo está. El entorno lo constituye un campo electromagnético, por lo que su espacio de estados se considera bidimensional $\{|1_E\rangle, |0_E\rangle\}$ (hay o no fotón). Por otra parte, p describe la probabilidad del átomo de perder la energía en su estado excitado.

Procediendo de forma similar al caso del amortigüamiento de fase, se obtiene que

$$\hat{M}_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}, \qquad \hat{M}_1 = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix}, \tag{2.83}$$

a partir de lo cual es fácil comprobar que

$$\hat{M}_0^\dagger \hat{M}_0 + \hat{M}_1^\dagger \hat{M}_1 = \hat{1}. \tag{2.84}$$

El operador $\hat{M}_1$ induce un salto entre estados, de $|1_A\rangle$ a $|0_A\rangle$, mientras que $\hat{M}_0$ describe su evolución cuando este salto no se produce. El superoperador que describe la dinámica de A se puede expresar como

$$\$(\hat{\rho}) = \hat{M}_0\hat{\rho}\hat{M}_0^\dagger + \hat{M}_1\hat{\rho}\hat{M}_1^\dagger = \begin{pmatrix} \rho_{00} + p\rho_{11} & \sqrt{1-p}\rho_{01} \\ \sqrt{1-p}\rho_{10} & (1-p)\rho_{11} \end{pmatrix}. \tag{2.85}$$

Queda por tanto claro que si se aplican los mismos razonamientos del canal de amortigüamiento de fase, esto es, considerar que $p = \tau\Delta t$, y esperar un tiempo superior a τ^{-1}, el operador densidad que describe el sistema A es

$$\$(\hat{\rho}) \longrightarrow \begin{pmatrix} \rho_{00} + p\rho_{11} & 0 \\ 0 & 0 \end{pmatrix}. \tag{2.86}$$

Esta expresión confirma la posibilidad de pasar de una superposición coherente a una incoherente a través de un superoperador.

Si el estado de A fuese $a|0_A\rangle + b|1_A\rangle$, la acción del canal lo haría evolucionar según

$$(a|0_A\rangle + b|1_A\rangle)|0_E\rangle \longrightarrow (a|0_A\rangle + b\sqrt{1-p}|1_A\rangle)|0_E\rangle + \sqrt{p}|0_A\rangle|1_E\rangle. \tag{2.87}$$

Esto significa que si se monitoriza la presencia o no de fotones, es decir, si se realiza una medida ortogonal en $\mathcal{E}_E$ (proyectando sobre los estados $|1_E\rangle$ y $|0_E\rangle$), el estado de A sería, respectivamente, $|0_A\rangle$ ó $a|0_A\rangle + b\sqrt{1-p}|1_A\rangle$. Este resultado coincide plenamente con lo ya avanzado en la expresión (2.86): cuando $t \to \infty$ ($p \to 1$), A se encuentra en el estado puro $|0_A\rangle$[13].

2.8.3 Canal de depolarización Finalmente, se presenta el canal de depolarización[14]. Este canal se comporta de manera parecida a un canal binario simétrico: con probabilidad $1-p$ el estado de A permanece inalterado, y con probabilidad p se produce un error. A su vez, estos errores pueden ser de tres tipos (todos equiprobables), y se caracterizan mediante las matrices de Pauli:

$$\hat{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \hat{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \hat{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.88}$$

[13] Nótese que sin que se produzca ningún cambio en el estado del entorno se puede llegar a inferir el estado cuántico de A.

[14] Reduce, en un factor $1 - \frac{4}{3}p$, la polarización del spin de una partícula.

En base a lo expuesto, el operador unitario que representa al canal es:

$$\hat{U}_{AE} : |\psi_A\rangle|0_E\rangle \longrightarrow \sqrt{1-p}|\psi_A\rangle|0_E\rangle$$

$$+\sqrt{\frac{p}{3}}[\hat{\sigma}_1|\psi_A\rangle|1_E\rangle + \hat{\sigma}_2|\psi_A\rangle|2_E\rangle + \hat{\sigma}_3|\psi_A\rangle|3_E\rangle], \tag{2.89}$$

donde ahora la dimensión de $\mathcal{E}_E$ es cuatro, con $\{|0_E\rangle, |1_E\rangle, |2_E\rangle, |3_E\rangle\}$ una base ortonormal de este espacio.

Los operadores de Kraus son:

$$\hat{M}_0 = \sqrt{1-p}\hat{1}, \qquad \hat{M}_1 = \sqrt{\frac{p}{3}}\hat{\sigma}_1, \qquad \hat{M}_2 = \sqrt{\frac{p}{3}}\hat{\sigma}_2, \qquad \hat{M}_3 = \sqrt{\frac{p}{3}}\hat{\sigma}_3; \tag{2.90}$$

a partir de los cuales se puede comprobar la condición de normalización

$$\sum_\mu \hat{M}_\mu^\dagger \hat{M}_\mu = [(1-p) + 3\frac{p}{3}]\hat{1} = \hat{1}. \tag{2.91}$$

Utilizando, al igual que en los otros dos ejemplos, la ecuación (2.73), se consigue el superoperador que muestra la evolución de A:

$$\$(\hat{\rho}) = (1-p)\hat{\rho} + \frac{p}{3}(\hat{\sigma}_1\hat{\rho}\hat{\sigma}_1 + \hat{\sigma}_2\hat{\rho}\hat{\sigma}_2 + \hat{\sigma}_3\hat{\rho}\hat{\sigma}_3) =$$

$$= \begin{pmatrix} \frac{2}{3}p\rho_{11} + (1-\frac{2}{3}p)\rho_{00} & (1-\frac{4}{3}p)\rho_{01} \\ (1-\frac{4}{3}p)\rho_{10} & \frac{2}{3}p\rho_{00} + (1-\frac{2}{3}p)\rho_{11} \end{pmatrix}. \tag{2.92}$$

Si se toma el peor valor de p ($p = 3/4$), es inmediato comprobar que cualquier estado puro de A se convierte en la superposición incoherente $\hat{1}/2$, en la cual ya no está presente el efecto de interferencia; es decir, se ha producido la decoherencia.

CAPÍTULO 3

Una teoría cuántica de la información

CONTENIDOS

Quizás la Teoría de la Información, tal y como fue formalizada por Claude Shannon hace ahora unos cincuenta años, sea la teoría, junto con la Mecánica Cuántica, de mayor relevancia de entre las concebidas durante el pasado siglo. De hecho, la revolución tecnológica en las comunicaciones acaecida en este período de tiempo recibe su soporte teórico de esta reciente disciplina. Hasta su concepción, el estudio de las limitaciones fundamentales de los sistemas de comunicaciones se realizaba sin desligar su operación de la realización tecnológica. Gracias a las aportaciones de esta teoría se consigue abstraer el soporte físico, que pasa a ser irrelevante, para centrarse en el estudio de la información como entidad abstracta pero cuantificable.

Sin embargo, el descubrimiento reciente[1] de que, cuando se aplican los métodos de análisis propios de la Mecánica Cuántica al estudio de la información, se obtienen resultados no reconciliables con las ideas de Shannon, ha favorecido el nacimiento de un nuevo paradigma de descripción formal de la información: la Teoría de la Información Cuántica. Algunos de los puntos en los que difiere esta nueva disciplina de su predecesora provienen de la propia naturaleza de la Física Cuántica, sin igual en la Física Newtoniana. A modo de ejemplo se podrían citar: la existencia de magnitudes incompatibles (el orden de la medida influye decisivamente en los resultados obtenidos), los resultados derivados del carácter "no local" de la naturaleza (sistemas "entangled"), o un aspecto que todavía no hemos analizado con total detenimiento y que resulta esencial: la imposibilidad de distinguir perfectamente estados no ortogonales.

En este capítulo pretendemos, además de introducir los principales resultados de la Teoría de la Información Cuántica, mostrar cómo, de manera análoga a la relación existente entre la

[1] Las primeras referencias, en el ámbito de la computación cuántica, son de mediados de los años 80.

Mecánica Cuántica y la Clásica, en la mayoría de las ocasiones la concepción de Shannon no es más que un caso particular de esta nueva teoría cuántica. Así, comprobaremos que, empleando los conocimientos introducidos en los capítulos precedentes, la mayoría de los resultados fundamentales de la Teoría de la Información Clásica representan casos particulares de sus equivalentes en la versión cuántica.

3.1. Unidad de información cuántica. El qubit

De manera similar a la teoría clásica, en la que la unidad de información es el bit (con un valor perfectamente definido: 0 ó 1), en la teoría cuántica se utiliza el qubit[2], el estado genérico asociado a un espacio de Hilbert bidimensional. Así, y en función de una base ortonormal cualquiera, que se denotará mediante $\{|0\rangle, |1\rangle\}$, la forma más general de expresar un qubit es

$$|\psi\rangle = a_0|0\rangle + a_1|1\rangle, \tag{3.1}$$

con a_0 y a_1 dos números complejos resultantes de proyectar el estado sobre los elementos de la base: $a_i = \langle i|\psi\rangle$ $(i = 1, 0)$, y que verifican $|a_0|^2 + |a_1|^2 = 1$. A la vista de la expresión (3.1), resultan evidentes las grandes diferencias entre las descripciones clásica y cuántica de la unidad elemental de información. Para ello no hay más que abstraerse de la entidad física del sistema y asignar un estado lógico "0" a $|0\rangle$ y otro estado lógico "1" a $|1\rangle$. Esto significa que, mientras un bit contiene una información concreta a la que se puede acceder sin perturbación alguna, una medida que proyecte un qubit en la base $\{|0\rangle, |1\rangle\}$ siempre proporciona un resultado probabilístico, y sólo en caso de que $a_0 = 0$ ó $a_1 = 0$ el estado del sistema permanece inalterado. Además, conviene recordar que en estos estados $|\psi\rangle$ dados por la ecuación (3.1), a los que no es posible atribuir un estado lógico concreto, no sólo son importantes los módulos de los coeficientes de expansión, sino también la interferencia presente entre ellos, es decir, su fase relativa. Por tanto, no es correcto interpretar un qubit desde una perspectiva probabilística de la información clásica; no se trata simplemente de un bit en el que no se conoce más que la probabilidad del valor almacenado. Esta mayor complejidad del formalismo es la responsable de la mayor riqueza de la descripción cuántica respecto a la clásica.

3.1.1 Imposibilidad de copiar qubits

Uno de los conceptos que determinan las grandes diferencias con la Teoría de la Información Clásica es la imposibilidad de copiar un estado cuántico desconocido con total fidelidad. Es un corolario evidente del principio de incertidumbre de Heisenberg, ya que si esto no fuera así, siempre sería posible medir magnitudes incompatibles de manera independiente en cada una de las copias.

De forma más cuantitativa tampoco resulta difícil comprobarlo. Así, dados dos espacios de estados bidimensionales, $\mathcal{E}_A$ y $\mathcal{E}_B$, y el espacio conjunto $\mathcal{E}_{AB} = \mathcal{E}_A \otimes \mathcal{E}_B$, se puede demostrar[3] que siempre es posible, a partir de un conjunto de estados ortogonales de $\mathcal{E}_A$, encontrar un operador unitario $\hat{U}_{AB}$ que consiga igualar el estado de $\mathcal{E}_B$ a cualquiera de los kets del conjunto:

$$\hat{U}_{AB} : |0\rangle_A|0\rangle_B \longrightarrow |0\rangle_A|0\rangle_B, \tag{3.2}$$

$$\hat{U}_{AB} : |1\rangle_A|0\rangle_B \longrightarrow |1\rangle_A|1\rangle_B. \tag{3.3}$$

Sin embargo, resulta evidente que este mismo operador no puede ser empleado cuando el estado de $\mathcal{E}_A$ está descrito por un ket que no es ortogonal a los estados para los que fue diseñado:

$$\hat{U}_{AB} : (a|0\rangle_A + b|1\rangle_A)|0\rangle_B \longrightarrow a|0\rangle_A|0\rangle_B + b|1\rangle_A|1\rangle_B \neq$$

[2]La terminología proviene de la denominación anglosajona "quantum bit".

[3]Para ello se utiliza una puerta cuántica XOR. Se verán cuando se trate la computación cuántica más adelante.

$$\neq (a|0\rangle_A + b|1\rangle_A) \otimes (a|0\rangle_B + b|1\rangle_B), \tag{3.4}$$

ya que claramente la operación de clonación falla. Luego, cuando el estado a copiar es desconocido (es, en general, una superposición), no se tiene en absoluto garantizado que la operación se vaya a realizar con éxito.

3.1.2 Indistinguibilidad de qubits no ortogonales Supóngase que en la preparación del estado de un sistema se escoge, de entre un conjunto ortogonal de estados puros[4], $\{|\psi_x\rangle\}_x$, el ket $|\psi_x\rangle$ con probabilidad p_x. De acuerdo con lo visto en los dos capítulos anteriores, para un observador que no conozca la decisión tomada, el sistema se encuentra caracterizado por la siguiente mezcla estadística:

$$\hat{\rho} = \sum_x p_x |\psi_x\rangle\langle\psi_x|. \tag{3.5}$$

Si en un instante dado se desea conocer la elección concreta que se efectuó, de acuerdo con lo visto anteriormente, lo más conveniente es realizar una medida de von Neumann:

$$\hat{Y} = \sum_y a_y \hat{P}_y, \tag{3.6}$$

donde $\hat{P}_y = |\psi_y\rangle\langle\psi_y|$. De esta forma, la obtención, como resultado de la medida, de un valor a_y, determina perfectamente el ket seleccionado, $|\psi_y\rangle$. Pero, ¿qué ocurre si el conjunto sobre el que se escoge el estado no es ortogonal? En este caso, no es posible determinar con total seguridad el estado puro que se eligió. Ello es debido a que, aunque se utilice el operador de la ecuación (3.6)[5], siempre habrá, para un cierto valor a_y de la medida, varios estados que con probabilidad distinta de cero podrían producir dicho resultado (aquellos estados $|\psi_x\rangle$ para los que $\langle\psi_y|\psi_x\rangle \neq 0$).

3.1.3 Sistemas multiqubit Sea ahora un sistema cuántico compuesto por dos qubits. Su espacio de estados, según hemos visto, será el producto tensorial de los espacios de estados de los respectivos qubits, $\mathcal{E}_{\text{Total}} = \mathcal{E}_1 \otimes \mathcal{E}_2$, por lo que su base natural será $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. En esta base, un estado genérico $|\psi\rangle$ del sistema se escribe de acuerdo con la superposición coherente

$$|\psi\rangle = a_{00}|00\rangle + a_{01}|01\rangle + a_{10}|10\rangle + a_{11}|11\rangle, \tag{3.7}$$

donde a_{ij} representa un número complejo. Obsérvese que, en este caso, para especificar completamente el estado del sistema son necesarios cuatro números complejos. Como este caso bi-qubit no es demasiado significativo, considérese la situación en la que se dispone de N qubits. Ahora, procediendo por analogía con el caso anterior, la base del espacio de estados del sistema consta de 2^N elementos, por lo que la especificación de un estado, que en el caso clásico requeriría conocer N valores reales, precisa en la teoría cuántica de 2^N valores complejos. Si se denota a cada uno de los elementos de la base mediante un ket $|x\rangle$, con $x = 0, \ldots, 2^N - 1$, se puede escribir de forma compacta el estado del sistema como

$$|\psi\rangle = \sum_{x=0}^{2^N-1} a_x |x\rangle, \tag{3.8}$$

con $a_x = \langle x|\psi\rangle$. Al igual que ocurría con un solo qubit, se advierte también aquí una mayor complejidad de la descripción cuántica con respecto a la clásica. Más adelante se tendrá ocasión de comprobar que una de las grandes ventajas potenciales de la Teoría de la Información Cuántica — la posibilidad de procesar información paralelamente de forma masiva (fundamento de la llamada

[4] La generalización al caso de mezclas estadísticas no supondría ninguna dificultad.

[5] Este caso se correspondería con la realización de una POVM.

computación cuántica)— reside, en parte, en este crecimiento exponencial de la complejidad del formalismo de descripción con el número de qubits.

3.2. Transmisión de la información cuántica

El modelo tradicional de un sistema de comunicación lo componen una fuente, un canal y un receptor. La fuente[6] de información genera símbolos de un determinado conjunto finito (alfabeto fuente) a intervalos regulares y de manera aleatoria e independiente de los símbolos anteriores. Los canales discretos transmiten símbolos de un determinado conjunto (alfabeto de entrada), y generan a su salida otros símbolos pertenecientes a otro conjunto (alfabeto de salida). Como el alfabeto fuente y el de entrada al canal no tienen por qué coincidir, en ocasiones se hace necesaria una codificación que garantice la máxima eficiencia en la transmisión. Básicamente, la codificación consiste en asignar a cada uno de los símbolos de la fuente una determinada palabra código (formada por elementos del alfabeto de entrada). Esto debe realizarse buscando una longitud media del código mínima, así como una decodificación unívoca en el receptor. Además, puesto que el canal normalmente no es ideal, la información recibida difiere de la enviada, y esta discrepancia se traduce en la existencia de una probabilidad de error en el funcionamiento del receptor, cuya misión no es otra que recuperar, con la máxima fidelidad posible, la información original.

En la Teoría de la Información Cuántica estos conceptos se mantienen, pero con ciertas matizaciones. En primer lugar, al símbolo generado por la fuente se le asocia un estado cuántico (o, equivalentemente, un operador densidad) definido sobre un cierto espacio de estados n-dimensional. Además, como normalmente el alfabeto de entrada del canal suele pertenecer a un espacio de Hilbert bidimensional (qubits), también suele ser necesaria una codificación de fuente que asigne a cada estado-símbolo una representación en qubits. En segundo lugar, cualquier proceso relacionado con la transmisión de información que altere el estado asociado a la misma se caracteriza mediante la acción de un superoperador[7]. Pero, en este caso, lo habitual es incluir el comportamiento ruidoso del canal, debido a la interacción del sistema con su entorno (más o menos pasivo), en el alfabeto fuente. Así, en los canales libres de errores se asocia con cada símbolo un estado puro, mientras que en los ruidosos se suele emplear una mezcla estadística. Se trata, por tanto, de sistemas cuánticos cerrados, compuestos por los subsistemas asociados a la información (sistema abierto) y al entorno ruidoso. Por último, la calidad de la comunicación se mide en el receptor mediante la función fidelidad, que introducimos a continuación.

3.3. Fidelidad

Supóngase que se desea enviar a través de un canal cuántico un estado puro descrito mediante el operador densidad $\hat{\rho} = |\psi\rangle\langle\psi|$. El comportamiento no ideal del canal, caracterizado por un superoperador \$, puede provocar una alteración en el símbolo asociado a $\hat{\rho}$:

$$\hat{\rho}' = \$(\hat{\rho}). \tag{3.9}$$

Es por ello necesario realizar una cuantificación de estas alteraciones indeseadas en el estado del sistema; es decir, calcular la probabilidad de recibir, a todos los efectos, el mismo símbolo que se ha enviado. Esto se consigue mediante la función fidelidad, la cual, en base al ejemplo expuesto, se define como

$$F(\hat{\rho}, \hat{\rho}') = \langle\psi|\hat{\rho}'|\psi\rangle. \tag{3.10}$$

Obviamente, cuanto más próxima a 1 se encuentre $F(\hat{\rho}, \hat{\rho}')$, mejor habrá sido la transmisión.

[6] Se considerará siempre que se trata de una fuente discreta y sin memoria.

[7] Recuérdense los ejemplos de canales cuánticos introducidos en el capítulo anterior.

Tal y como acaba de ser introducida la función fidelidad, su utilidad parecería limitada al envío de estados puros. Para el caso de mezclas estadísticas, el formalismo de descripción se complica, por lo que, por simplicidad, únicamente se incluirá la nueva definición[8]:

$$F(\hat{\rho},\hat{\rho}') = \mathrm{tr}^2[(\sqrt{\hat{\rho}}\hat{\rho}'\sqrt{\hat{\rho}})^{1/2}], \tag{3.11}$$

donde se presupone que tanto $\hat{\rho}$ como $\hat{\rho}'$ se corresponden con mezclas estadísticas. La ecuación anterior también puede reescribirse de la forma[9]

$$F(\hat{\rho},\hat{\rho}') = \max|\langle\phi|\phi'\rangle|^2, \tag{3.12}$$

con $|\phi\rangle$ y $|\phi'\rangle$ sendas purificaciones de $\hat{\rho}$ y $\hat{\rho}'$.

Debido a la importancia de la función fidelidad, se enuncian a continuación, sin demostración [45], algunas de sus propiedades:

1. $0 \leq F(\hat{\rho},\hat{\rho}') \leq 1$, donde la igualdad, $F(\hat{\rho},\hat{\rho}') = 1$, se verifica si y sólo si $\hat{\rho} = \hat{\rho}'$.
2. $F(\hat{\rho},\hat{\rho}') = F(\hat{\rho}',\hat{\rho})$.
3. Dados dos números reales positivos, p_1 y p_2, tales que $p_1 + p_2 = 1$, entonces

$$F(\hat{\rho}, p_1\hat{\rho}_1 + p_2\hat{\rho}_2) \geq p_1 F(\hat{\rho},\hat{\rho}_1) + p_2 F(\hat{\rho},\hat{\rho}_2). \tag{3.13}$$

   Además:

$$F(\hat{\rho},\hat{\rho}') \geq \mathrm{tr}(\hat{\rho}\hat{\rho}'). \tag{3.14}$$

4. Si $\hat{\rho}$ es un estado puro ($\hat{\rho} = |\psi\rangle\langle\psi|$), entonces:

$$F(\hat{\rho},\hat{\rho}') = \langle\psi|\hat{\rho}'|\psi\rangle = \mathrm{tr}(\hat{\rho}\hat{\rho}'). \tag{3.15}$$

5. $F(\hat{\rho}_1\otimes\hat{\rho}_2, \hat{\rho}_3\otimes\hat{\rho}_4) = F(\hat{\rho}_1,\hat{\rho}_3)F(\hat{\rho}_2,\hat{\rho}_4)$.
6. Cualquier medida que transforme los estados $\hat{\rho}_1$ y $\hat{\rho}_2$ en, respectivamente, $\hat{\rho}_1'$ y $\hat{\rho}_2'$, verifica

$$F(\hat{\rho}_1',\hat{\rho}_2') \geq F(\hat{\rho}_1,\hat{\rho}_2). \tag{3.16}$$

3.4. Entropía de von Neumann

Supóngase ahora que una fuente con un alfabeto de entrada de n elementos genera, con probabilidad p_x, un símbolo caracterizado por el operador densidad $\hat{\rho}_x$. De este modo, el estado cuántico proporcionado por la fuente se puede describir mediante

$$\hat{\rho} = \sum_x p_x\hat{\rho}_x. \tag{3.17}$$

Pues bien, se define la entropía de von Neumann como la función de $\hat{\rho}$ dada por

$$S(\hat{\rho}) = -\mathrm{tr}(\hat{\rho}\log\hat{\rho}). \tag{3.18}$$

[8] Al lector interesado se le remite a [57, 58, 3, 2]

[9] Al lector interesado se le remite a [45].

Es inmediato comprobar que, si se elige una base ortonormal que proporcione una representación matricial diagonal de $\hat{\rho}$,

$$\hat{\rho} = \sum_a \lambda_a |a\rangle\langle a|, \tag{3.19}$$

la entropía de von Neumann coincide con la de Shannon, con tal que se considere una fuente clásica que genere símbolos a con probabilidad λ_a. La entropía de Shannon, por tanto, no es más que la particularización de $S(\hat{\rho})$ a un alfabeto fuente compuesto por símbolos ortogonales.

Además, también es interesante observar cómo $S(\hat{\rho})$ refleja la distinción entre un estado puro y una mezcla estadística. De hecho, es inmediato constatar[10] que sólo en el primer caso $S(\hat{\rho}) = 0$. Igualmente, no resulta difícil verificar que la evolución unitaria determinada por la ecuación de Schrödinger no modifica el valor de $S(\hat{\rho})$ (evolución denominada reversible), mientras que la decoherencia producida por la influencia del entorno sí provoca un incremento de la entropía de von Neumann (se habla entonces de evolución irreversible). Posteriormente veremos que, mientras un incremento de $S(\hat{\rho})$ está relacionado con la pérdida de información, una reducción de la entropía implica una ganancia de ésta.

A continuación se presentan sin demostración [60, 50] un conjunto de propiedades de $S(\hat{\rho})$ que resultan a menudo muy útiles. Algunas ya han sido avanzadas en su definición.

1. **Rango:** Para cualquier número real, $c \in \mathbb{R}$ y $0 \leq c \leq \infty$, siempre existe un operador densidad $\hat{\rho}$ que verifica $S(\hat{\rho}) = c$.

2. **Valor máximo:** Si un operador densidad $\hat{\rho}$ posee N autovalores distintos de cero, entonces

$$S(\hat{\rho}) \leq \log N. \tag{3.20}$$

   La igualdad sólo se consigue cuando los N autovalores son iguales.

3. **Invariancia:** Una evolución unitaria del estado no modifica su entropía:

$$S(\hat{U}\hat{\rho}\hat{U}^{-1}) = S(\hat{\rho}). \tag{3.21}$$

4. **Concavidad:** Dados $\lambda_1, \lambda_2, \ldots, \lambda_n \geq 0$ y tales que $\lambda_1 + \lambda_2 + \ldots + \lambda_n = 1$, se cumple que:

$$S(\lambda_1\hat{\rho}_1 + \lambda_2\hat{\rho}_2 + \ldots + \lambda_n\hat{\rho}_n) \geq \lambda_1 S(\hat{\rho}_1) + \lambda_2 S(\hat{\rho}_2) + \ldots + \lambda_n S(\hat{\rho}_n). \tag{3.22}$$

   Este resultado es consecuencia de la propiedad de convexidad de los operadores densidad, ya que éstos, cuando no son puros, siempre pueden ser expresados de múltiples maneras como suma convexa de otros estados. Luego, a partir de una determinada mezcla estadística $\hat{\rho}$, es imposible determinar la suma convexa concreta que se utilizó para su realización, esto es, se ha perdido información.

5. **Entropía de una mezcla estadística:** Supóngase un sistema cuántico caracterizado por la mezcla estadística

$$\hat{\rho} = \sum_x p_x |\psi_x\rangle\langle\psi_x|. \tag{3.23}$$

   Se puede demostrar que

$$H(X) \geq S(\hat{\rho}), \tag{3.24}$$

[10] Hay que tener en cuenta que el operador densidad de un estado puro es un proyector, $\hat{\rho}^2 = \hat{\rho}$.

donde X es una variable aleatoria que toma un valor x con probabilidad p_x, y $H(x)$ es la entropía de Shannon. La igualdad se verifica si el conjunto $\{|\psi_x\rangle\}_x$ es ortogonal. Físicamente, este resultado no es más que un reflejo de lo ya mencionado en la Subsección 3.1.2: la imposibilidad de distinguir perfectamente estados no ortogonales.

6. **Entropía del resultado de una medida:** Dado un sistema descrito por el operador densidad $\hat{\rho}$, y un observable $\hat{A}$,

$$\hat{A} = \sum_n a_n |\psi_n\rangle\langle\psi_n|, \tag{3.25}$$

se cumple que

$$H(X) \geq S(\hat{\rho}), \tag{3.26}$$

donde X representa una variable aleatoria que toma un valor a_n con probabilidad $p_n = \langle a_n|\hat{\rho}|a_n\rangle$, y $H(x)$ es la entropía de Shannon. La igualdad sólo es cierta si $[\hat{A}, \hat{\rho}] = 0$. Esta propiedad evidencia que el resultado de medir una magnitud es menos predecible si su observable asociado no conmuta con el estado cuántico del sistema[11].

7. **Subaditividad:** La entropía de von Neumann de un sistema bipartito, AE, es siempre menor o igual que la suma de las entropías de cada subsistema,

$$S(\hat{\rho}_{AE}) \leq S(\hat{\rho}_A) + S(\hat{\rho}_E), \tag{3.27}$$

donde $\hat{\rho}_A = \mathrm{tr}_E\{\hat{\rho}_{AE}\}$ y $\hat{\rho}_E = \mathrm{tr}_A\{\hat{\rho}_{AE}\}$. La igualdad sólo se cumple cuando $\hat{\rho}_{AE} = \hat{\rho}_A \otimes \hat{\rho}_E$. Esta propiedad pone de manifiesto que la entropía del sistema conjunto nunca decrece. Para comprobarlo, supóngase que en un principio A y E están incorrelados,

$$\hat{\rho}_{AE} = \hat{\rho}_A \otimes \hat{\rho}_E, \tag{3.28}$$

por lo que la entropía del sistema bipartito es

$$S(\hat{\rho}_{AE}) = S(\hat{\rho}_A) + S(\hat{\rho}_E). \tag{3.29}$$

Permitamos ahora que A evolucione durante unos instantes. Sabemos que esta evolución puede ser modelada mediante un operador unitario $\hat{U}_{AE}$ actuando en AE,

$$\hat{U}_{AE} : \hat{\rho}_{AE} \longrightarrow \hat{\rho}'_{AE} = \hat{U}_{AE}\hat{\rho}_{AE}\hat{U}_{AE}^{-1}. \tag{3.30}$$

Por último, y teniendo en cuenta las expresiones (3.21) y (3.27), se concluye que

$$S(\hat{\rho}_A) + S(\hat{\rho}_E) \leq S(\hat{\rho}'_A) + S(\hat{\rho}'_E). \tag{3.31}$$

De hecho, se puede demostrar que la interacción entre un sistema A y su entorno E induce correlaciones que provocan un aumento asintótico de la entropía de von Neumann hasta lograr su máximo teórico; es decir, la información contenida en A se traslada a sus correlaciones con el entorno. Esto significa que, aunque teóricamente es posible recuperar la información, en la práctica, no es así.

La propiedad análoga en la teoría clásica sería

$$H(X, Y) \leq H(X) + H(Y). \tag{3.32}$$

[11] Se trata de una cuestión que ya mencionamos al estudiar las magnitudes compatibles e incompatibles.

8. **Subaditividad fuerte:** Dado un estado $\hat{\rho}_{ABC}$ de un sistema tripartito, se cumple que

$$S(\hat{\rho}_{ABC}) + S(\hat{\rho}_B) \leq S(\hat{\rho}_{AB}) + S(\hat{\rho}_{BC}). \tag{3.33}$$

9. **Desigualdad de Araki-Lieb:** Mientras que la entropía de Shannon verifica

$$H(X,Y) \geq H(X), H(Y), \tag{3.34}$$

es decir, hay más información en un sistema compuesto que en cada una de sus partes, su versión cuántica establece que

$$S(\hat{\rho}_{AB}) \geq |S(\hat{\rho}_A) - S(\hat{\rho}_B)|. \tag{3.35}$$

De esta manera, y para un estado puro de un sistema bipartito que cumpla $S(\hat{\rho}_A) = S(\hat{\rho}_B) \neq 0$, se obtiene $S(\hat{\rho}_{AB}) = 0$, lo contrario de lo estipulado en la teoría clásica. Esto no debe suponer ninguna sorpresa; es más, resulta inmediato a partir de lo ya visto. No hay más que recordar cómo en el capítulo anterior dijimos que, debido al carácter no local de la naturaleza, es imposible acceder al conocimiento de un estado "entangled" mediante medidas restringidas a subsistemas, ya que éstas proporcionan resultados totalmente aleatorios.

10. **Discontinuidad:** Dado un operador densidad $\hat{\rho}$ y un número cualquiera $\epsilon > 0$, siempre existe $\hat{\rho}'$ tal que

$$\mathrm{tr}|\hat{\rho} - \hat{\rho}'| \leq \epsilon \qquad \text{y} \qquad S(\hat{\rho}') = \infty. \tag{3.36}$$

Luego, claramente, $S(\hat{\rho})$ es discontinua.

3.5. Codificación de fuente

Dada una fuente de información clásica (discreta y sin memoria), y un canal libre de errores, el teorema de codificación de fuente de Shannon [30] establece la posibilidad de encontrar un código unívocamente decodificable cuya longitud media sea tan próxima a la entropía como se desee, y la fidelidad del esquema sea arbitrariamente alta[12]. La demostración se basa en la codificación conjunta de secuencias de N símbolos consecutivos (N suficientemente grande), de manera que siempre se puede conseguir realizar su transmisión con una fidelidad $F > 1-\varepsilon$ utilizando $N(H+\delta)$ bits (H denota la entropía de la fuente y δ y ε dos números reales cualesquiera estrictamente positivos). Además, la entropía es la longitud media mínima, y representa, en promedio, la cantidad de información clásica almacenada en cada símbolo.

A continuación se presentarán sin demostración [46, 51], y para un alfabeto de entrada al canal formado por estados de un espacio bidimensional (qubits), dos resultados análogos al anterior en la teoría cuántica. Primero se analizará el caso en que la fuente genera estados puros, para posteriormente extender el concepto a las mezclas estadísticas. En ambos casos quedará de manifiesto el papel fundamental que desempeña la entropía de von Neumann.

[12] Pero nunca $F = 1$.

3.5.1 Codificación de estados puros. Teorema de Schumacher

Sea una fuente que genera estados cuánticos puros (pertenecientes a un espacio de Hilbert n-dimensional) de un conjunto no necesariamente ortogonal ($\{|\psi_1\rangle, \ldots, |\psi_x\rangle\}$), con probabilidades $p_1, \ldots, p_x$. El símbolo generado por la fuente se encuentra caracterizado, por tanto, por el siguiente operador densidad:

$$\hat{\rho} = \sum_x p_x |\psi_x\rangle\langle\psi_x|. \tag{3.37}$$

De igual forma, una secuencia compuesta por N símbolos se describe, utilizando el formalismo del producto tensorial, como[13]

$$\hat{\rho}^N = \hat{\rho} \otimes \ldots \otimes \hat{\rho}. \tag{3.38}$$

Pues bien, Schumacher demostró que, dados dos números reales positivos cualesquiera, δ y ε, y secuencias de N símbolos (N suficientemente grande), siempre es posible codificar, en promedio, cada estado emitido por la fuente mediante $S(\hat{\rho})+\delta$ qubits, manteniendo la fidelidad tan alta como se quiera ($F > 1-\varepsilon$). Además, en caso de utilizar $S(\hat{\rho})-\delta$ qubits, la fidelidad nunca es superior a ε. Esto significa que $S(\hat{\rho})$ qubits representan el contenido de información cuántico, en promedio, de cada símbolo de la fuente, y constituye, por tanto, la mejor codificación posible.

Si se observan los resultados expuestos, es evidente el paralelismo existente entre el teorema de Shannon y el de Schumacher. Sin embargo, y de acuerdo con la expresión (3.24), la teoría cuántica puede llegar a permitir una compresión mayor de la información. ¿A qué se debe esta posibilidad? Nuevamente, a la imposibilidad de distinguir perfectamente estados cuánticos no ortogonales: cuando el alfabeto fuente lo constituyen kets ortogonales, la teoría clásica y cuántica coinciden, mientras que, en el caso contrario, la indistinguibilidad intrínseca de dichos estados convierte cierta información clásica en redundante y, por tanto, permite una mayor compresión.

3.5.2 Codificación de mezclas estadísticas. Información de Holevo

Sea una fuente cuántica como la de la subsección precedente, pero que, en lugar de emitir estados puros, genera una mezcla estadística $\hat{\rho}_x$ ($S(\hat{\rho}_x) \neq 0$) con probabilidad 1. De acuerdo con el teorema de Schumacher, sería necesario transmitir $S(\hat{\rho}_x)$ qubits por cada símbolo. Sin embargo, esto carece de sentido, ya que al tratarse de una fuente determinista, no se está enviando realmente ninguna información al receptor. Se propone entonces, para paliar esta deficiencia de la entropía de von Neumann, el concepto de información de Holevo. En realidad no es más que una generalización del teorema anterior, ya que establece el número mínimo de qubits necesarios para codificar los símbolos generados por la fuente (con la fidelidad que se desee), pero tanto en el caso de estados puros como de mezclas estadísticas.

En efecto, supóngase ahora que la fuente emite símbolos caracterizados por el operador densidad

$$\hat{\rho} = \sum_x p_x \hat{\rho}_x. \tag{3.39}$$

Se define la información de Holevo, y se denota por $\chi(\mathcal{E})$, como la función que actúa sobre el estado cuántico del sistema de acuerdo con la expresión

$$\chi(\mathcal{E}) = S(\hat{\rho}) - \sum_x p_x S(\hat{\rho}_x). \tag{3.40}$$

El hecho de restar el término $\sum_x p_x S(\hat{\rho}_x)$ a $S(\hat{\rho})$ tiene una interpretación similar al concepto de información mutua en la teoría clásica: el conocimiento de la preparación concreta de $\hat{\rho}$ implica necesariamente una reducción en la entropía de von Neumann. Nótese que, de acuerdo con su significado (información cuántica almacenada, en promedio, en cada símbolo de la fuente), se trata

[13] Recuérdese que se supone siempre una fuente discreta y sin memoria.

de una cantidad que no puede ser negativa. De hecho, es inmediato comprobar, véase la Sección 3.4, que

$$S(\sum_x p_x \hat{\rho}_x) \geq \sum_x p_x S(\hat{\rho}_x). \tag{3.41}$$

Además, se puede demostrar que la acción de un superoperador nunca puede incrementar la información de Holevo:

$$\$: \mathcal{E} = \{\hat{\rho}_x, p_x\} \longrightarrow \mathcal{E}' = \{\$(\hat{\rho}_x), p_x\}, \qquad \chi(\mathcal{E}') \leq \chi(\mathcal{E}). \tag{3.42}$$

Es decir, un canal puede mantener o reducir el contenido de información del símbolo enviado, pero en ningún caso incrementarlo.

3.6. Información accesible

Hasta el momento, la discusión se ha centrado en analizar el contenido de información cuántica (en qubits), de cada símbolo de una fuente cuántica, con especial hincapié en las repercusiones, sobre la compresión de la información, de la imposibilidad de distinguir perfectamente entre estados no ortogonales. Para finalizar el capítulo, estudiaremos, en esta sección y en la siguiente, la posibilidad de utilizar canales cuánticos para enviar información clásica. Pretendemos con ello cuantificar el contenido de información (en bits) que puede ser, en promedio, recuperado del estado de un sistema cuántico. A esta cantidad se la denomina información accesible. Distinguimos entre los casos de canales libres de errores y canales ruidosos.

3.6.1 Canal cuántico libre de errores

Supóngase que se desea transmitir la información clásica generada por una fuente, cuya entropía es $H(X)$, a través de un canal cuántico libre de errores. Obviamente, para lograr dicha transmisión es necesario emplear una fuente cuántica. Una posible solución podría consistir en establecer una aplicación biyectiva entre los símbolos del alfabeto fuente original, con sus probabilidades asociadas, y un conjunto de estados cuánticos $\{|\psi_x\rangle\}_x$. Pero, ¿qué fiabilidad brinda este esquema? Es decir, ¿en qué cantidad se reduce la incertidumbre de X por el conocimiento del resultado de la mejor de las medidas sobre el estado cuántico recibido? Para responder a estas preguntas se recurre al concepto de información accesible, $\mathrm{Acc}(\mathcal{E})$, que se define como la máxima información (en bits) que puede ser recuperada, mediante la observación de los resultados de la medida que maximiza dicha información, del estado de un sistema cuántico,

$$\mathrm{Acc}(\mathcal{E}) = \mathrm{Máx}_{\{\hat{F}_y\}} I(X, Y), \tag{3.43}$$

donde $\{\hat{F}_y\}$ identifica una POVM, Y es la variable aleatoria formada por los resultados de la medida y sus probabilidades asociadas, e $I(X, Y)$ representa la información mutua clásica[14] entre X e Y.

De acuerdo con lo visto en la Subsección 3.1.2, es inmediato comprobar que si el conjunto $\{|\psi_x\rangle\}_x$ es ortogonal, la mejor de las medidas posibles es una medida de von Neumann como la de la ecuación (3.6), con $\hat{P}_y = |\psi_y\rangle\langle\psi_y|$. De esta forma, la obtención del valor a_y determina perfectamente la preparación del sistema, ya que

$$p(a_y/x) = \delta_{yx}, \tag{3.44}$$

donde queda de manifiesto la perfecta concordancia con lo establecido por Shannon en la teoría clásica para un canal libre de errores, $H(X/Y) = 0$ e $I(X, Y) = H(X)$.

[14] La información mutua cuantifica la cantidad (en bits) en que se reduce la entropía de X por el conocimiento de Y. Se recuerda que su definición es: $I(X, Y) = H(X) - H(X/Y) = H(Y) - H(Y/X)$.

Sin embargo, resulta mucho más interesante analizar el caso de un alfabeto fuente cuántico formado por estados puros pero no mutuamente ortogonales. Pese a que no existe ninguna expresión para $\mathrm{Acc}(\mathcal{E})$, sí se conoce un límite superior:

$$\mathrm{Acc}(\mathcal{E}) \leq S(\hat{\rho}). \tag{3.45}$$

Además, se puede demostrar [42] que, mediante una codificación y decodificación adecuadas, siempre es posible acercarse de forma asintótica a dicho límite. Esto significa, por tanto, que pese a tratarse de un canal libre de errores, debido a la imposibilidad de distinguir perfectamente estados no ortogonales, se produce una degradación de la información clásica. La información accesible nunca puede superar la entropía de von Neumann, que es menor que $H(X)$.

3.6.2 Canal cuántico ruidoso El análisis de la información accesible en un canal cuántico ruidoso es casi igual al del canal libre de errores. La única diferencia estriba en que ahora también se permite que el alfabeto de la fuente cuántica esté formado por mezclas estadísticas. Debido al efecto de decoherencia, que transforma estados puros en mezclas estadísticas, un canal cuántico ruidoso se puede modelar mediante un superoperador, \$, que actúa directamente sobre los símbolos del alfabeto fuente,

$$|\psi_x\rangle\langle\psi_x| \longrightarrow \$(|\psi_x\rangle\langle\psi_x|) = \hat{\rho}_x. \tag{3.46}$$

Pues bien, se puede demostrar [51] que en este caso, y para canales sin memoria, el límite superior de la información accesible, que se denomina ahora límite de Holevo, lo constituye la información de Holevo del conjunto $\{\hat{\rho}_x, p_x\}$

$$\mathrm{Acc}(\mathcal{E}) \leq \chi(\mathcal{E}). \tag{3.47}$$

Además, y de manera análoga a la sección precedente, siempre es posible, mediante una codificación y decodificación adecuadas[15], acercarse de manera asintótica al valor $\chi(\mathcal{E})$. De esta manera, y de acuerdo con la expresión

$$\chi(\mathcal{E}) \leq S(\hat{\rho}) \leq 1, \tag{3.48}$$

es evidente que tanto para el caso de canales cuánticos libres de errores como ruidosos, la máxima información clásica que puede ser almacenada en un qubit es un bit.

3.7. Capacidad clásica de un canal cuántico

Dado un canal cuántico sin memoria y caracterizado por un superoperador \$, se define su capacidad clásica como la máxima información (en bits) que puede transmitir con una probabilidad de error tan pequeña como se desee. Dado que la información mutua es convexa respecto a las distribuciones de probabilidad de la fuente, la capacidad no es más que su máximo,

$$C = \mathrm{Máx}_{\{p_x, \hat{F}_y\}} I(X, Y). \tag{3.49}$$

Así, teniendo en cuenta el límite de Holevo, y que mediante una codificación adecuada es posible acercarse asintóticamente a su valor, se puede reescribir la expresión anterior como

$$C(\$) = \mathrm{Máx}_{\{\mathcal{E}\}} \chi(\$(\mathcal{E})), \tag{3.50}$$

en donde se han incluido, mediante \$, las posibles alteraciones introducidas por el canal.

[15] Al lector interesado se le remite a [54].

Para finalizar, conviene resaltar que todos los resultados que se han obtenido en este capítulo presuponen siempre que en la codificación de las secuencias de símbolos generadas por la fuente no se utilizan estados "entangled". De hecho, la posibilidad de conseguir tasas de transmisión mayores, mediante el empleo de este último tipo de estados, es una cuestión abierta.

CAPÍTULO 4

Panorámica de aplicaciones

CONTENIDOS

El análisis realizado en 1964 por John Bell sobre la imposibilidad de imitar, mediante cualquier ley física clásica, las correlaciones existentes entre sistemas cuánticos que hayan interactuado en un pasado, constituye uno de los pilares esenciales de la Teoría de la Información Cuántica. De hecho, el mayor potencial ofrecido por el procesado cuántico de la información respecto al clásico se basa, en gran medida, en la utilización adecuada de los sistemas "entangled". De este modo, se puede conseguir transmitir a través de canales cuánticos información clásica al doble de la velocidad impuesta por el límite de Holevo (códigos densos); utilizar la mayor robustez de los bits clásicos para enviar qubits desconocidos (teleportación cuántica); establecer la clave a utilizar en la encriptación de un mensaje de manera totalmente segura (criptografía cuántica); o procesar información paralelamente de forma masiva (computación cuántica). A continuación presentamos, de manera concisa, las principales ideas subyacentes de todas estas aplicaciones.

4.1. Codificación densa

Considérese el escenario siguiente: Dos entidades cualesquiera (A y B) desean intercambiarse información a través de un canal cuántico ruidoso. Una posibilidad es establecer una aplicación biyectiva tal que: $0 \longrightarrow |0\rangle$ y $1 \longrightarrow |1\rangle$, de manera que para transmitir, por ejemplo, la secuencia de bits 01, se utilicen dos qubits preparados en un estado $|01\rangle$. De este modo, un receptor siempre podría recuperar con total fidelidad la información transmitida; únicamente precisaría realizar una medida de von Neumann en la base $\{|0\rangle, |1\rangle\}$ sobre cada qubit.

Supóngase ahora que además las entidades A y B comparten dos qubits en un estado "entangled": $|\phi^+\rangle_{AB} = (|00\rangle + |11\rangle)/\sqrt{2}$; es decir, cada entidad posee un qubit del par. En este caso, A podría enviar la información clásica con el mismo método descrito en el párrafo anterior, o intentar

pensar en algún nuevo esquema de transmisión que le permitiese rentabilizar este recurso añadido. Pues bien, si A actúa convenientemente, se puede demostrar que es posible conseguir transmitir los 2 bits del ejemplo mediante un único qubit.

En efecto, considérese que A, antes de enviar el qubit del par "entangled" que comparte con B, y dependiendo de los dos bits que desea transmitir, efectúa en él una de las cuatro siguientes transformaciones unitarias:

$$\hat{1} : |\phi^+\rangle_{AB} \longrightarrow |\phi^+\rangle_{AB}, \tag{4.1}$$

$$\hat{X} : |\phi^+\rangle_{AB} \longrightarrow |\psi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \tag{4.2}$$

$$\hat{Y} : |\phi^+\rangle_{AB} \longrightarrow |\psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle), \tag{4.3}$$

$$\hat{Z} : |\phi^+\rangle_{AB} \longrightarrow |\phi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle). \tag{4.4}$$

De esta manera, cuando dicho qubit llegue al receptor (entidad B), éste siempre podrá, mediante una medida ortogonal en el par "entangled" (proyectando sobre la base $\{|\phi^+\rangle_{AB}, |\phi^-\rangle_{AB}, |\psi^+\rangle_{AB}, |\psi^-\rangle_{AB}\}$), conocer qué operación realizó A, y, por tanto, determinar qué dos bits se pretendían enviar. En definitiva, se ha logrado, mediante la transmisión de un solo qubit, proporcionar al receptor dos bits de información clásica. A este procedimiento se le llama codificación densa [25, 9].

Resulta interesante comprobar que mediante este procedimiento se asegura la confidencialidad de la transmisión, ya que aunque un desconocido interceptase el qubit emitido por A, nunca podría acceder a ninguna información[1], ya que ésta se encuentra codificada en las correlaciones existentes entre los dos qubits "entangled", y, por consiguiente, la única manera de acceder a ella es mediante una medida conjunta en el par.

Para concluir, queda por determinar si este resultado supone alguna violación del límite de Holevo. Resulta trivial constatar que no es así, porque si se analiza con detenimiento el modelo de transmisión expuesto, es evidente que en realidad se utilizan dos qubits: el que envía A y el que previamente tuvo que ser transmitido para poder compartir el par de qubits "entangled". Los códigos densos, por tanto, no son más que un simple reflejo de la posibilidad de utilizar las correlaciones pre-existentes entre qubits de diferentes entidades como recurso para comunicaciones más eficientes[2].

4.2. Teleportación cuántica

Supóngase que dos entidades (A y B), comunicadas únicamente por un canal clásico, necesitan intercambiar un determinado qubit $|\psi\rangle_C = a|0\rangle_C + b|1\rangle_C$. Si el emisor (por ejemplo A) conociese perfectamente el estado de dicho qubit, siempre podría proporcionar suficiente información clásica a B (por ejemplo, "el estado del qubit es $a|0\rangle_C + b|1\rangle_C$") de manera que éste pudiese prepararlo. Pero, ¿qué sucede cuando el estado de $|\psi\rangle_C$ es desconocido? En este caso, A no puede realizar ninguna medida que le facilite información acerca del qubit, ya que ésta inevitablemente provocaría un colapso del estado del mismo. ¿Qué hacer entonces? De manera análoga al caso de códigos densos, se puede pensar en compartir entre las entidades A y B dos qubits en un estado "entangled": $|\phi^+\rangle_{AB} = (|00\rangle + |11\rangle)/\sqrt{2}$, y tratar de aprovechar estos nuevos recursos. Así, el estado de estos tres qubits puede describirse ahora como

[1] Recuérdese de que $\hat{\rho}_A = \frac{1}{2}\hat{1}_A$.

[2] Experimentos que corroboran estos resultados se pueden encontrar en [48, 61, 27].

$$|\psi\rangle_C|\phi^+\rangle_{AB} = (a|0\rangle_C + b|1\rangle_C)\frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB})$$

$$= \tfrac{1}{\sqrt{2}}(a|000\rangle_{CAB} + a|011\rangle_{CAB} + b|100\rangle_{CAB} + b|111\rangle_{CAB})$$

$$= \tfrac{1}{2}a(|\phi^+\rangle_{CA} + |\phi^-\rangle_{CA})|0\rangle_B + \tfrac{1}{2}a(|\psi^+\rangle_{CA} + |\psi^-\rangle_{CA})|1\rangle_B$$

$$+\tfrac{1}{2}b(|\psi^+\rangle_{CA} - |\psi^-\rangle_{CA})|0\rangle_B + \tfrac{1}{2}b(|\phi^+\rangle_{CA} - |\phi^-\rangle_{CA})|1\rangle_B$$

$$= \tfrac{1}{2}|\phi^+\rangle_{CA}(a|0\rangle_B + b|1\rangle_B) + \tfrac{1}{2}|\psi^+\rangle_{CA}(a|1\rangle_B + b|0\rangle_B)$$

$$+\tfrac{1}{2}|\psi^-\rangle_{CA}(a|1\rangle_B - b|0\rangle_B) + \tfrac{1}{2}|\phi^-\rangle_{CA}(a|0\rangle_B - b|1\rangle_B)$$

$$= \frac{1}{2}|\phi^+\rangle_{CA}|\psi\rangle_B + \frac{1}{2}|\psi^+\rangle_{CA}\hat{X}|\psi\rangle_B + \frac{1}{2}|\psi^-\rangle_{CA}(-i\hat{Y})|\psi\rangle_B + \frac{1}{2}|\phi^-\rangle_{CA}\hat{Z}|\psi\rangle_B \tag{4.5}$$

A la vista de este resultado, es evidente que si A realiza una medida de von Neumann que proyecte los dos qubits que posee en la base formada por los kets $\{|\phi^+\rangle_{CA}, |\phi^-\rangle_{CA}, |\psi^+\rangle_{CA}, |\psi^-\rangle_{CA}\}$, seguida de la transmisión del resultado obtenido (se necesitan dos bits para identificarlo) a la entidad B, ésta última siempre podrá preparar un qubit en el estado $|\psi\rangle$. Para ello, de acuerdo con la ecuación (4.5), y dependiendo de los dos bits recibidos, únicamente tendría que aplicar a su qubit del par "entangled" uno de los cuatro operadores siguientes[3]:

$$\begin{aligned} |\phi^+\rangle_{CA} &\longrightarrow \hat{1}_B, \\ |\psi^+\rangle_{CA} &\longrightarrow \hat{X}_B, \\ |\psi^-\rangle_{CA} &\longrightarrow \hat{Y}_B, \\ |\phi^-\rangle_{CA} &\longrightarrow \hat{Z}_B. \end{aligned} \tag{4.6}$$

En resumen, se puede conseguir que el qubit que poseía B adquiera el estado $|\psi\rangle$[4]. A este proceso se le denomina teleportación cuántica [22, 17].

Una de las grandes aplicaciones potenciales de este procedimiento de comunicación es la transmisión de información cuántica a través de canales cuánticos ruidosos [24]; esto es, cabe emplear dichos canales para distribuir qubits de pares "entangled", y posteriormente utilizar canales clásicos (más fiables) para la transferencia de la información. Además, si existiese una entidad independiente que realizase el reparto de los pares "entangled", no sería necesario conocer la dirección del receptor (se podrían emplear técnicas de difusión).

4.3. Criptografía cuántica

La mayoría de las técnicas que permiten garantizar una comunicación confidencial entre dos entidades cualquiera, A y B, está basada en la utilización de una secuencia aleatoria de bits que se denomina clave y que se emplea en la encriptación del mensaje. Los inconvenientes surgen al intentar establecer, de forma segura, la clave de una comunicación concreta. Lo habitual es emplear los denominados criptosistemas de clave pública, como por ejemplo el conocido RSA [52, 43], que

[3] No hay más que tener en cuenta que: $\hat{X}\hat{X} = \hat{1}$,$\hat{Y}\hat{Y} = \hat{1}$, $\hat{Z}\hat{Z} = \hat{1}$.

[4] Nótese que no se está incumpliendo la imposibilidad de copiar qubits desconocidos, ya que antes de que el qubit de B adopte el estado $|\psi\rangle_C$, éste desaparece en el transmisor A (recuérdese que se efectúa una medida que produce un colapso). Luego, realmente el estado se destruye en el emisor y posteriormente (tras recibir los dos bits y efectuar la operación correspondiente) se crea en el receptor.

basan su invulnerabilidad en la dificultad de descomponer grandes números en sus factores primos. El problema de estos criptosistemas es que su seguridad no es incondicional, sino computacional.

La criptografía cuántica fue establecida a principios de los años 80 por Bennett y Brassard [23, 21, 20, 19], si bien sería preciso considerar a su vez los trabajos preliminares de Wiesner [62]. A continuación, y a modo de ejemplo, simplemente se mostrará uno de los muchos métodos posibles.

Supóngase que A y B desean establecer una clave criptográfica de n bits. Para ello, A transmite a B un conjunto de $4n$ qubits preparados cada uno, de manera aleatoria, en alguno de los siguientes cuatro estados: $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$[5]. A continuación, B debe realizar una medida ortogonal sobre la secuencia de qubits recibida, eligiendo de manera aleatoria entre las bases $\{|0\rangle, |1\rangle\}$ y $\{|+\rangle, |-\rangle\}$. Por último, ambas entidades se comunican de manera pública (no es necesario que sea confidencial) la base concreta que utilizaron para preparar o medir, respectivamente, cada qubit, y se quedan únicamente con aquellos resultados en los que coinciden (utilizaron la misma base). De este modo, y con tal de asignar posteriormente a los estados $|0\rangle$ y $|+\rangle$ el bit 0, y a $|1\rangle$ y $|-\rangle$ el bit 1, ya se consigue compartir una clave con longitud $2n$[6].

¿Por qué utilizar este método y no enviar directamente n bits aleatorios como clave? Supóngase que C intercepta el mensaje que viaja por el canal, lo lee, y lo reenvía a B. Con los métodos clásicos, y si C es suficientemente rápido, resulta imposible detectar cuándo se ha producido esta acción. Sin embargo, si se utilizan $4n$ qubits para la transmisión de la clave, cuando C intente determinar su valor, tan sólo podrá conocer, en promedio, $2n$ (aquellos que proyectó en la base que eligió A), mientras que los restantes resultarán inevitablemente alterados[7]. Cuando B realice su medida, por el mismo motivo expuesto anteriormente, se quedará sólo con $2n$ qubits (aquellos que proyectó en la misma base que utilizó A). Ahora bien, de estos $2n$ qubits que comparten A y B, debido a la actuación de C, $n/2$ no tienen por qué coincidir. Esto es así porque en promedio, estos $2n$ qubits provienen de n qubits que fueron alterados por C y otros n qubits que no fueron modificados. El problema reside en los n que variaron su valor. Cuando B mida este conjunto, únicamente conseguirá que la mitad $(n/2)$, en promedio, colapsen a su estado original, mientras que la otra mitad se mantendrá en un estado incorrecto. Este resultado es crucial, ya que permite establecer un mecanismo muy sencillo para detectar la presencia de C. Para ello, A y B simplemente necesitan intercambiar públicamente n bits de la clave. De este modo, si éstos coinciden se tiene asegurado que C no estuvo presente[8], y por tanto, la clave la constituyen los restantes n bits que no se dieron a conocer.

En la práctica, el protocolo es más complicado de lo que se acaba de mostrar, ya que C puede actuar con otras estrategias (por ejemplo, tan sólo interceptar ciertos qubits) y además la propia utilización de canales cuánticos ruidosos puede producir errores aun en ausencia de C. En este caso más general nunca se rechaza la clave salvo que el número de bits no coincidentes supere un cierto porcentaje. Posteriormente se procesa la información en dos pasos. En primer lugar se corrigen los errores, para lo cual se realiza una comparación pública de la paridad de ciertos subconjuntos de bits escogidos aleatoriamente[9], y, por último, se vuelven a calcular diversos valores de la paridad de distintos subconjuntos de la clave corregida, y a partir de estos valores se establece una nueva[10].

---

[5] $\{|0\rangle, |1\rangle\}$ y $\{|+\rangle, |-\rangle\}$ simplemente representan dos bases diferentes de un espacio de Hilbert bidimensional, que verifican $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ y $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$.

[6] Esto es así porque, en promedio, B únicamente elegirá en la mitad de los casos medir los qubits recibidos con la misma base que utilizó A en su preparación.

[7] Es trivial comprobar que si se proyecta, por ejemplo, el qubit $|0\rangle$ en la base $\{|+\rangle, |-\rangle\}$ su estado colapsará con la misma probabilidad a: $|+\rangle$ ó $|-\rangle$. Del mismo modo se pueden deducir los demás casos.

[8] La probabilidad de que C estuviese presente y se seleccionasen n bits coincidentes es: $(3/4)^{n/2} \simeq 10^{-125}$ para $n = 1000$.

[9] Además, y para no incrementar la información de C, se descartan ciertos bits.

[10] Con un porcentaje de error del 4 % (debido básicamente a la actuación de C) se puede reducir la clave original de 2000 bits a 754, de manera que C conozca menos que 10^{-6} de un bit [37].

4.4. Computación cuántica

El concepto de computación cuántica fue desarrollado a principios de la década de los 80, de manera independiente, por Paul Benioff [14, 15, 16] y Richard Feynman [38, 39]. Básicamente, una computación cuántica consiste en aplicar una determinada transformación unitaria a un conjunto de n qubits. Para lograrlo, se pueden emplear, de manera sucesiva, varias puertas cuánticas sobre distintos subconjuntos de uno o dos qubits. Finalmente, es necesario medir el estado de todos ellos para obtener el resultado[11].

Sin embargo, y pese al cariz "mágico" que parece rodear al adjetivo cuántico, es importante resaltar una cuestión que muchas veces no queda muy clara: un computador cuántico no puede realizar nada que no sea posible en uno clásico, es decir, la idea de computabilidad es la misma. Bien es cierto que los principios físicos en los que descansan uno y otro son diferentes, pero de la misma manera que un computador clásico puede almacenar vectores, rotarlos, o modelar el proceso de medición cuántica mediante una proyección ortogonal, éste también puede simular cualquier computador cuántico. Lo realmente significativo es lo ineficiente de esta simulación. Así, si se quisiese describir el estado de un computador cuántico de, por ejemplo, tan sólo 100 qubits, sería necesario disponer de una memoria que pudiese almacenar $2^{100} \sim 10^{30}$ números complejos. Esto sugiere, aunque todavía no está probado, la imposibilidad de realizar una simulación polinómica de un computador cuántico empleando uno clásico; es decir, la máquina de Turing [56] no es un modelo adecuado para las computaciones que se pueden llevar a cabo en el mundo físico. Estos resultados fueron los que llevaron a Benioff y Feynman a especular con la posibilidad de emplear computadores cuánticos para resolver problemas de complejidad muy elevada.

En las siguientes subsecciones trataremos un poco más estas importantes cuestiones, comentando tanto la posibilidad de procesar paralelamente la información cuántica, como la manera de realizar las complejas transformaciones unitarias que habitualmente se requieren. Para finalizar, se expondrán algunos algoritmos concretos en los que quedan de manifiesto las grandes ventajas potenciales de la computación cuántica.

4.4.1 Procesado paralelo de la información

La computación cuántica posee el potencial de resolver, en mucho menor tiempo, problemas de complejidad intratable para la computación clásica. Este potencial reside en la capacidad de procesar información paralelamente de forma masiva[12]. De hecho, mediante una transformación unitaria adecuada en los 100 qubits del ejemplo anterior, se puede conseguir el equivalente a 2^{100} computaciones simultáneas sobre un conjunto de bits. El secreto se encuentra en la explotación oportuna de la información codificada en las correlaciones "no locales" existentes entre las diferentes partes de un sistema cuántico. Para esclarecer esta idea, a continuación se expondrá el problema enunciado por Deutsch.

Supóngase que se dispone de un dispositivo clásico que transforma un bit cualquiera x en otro bit $f(x)$, y se desea conocer si $f(0) = f(1)$ ó $f(0) \neq f(1)$. Trivialmente, la única forma de solucionar esta cuestión es evaluando la función para cada caso.

Supóngase ahora que se dispone de un computador cuántico capaz de hacer evolucionar dos qubits cualesquiera ($|x\rangle$ e $|y\rangle$) según una transformación unitaria $\hat{U}$:

$$\hat{U} : |x\rangle|y\rangle \longrightarrow |x\rangle|y \oplus f(x)\rangle. \tag{4.7}$$

Si $|x\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ y $|y\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, es inmediato comprobar que

[11] Nótese que, debido al carácter intrínsecamente aleatorio del proceso de medición (cuarto postulado), la ejecución consecutiva del mismo algoritmo en un computador cuántico no tiene por qué proporcionar en general el mismo resultado.

[12] Este hecho fue anunciado por primera vez por David Deutsch [31] en 1985.

$$\hat{U} : \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \longrightarrow \frac{1}{\sqrt{2}}[(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle]\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \tag{4.8}$$

De esta manera, con tal de realizar una proyección ortogonal del primer qubit en una base $\{|+\rangle, |-\rangle\}$, donde

$$|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle), \tag{4.9}$$

ya se habría resuelto, con una única computación, el problema planteado[13]. Esto ha sido posible porque al utilizar un dispositivo cuántico, éste no tiene por qué limitarse a evaluar $f(0)$ ó $f(1)$, sino que puede actuar sobre una superposición de los estados $|0\rangle$ y $|1\rangle$, y, por tanto, permite extraer información "global" de la función, esto es, información que depende tanto de $f(0)$ como de $f(1)$.

De igual modo, se puede emplear este procesado paralelo de la información para conocer propiedades de funciones más complicadas. Así, si se quisiese calcular, para todas las posibles combinaciones de un conjunto de N bits, el valor de una función cualquiera f, un computador clásico necesitaría realizar 2^N evaluaciones de la función. Sin embargo, utilizando uno cuántico, solamente se precisaría efectuar, sobre un conjunto de N qubits, una transformación unitaria $\hat{U}$ tal que

$$\hat{U} : |x\rangle|0\rangle \longrightarrow |x\rangle|f(x)\rangle, \tag{4.10}$$

ya que con tal de preparar cada qubit en un estado $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, el estado del conjunto sería

$$\overbrace{[\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)] \otimes \ldots \otimes [\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)]}^{N} = \frac{1}{2^{N/2}} \sum_{x=0}^{2^N-1} |x\rangle, \tag{4.11}$$

y, en consecuencia, al aplicar $\hat{U}$ se tendría

$$\frac{1}{2^{N/2}} \sum_{x=0}^{2^N-1} |x\rangle|f(x)\rangle, \tag{4.12}$$

estado que ya recoge las propiedades globales de la función[14]. Posteriormente, cuando introduzcamos algún algoritmo cuántico concreto, mostraremos diversas maneras de acceder a esta información.

4.4.2 Puertas cuánticas En la sección precedente siempre se ha supuesto que un computador cuántico es capaz de realizar la transformación unitaria requerida (ecuaciones (4.7) y (4.10)). A continuación mostraremos cómo, a partir de sencillas puertas cuánticas que actúen en uno o dos qubits, esto es posible.

Una puerta cuántica [31, 32] es simplemente una operación unitaria sobre un conjunto de qubits. De hecho, aunque no se ha indicado de forma explícita a lo largo de esta obra, las puertas cuánticas ya se han empleado en múltiples ocasiones. Como ejemplo se pueden citar los operadores de Pauli,

$$\hat{I} = |0\rangle\langle 0| + |1\rangle\langle 1|, \tag{4.13}$$

$$\hat{X} = |0\rangle\langle 1| + |1\rangle\langle 0|, \tag{4.14}$$

[13] Obviamente, si el resultado de la proyección ortogonal del primer qubit en la base $\{|+\rangle, |-\rangle\}$ es $|+\rangle$ entonces $f(0) = f(1)$, y, en caso contrario, $f(0) \neq f(1)$.

[14] La información global de la función se encuentra ahora codificada en las correlaciones "no locales" existentes entre los qubits preparados en un estado $|x\rangle$ y los que se encuentran en un estado $|f(x)\rangle$. Nótese que no es sencillo acceder a ella, ya que una medida de von Neumann provocaría un colapso al estado $|x_0\rangle|f(x_0)\rangle$, y, por tanto, no se adquiriría ninguna ventaja respecto a la computación clásica.

$$\hat{Z} = |0\rangle\langle 0| - |1\rangle\langle 1|, \tag{4.15}$$

$$\hat{Y} = \hat{X}\hat{Z}. \tag{4.16}$$

Ahora bien, mientras que en la teoría clásica las puertas lógicas constituyen un conjunto claramente finito[15], debido a que el espacio de estados de un qubit es continuo, el número de posibles transformaciones unitarias también lo es, y, en consecuencia, existen infinitas puertas cuánticas. Sin embargo, es posible demostrar [7, 33, 34, 47] que cualquier transformación unitaria en un conjunto de n qubits puede realizarse mediante la aplicación sucesiva de tan sólo dos puertas cuánticas[16]: la asociada a la operación XOR, y la de rotación, $V(\theta, \phi)$.

El operador XOR es un caso particular de un conjunto de operadores que actúan sobre un par de qubits y que pueden ser representados mediante la expresión

$$|0\rangle\langle 0|\hat{I} + |1\rangle\langle 1|\hat{U}, \tag{4.17}$$

donde $\hat{U}$ identifica una transformación unitaria cualquiera sobre un único qubit. Es decir, mientras que el primer qubit permanece inalterado, al segundo se le aplica $\hat{I}$ ó $\hat{U}$ dependiendo del estado del primero. Concretamente, la puerta XOR transforma un estado $|x\rangle|y\rangle$ en $|x\rangle|x \oplus y\rangle$, donde $\oplus$ identifica la operación lógica O-exclusiva (simplemente habría que sustituir en la ecuación (4.17) el operador $\hat{U}$ por $\hat{X}$).

Por su parte, la operación de rotación $V(\theta, \phi)$ se puede describir como

$$V(\theta,\phi) = |0\rangle(\cos(\theta/2)\langle 0| - ie^{-i\phi}\text{sen}(\theta/2)\langle 1|) + +|1\rangle(-\text{ie}^{-\text{i}\phi}\text{sen}(\theta/2)\langle 0| + \cos(\theta/2)\langle 1|), \tag{4.18}$$

donde θ y ϕ son dos números irracionales[17].

4.4.3 Ejemplos de algoritmos cuánticos Para finalizar con este breve recorrido por la computación cuántica, introduciremos algunos algoritmos especialmente interesantes: el cálculo del período de un función (muchos problemas se pueden reducir a esta operación), la factorización de un número (como ya se ha visto, la mayoría de los criptosistemas actuales se basan el la intratabilidad de este problema), y el algoritmo de Grover (permite, entre otras cosas, buscar eficientemente en bases de datos desordenadas). Además, tal y como se ha avanzado, estos ejemplos servirán para ilustrar diferentes maneras de acceder a la información codificada en las correlaciones "no locales" entre los distintos qubits.

4.4.3.1 Cálculo del período de una función Los algoritmos clásicos existentes para calcular el período p de una función son sumamente ineficientes, ya que el número de veces que es necesario evaluar la función crece exponencialmente con $\log_2 N$, donde N representa un número tal que $N/2 < p < N$. A continuación se mostrará cómo, utilizando un computador cuántico, es posible realizar un algoritmo de orden polinómico [55].

Supongamos que se dispone de un computador cuántico que posee dos registros (x e y) de n qubits cada uno, donde $n = \lceil 2\log_2 N \rceil$. De acuerdo con la Subsección 4.4.1, si se prepara cada qubit del primer registro en un estado $(|0\rangle + |1\rangle)/\sqrt{2}$, y se mantiene el segundo en un estado $|0\rangle$[18], siempre se puede efectuar una transformación unitaria $\hat{U}$ tal que

[15] Por ejemplo, únicamente hay dos puertas lógicas que puedan aplicarse a un bit: la identidad y la negación. Del mismo modo, se puede comprobar que sobre 2 bits sólo pueden actuar 16 puertas lógicas distintas, y así sucesivamente.

[16] El concepto de puerta cuántica universal es totalmente equivalente al existente en la computación clásica, donde con únicamente puertas NAND se puede efectuar cualquier operación lógica.

[17] Es necesario que θ y ϕ sean irracionales para poder conseguir efectuar una transformación $V(\theta, \phi)$ (con θ y ϕ continuos) mediante el empleo reiterado de una misma puerta cuántica con estos valores perfectamente determinados.

[18] Para aliviar la notación, con el ket $|0\rangle$ se denota el estado de n qubits. Lo correcto sería emplear $\overbrace{|0\rangle \otimes \ldots \otimes |0\rangle}^{n}$.

$$\frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} |x\rangle|0\rangle \longrightarrow \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} |x\rangle|f(x)\rangle. \tag{4.19}$$

De este modo, al realizar una medida ortogonal del registro y que proporcione, por ejemplo, un resultado t ($f(x) = t$), el estado de los $2n$ qubits del computador colapsará a

$$\frac{1}{M^{1/2}} \sum_{k=0}^{M-1} |d_u + cp\rangle|t\rangle, \tag{4.20}$$

donde $d_u + cp$, con $c = 1, \ldots, M-1$, representa todos aquellos valores x para los que $f(x) = t$. Es decir, el registro x queda preparado en una superposición de $M \simeq 2^n/p$ estados separados entre sí una distancia p. El parámetro d_u simplemente identifica un desplazamiento que, obviamente, depende del resultado t de la medida. Así, con tal de aplicar el operador unitario $\hat{U}_{FFT}$,[19]

$$\hat{U}_{FFT} : |x\rangle \longrightarrow \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} e^{i2\pi kx/2^n} |k\rangle, \tag{4.21}$$

al registro x, se obtiene[20]

$$\hat{U}_{FFT} : \frac{1}{\sqrt{2^n/p}} \sum_{c=0}^{2^n/p-1} |d_u + cp\rangle \longrightarrow \frac{1}{p^{1/2}} \sum_k \tilde{f}(k)|k\rangle, \tag{4.22}$$

donde

$$|\tilde{f}(k)| = \begin{cases} 1, & \text{si } k \text{ es un múltiplo de } 2^n/p; \\ 0, & \text{en otro caso.} \end{cases} \tag{4.23}$$

Esto significa que, con tal de medir finalmente este registro, ya se obtendría un múltiplo de $2^n/p$ ($x = \lambda 2^n/p$) a partir del cual es posible deducir p. Para ello, únicamente hay que tener en cuenta que si el máximo común divisor de λ y p es uno, es posible reducir $x/2^n$ a su fracción irreducible y ya inmediatamente obtener λ y p. El problema surge cuando tal máximo común divisor no es uno; en este el algoritmo falla, y, en consecuencia, sería necesario volver a ejecutarlo[21]. En cualquier caso, se puede demostrar [36] que se requerirían como máximo $\log_2 p$ computaciones para conseguir, con probabilidad prácticamente 1, el valor p buscado.

4.4.3.2 Factorización de un número Hallar los factores primos de un número compuesto cualquiera es un problema sumamente interesante, y no sólo desde la perspectiva abstracta de la teoría de la complejidad, sino también desde un punto de vista práctico (es la base de los criptosistemas de clave pública). Aunque no está demostrado, se cree con cierta seguridad que el tiempo requerido es superpolinómico en $\log_2 N$ (N es el número que se desea descomponer); esto es, a medida que N aumenta, el tiempo necesario (en el peor caso) crece más que cualquier potencia de $\log_2 N$. De hecho, en la actualidad, el mejor algoritmo que se conoce precisa de un tiempo aproximado [51]:

$$\exp\,[c(\ln N)^{1/3}(\ln \ln N)^{2/3}], \tag{4.24}$$

[19] Se trata de la versión cuántica de la FFT ("Fast Fourier Transform"). Además, es posible realizar esta transformación en un tiempo $(\log_2 N)^2$ [8, 29, 36, 51]. Nótese que los mejores algoritmos clásicos son de orden $N \log_2 N$.

[20] Por simplicidad en la exposición, se considerará que $M = 2^n/p$. Al interesado en el cálculo exacto se le remite a [55, 36].

[21] Recuérdese que el proceso de medida es probabilístico, esto es, cada realización de un algoritmo cuántico puede dar un resultado diferente.

donde $c = (64/9)^{1/3} \sim 1{,}9$. Esto significa que, mientras que para factorizar un número de 130 dígitos se tarda cerca de un mes[22], para factorizar uno de 400 es necesario 10^{10} años (la edad del universo).

A continuación se mostrará el algoritmo de factorización de Peter Shor [55] que, empleando un computador cuántico, es capaz de reducir el problema a un tiempo polinómico ($O[(\ln N)]^3$). De este modo, para el número anterior de 130 dígitos tan sólo se precisaría de unos pocos segundos[23], y para el de 400 apenas unos minutos.

Este algoritmo cuántico basa su potencia en localizar, utilizando el método de la sección precedente, el período de una función adecuada, ya que el resto es simplemente teoría de números. Así, para factorizar un número cualquiera N, la idea esencial consiste en determinar el período p de $f_N(x)$, dada por

$$f_N(x) = y^x \mathrm{mod} N, \tag{4.25}$$

donde y es un número elegido aleatoriamente que únicamente tiene que verificar $y < N$, y que el máximo común divisor entre él y N sea 1. Nótese que la última condición se puede comprobar fácilmente empleando el algoritmo de Euclides [36, 41]. En cualquier caso, si el máximo común divisor no fuera uno, tampoco sería un problema, ya que esto significaría sencillamente que ya se encontró un factor de N.

Una vez obtenido p, se calcula $y^{p/2}\mathrm{mod}N$[24], y, si el resultado es c ó $N - c$, con $c > 1$, ya se puede detener el algoritmo (los factores primos son el máximo común divisor de $c + 1$ y n, y el máximo común divisor de $c - 1$ y N[25]. En caso contrario ($c = 1$), simplemente se selecciona otro y y se repite el proceso[26].

Esquemas completos de las redes de puertas cuánticas necesarias para realizar este algoritmo se pueden encontrar en [49, 59, 10].

4.4.3.3 Algoritmo de Grover Se trata de un algoritmo muy atractivo, ya que puede ser empleado tanto para localizar de manera eficiente un determinado elemento en una base de datos desorganizada[27], como para resolver aquellos problemas en los que es muy complicado encontrar una solución, pero a la vez muy sencillo probar posibles candidatas.

Matemáticamente, el problema se puede reducir a encontrar un $x \in \{0, 1, 2 \ldots, N - 1\}$ con $N = 2^n \gg 1$, tal que $f(x) = a$ para un a conocido. La idea propuesta por Grover se basa en la utilización de una función f_s tal que, si s denota el x buscado,

$$\begin{aligned} f_s(x) &= 0, \qquad x \neq s, \\ f_s(x) &= 1, \qquad x = s. \end{aligned} \tag{4.26}$$

Desde la perspectiva de un computador cuántico (con dos registros de n y 1 qubits, respectivamente), lo que se necesita es una transformación unitaria $\hat{U}_{f_s}$ del tipo

[22] Suponiendo que se dispone de cientos de estaciones de trabajo colaborando en red.

[23] Suponiendo que el computador cuántico es capaz de efectuar tantas puertas cuánticas por segundo como puertas lógicas por segundo pueden ejecutar las estaciones de trabajo colaborando en red del ejemplo anterior.

[24] Es evidente que se necesita que p sea par. En caso contrario, se escoge otro y.

[25] Esto es así porque si p es el período de la función definida en la ecuación (4.25), $y^p \mathrm{mod} N = 1$, ó, equivalentemente, $z^2 \mathrm{mod} N = 1$ con $z = y^{p/2}$. Esto significa que las soluciones triviales de $z \mathrm{mod} N$ son 1 y $N - 1$. Ahora bien, también es posible que sean c y $N - c$, para $c > 1$. En este último caso se tiene que $(c - 1) \neq 0$ y $(c + 1) \neq N$, pero $(c + 1)(c - 1)\mathrm{mod} N = 0$, esto es, $(c + 1)(c - 1)$ es un múltiplo de N aunque $(c + 1)$ y $(c - 1)$ no lo sean. Luego, el máximo común divisor de N y $(c \pm 1)$ no puede ser N, y, por tanto, se trata de un factor de N.

[26] Se puede demostrar [36] que la probabilidad de tener que cambiar y es muy pequeña.

[27] Si se dispone de una lista de tamaño N, la teoría clásica necesita, en promedio, leer $N/2$ valores. Sin embargo, con el algoritmo de Grover [40] se requieren tan sólo del orden de $\sqrt{N}$ iteraciones. Además, se ha demostrado [18] que este algoritmo es óptimo, esto es, no existe ningún otro algoritmo cuántico de menor orden que pueda resolver el problema.

$$\hat{U}_{f_s} : |x\rangle|y\rangle \longrightarrow |x\rangle|y \oplus f_s(x)\rangle, \tag{4.27}$$

donde $|x\rangle$ es el estado del primer registro y $|y\rangle$ es el estado del qubit restante. Así, si se prepara este qubit como $(|0\rangle - |1\rangle)/\sqrt{2}$, se tiene que

$$\hat{U}_{f_s} : |x\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \longrightarrow (-1)^{f_s(x)}|x\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \tag{4.28}$$

en donde, si se ignora el segundo registro[28],

$$\hat{U}_s : |x\rangle \longrightarrow (-1)^{f_s(x)}|x\rangle, \tag{4.29}$$

o, equivalentemente,

$$\hat{U}_s = \hat{I} - 2|s\rangle\langle s|. \tag{4.30}$$

De este modo, y para aprovechar el paralelismo cuántico, con tal de preparar el primer registro en un estado

$$|x\rangle = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} |t\rangle = \mathrm{sen}(\theta)|\mathrm{s}\rangle + \frac{\cos(\theta)}{\sqrt{\mathrm{N}-1}} \sum_{\mathrm{t}\neq\mathrm{s}}^{\mathrm{N}-1} |\mathrm{t}\rangle, \tag{4.31}$$

donde inicialmente $\mathrm{sen}(\theta) = 1/\sqrt{\mathrm{N}}$, se puede comprobar que al aplicar la transformación unitaria

$$\hat{U}_{\mathrm{Grover}} = \hat{U}_x \hat{U}_s, \tag{4.32}$$

con $\hat{U}_x = 2|x\rangle\langle x| - \hat{I}$, el parámetro θ dobla su valor, es decir, se convierte en 2θ. Esto significa, por tanto, que con cada iteración que se realice, cada vez es más probable que una medida ortogonal en el registro ya proporcione el valor buscado (θ se aproxima a $\pi/2$). Ahora bien, hay que tener cuidado con no aplicar $\hat{U}_{\mathrm{grover}}$ en demasiadas ocasiones, ya que la probabilidad de éxito volvería a decrecer[29].

[28] Su inclusión se justifica para garantizar que $\hat{U}_{f_s}$ sea unitario.

[29] El número óptimo de iteraciones se puede encontrar en [26]. Un valor aproximado es: $(\pi/4)\sqrt{N}$.

Bibliografía

[1] L. Abellanas y A. Galindo. "Espacios de Hilbert". Eudema, Madrid, segunda edición (1991).

[2] P. Albertini. *Lett. Math. Phys.* **7**, 25 (1983).

[3] P. Albertini y A. Uhlmann. *Lett. Math. Phys.* **7**, 107 (1983).

[4] A. Aspect. Testing Bell's inequalities. *Europhys. News* **22**, 73–75 (1991).

[5] A. Aspect, J. Dalibard y G.Roger. Experimental test of Bell's inequalities using time-varying analyzers. *Phys. Rev. Lett.* **49**, 1804–1807 (1982).

[6] A. Aspect, P. Grangier y G. Roger. Experimental tests of realistic local theories via Bell's theorem. *Phys. Rev. Lett.* **47**, 460–463 (1981).

[7] A. Barenco. A universal two-bit gate for quantum computation. *Proc. Roy. Soc. Lond.* **A 449**, 679–683 (1995).

[8] A. Barenco, A. Ekert, K. A. Suominen y P. Torma. Approximate quantum Fourier transform and decoherence. *Phys. Rev.* **A 54**, 139–146 (1996).

[9] A. Barenco y A. K. Ekert. Dense coding based on quantum entanglement. *Journal of Modern Optics* **42**, 1253–1259 (1995).

[10] D. Beckman, A. Chari, S. Devabhaktuni y J. Preskill. Efficient networks for quantum factoring. *Phys. Rev.* **A 54**, 1034–1063 (1996).

[11] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Phys.* **1**, 195–200 (1964).

[12] J. S. Bell. On the problem of hidden variables in quantum mechanics. *Rev. Mod. Phys.* **38**, 447–452 (1966).

[13] J. S. Bell. "Speakable and Unspeakable in Quantum Mechanics". Cambridge Univ. Press, Cambridge, U.K (1987).

[14] P. Benioff. The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines. *J. Stat. Phys.* **22**, 563–591 (1980).

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

[16] P. Benioff. Quantum mechanical models of Turing machines that dissipate no energy. *Phys. Rev. Lett.* **48**, 1581–1585 (1982).

[17] C. H. Bennett. Quantum information and computation. *Phys. Today* **48**(10), 24–30 (1995).

[18] C. H. Bennett, E. Bernstein, G. Brassard y U. Vazirani. Strengths and weaknesses of quantum computing. (preprint quant-ph/9701001) (1997).

[19] C. H. Bennett y Brassard. Quantum public key distribution system. *IBM Techn. Disclosure Bull.* **28**, 3153–3163 (1985).

[20] C. H. Bennett y G. Brassard. Quantum cryptography: Public key distribution and coin tossing. En "Proc. IEEE Conf. on Computers, Syst. and Signal Process.", páginas 175–179 (1984).

[21] C. H. Bennett, G. Brassard, S. Briedbart y S. Wiesner. Quantum cryptography, or unforgeable subway tokens. En "Advances en Cryptography: Proceedings of Crypto '82", páginas 267–275, New York (1982). Plenum.

[22] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres y W. K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.* **70**, 1895–1898 (1993).

[23] C. H. Bennett, G. Brassard y A. K. Ekert. Quantum cryptography. *Scientif. Amer.* páginas 23–33 (Octubre 1992).

[24] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin y W. K. Wootters. Purification of noisy entanglement and faithful teleportation via noisy channel. *Phys. Rev. Lett.* **76**, 722–725 (1996).

[25] C. H. Bennett y S. J. Wiesner. Communication via one and two-particle operations on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.* **69**, 2881–2884 (1992).

[26] M. Boyer, G. Brassard, P. Hoyer y A. Tapp. Tight bounds on quantum serching. (preprint quant-ph/9605034) (1996).

[27] S. L. Braunstein y A. Mann. Measurement of the Bell operator and quantum teleportation. *Phys. Rev.* **A 51**, R1727–R1730 (1995).

[28] C. Cohen-Tannoudji, Bernard Diu y Franck Laloë. "Quantum Mechanics". Hermann, París, segunda edición (1977).

[29] D. Coppersmith. An approximate Fourier transform useful in quantum factoring. *IBM Research Report RC 19642* (1994).

[30] T. Cover y J. Thomas. "Elements of Information Theory". Wiley, New York (1991).

[31] D. Deutsch. Quantum theory, the Church-Turing principle and the universal quantum computer. *Proc. Roy. Soc. Lond.* **A 400**, 97–117 (1985).

[32] D. Deutsch. Quantum computational networks. *Proc. Roy. Soc. Lond.* **A 425**, 73–90 (1989).

[33] D. Deutsch, A. Barenco y A. Ekert. Universality in quantum computation. *Proc. Roy. Soc. Lond.* **A 449**, 669–677 (1995).

[34] D. DiVincenzo. Two-bit gates are universal for quantum computation. *Phys. Rev.* **A 51**, 1015–1022 (1995).

[35] A. Einstein, B. Podolsky y N. Rosen. Can quantum mechanical description of physical reality be considered complete? *Phys. Rev.* **47**, 777–780 (1935).

[36] A. Ekert y R. Jozsa. Quantum computation and Shor's factoring algorithm. *Rev. Mod. Phys.* **68**, 733 (1996).

[37] C. H. Bennett et al. Experimental quantum cryptography. *J. Cryptol.* **5**, 3–28 (1992).

[38] R. P. Feynman. Simulating physics with computers. *Int. J. Theor. Phys.* **21**, 467–488 (1982).

[39] R. P. Feynman. Quantum mechanical computers. *Opt. News* **11**, 11 (1985).

[40] L. K. Grover. Quantum mechanics helps in serching for a needle in a haystack. *Phys. Rev. Lett.* **79**, 325–328 (1997).

[41] G. H. Hardy y E. M. Wright. "An Introduction to the Theory of Numbers". Clarendon Press, Oxford (1979).

[42] P. Hausladen, R. Jozsa, B. Schumacher y M. Westmoreland y W. K. Wootters. Classical information capacity of a quantum channel. *Phys. Rev.* **54**, 1869–1876 (1996).

[43] M. E. Hellman. The mathematics of public-key cryptography. *Scientif. Amer.* **241**, 130–139 (Agosto 1979).

[44] C. Jack. Sherlock Holmes investigates the EPR paradox. *Phys. World* **8**, 39–42 (1995).

[45] R. Jozsa. Fidelity for mixed quantum states. *Journal of Modern Optics* **41**(12), 2315–2323 (1994).

[46] R. Jozsa y B. Schumacher. A new proof of the quantum noiseless coding theorem. *Journal of Modern Optics* **41**(12), 2343–2349 (1994).

[47] S. Lloyd. Almost any quantum logic gate is universal. *Phys. Rev. Lett.* **75**, 346–349 (1995).

[48] K. Mattle, H. Weinfurter, P. G. Kwiat y A. Zeilinger. Dense coding in experimental quantum communication. *Phys. Rev. Lett.* **76**, 4656–4659 (1996).

[49] C. Miquel, J. P. Paz y R. Perazzo. Factoring in a dissipative quantum computer. *Phys. Rev.* **A 54**, 2605–2613 (1996).

[50] Asher Peres. "Quantum Theory: Concepts and Methods". Kluwer Academic Publishers (1993).

[51] J. Preskill. Quantum information and computation. http://www.theory.caltech.edu/∼preskill/ph229 (1998).

[52] R. Rivest, A. Shamir y L. Adleman. On digital signatures and public key cryptosystems. Tech. rep. mit/lcs/tr-212, MIT Lab. Computer. Sci. (Enero 1979).

[53] B. Schumacher. Sending quantum entanglement through noisy channels. (preprint quant-ph/9604023) (1996).

[54] B. Schumacher y M. Westmoreland. Sending classical information via noise quantum channels. *Phys. Rev.* **56**, 131–138 (1997).

[55] P. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *IEEE Computer Society Press* (1994). Proc. of the 35th Annual Symp. on Foundations of Computer Science.

[56] A. M. Turing. On computable numbers, with an application to the Entschneidungsproblem. *Proc. Lond. Math. Soc. Ser.* **42**, 230 (1936).

[57] A. Uhlmann. *Rep. Math. Phys.* **9**, 273 (1976).

[58] A. Uhlmann. *Rep. Math. Phys.* **24**, 229 (1986).

[59] V. Vedral, A. Barenco y A. Ekert. Quantum networks for elementary arithmetic operations. *Phys. Rev.* **A 54**, 147–153 (1996).

[60] A. Wehrl. General properties of entropy. *Rev. Mod. Phys.* **50**(2), 221–260 (Abril 1978).

[61] H. Weinfurter. Experimental Bell-state analysis. *Europhys. Lett.* **25**, 559–564 (1994).

[62] S. Wiesner. Conjugate coding. *Sigact News* **28**, 78–88 (1983).