



UNIVERSIDAD CARLOS III DE MADRID

TESIS DOCTORAL

Medidas de información, incertidumbre y entrelazamiento en Mecánica Cuántica

Autor:

Julio I. de Vicente Majúa

Director:

Jorge Sánchez Ruiz

DEPARTAMENTO DE MATEMÁTICAS

Leganés, abril de 2008

TESIS DOCTORAL

MEDIDAS DE INFORMACIÓN, INCERTIDUMBRE Y ENTRELAZAMIENTO EN MECÁNICA CUÁNTICA

Autor: Julio I. de Vicente Majúa

Director: Jorge Sánchez Ruiz

Firma del Tribunal Calificador:

Firma

Presidente:

Vocal:

Vocal:

Vocal:

Secretario:

Calificación:

Leganés, de de

Agradecimientos

El principal y más merecido agradecimiento se lo debo a mi director de tesis, Jorge Sánchez Ruiz, por haberme introducido y guiado con ilusión en el mundo de la investigación. Tanto en las ocasiones en las que hemos trabajado codo con codo como en los momentos en los que ha tenido a bien darme libertad de acción siempre he contado con total confianza, ánimo y valoración de su parte. En segundo lugar, quisiera expresar mi agradecimiento a mis compañeros del grupo de Análisis Aplicado, en especial a sus responsables Francisco Marcellán y Guillermo López Lagomasino por haberme permitido formar parte de él y de los proyectos de investigación. Sin duda, he aprendido mucho gracias a todos ellos. Debo extender este agradecimiento al resto de miembros del departamento de Matemáticas, en particular, por haber tenido acceso a la beca de formación de doctores y al contrato de profesor ayudante específico. No quisiera dejar de mencionar a mis compañeros en el doctorado a lo largo de estos años por el buen ambiente que ha habido siempre entre nosotros. Así mismo, es un placer recordar a los diversos colegas con los cuales, bien sea personalmente o por correspondencia, he podido intercambiar intereses científicos. En este sentido quiero destacar especialmente a Otfried Gühne.

Quisiera expresar mi más profundo agradecimiento a mi madre por su constante cariño, confianza y apoyo, a la cual quiero dedicar esta tesis ya que es la persona en el mundo que recibe con más ilusión mis infinitesimales contribuciones al desarrollo científico. Por último, deseo mencionar a Elsa, la única persona con la que me olvido completamente del contenido de esta tesis y demás pájaros que pululan por mi cabeza.

Índice general

I	Introducción	1
1.	Medidas de incertidumbre	3
1.1.	Mayorización	4
1.2.	Entropía de información de Shannon	6
1.2.1.	Derivación axiomática	7
1.2.2.	Propiedades	9
1.2.3.	El caso continuo	10
1.3.	Otras medidas	11
1.3.1.	Entropías generalizadas	11
1.3.2.	Desviación estándar	12
2.	Mecánica Cuántica. Principio de incertidumbre y entrelazamiento	15
2.1.	Conceptos básicos de la Mecánica Cuántica	15
2.1.1.	Estados	15
2.1.2.	Estados compuestos	17
2.1.3.	Medidas	18
2.1.4.	Evolución	21
2.1.5.	Los postulados de la Mecánica Cuántica	21
2.2.	Principio de incertidumbre y relaciones de incertidumbre	22
2.2.1.	Formulación estándar. Relación de incertidumbre de Heisenberg-Robertson	24
2.2.2.	Crítica a la formulación estándar. Formulación entrópica	25
2.3.	Incetidumbre en sistemas atómicos y moleculares	33
2.3.1.	Entropía de polinomios ortogonales	35
2.4.	Entrelazamiento	36
2.4.1.	El problema de la separabilidad. Antecedentes	39
2.4.2.	Separabilidad y relaciones de incertidumbre	45
2.4.3.	Medidas de entrelazamiento	48
II	Incetidumbre	55
3.	Entropía de polinomios ortogonales	57

3.1. Entropía de polinomios de Gegenbauer de parámetro entero . . .	61
3.1.1. Representaciones trigonométricas para los polinomios de Gegenbauer	61
3.1.2. Evaluación de la integral entrópica	63
3.1.3. Resultados	72
3.2. Entropía de polinomios de Jacobi de parámetros semienteros . .	74
3.2.1. Representaciones trigonométricas para los polinomios de Jacobi	77
3.2.2. Evaluación de la integral entrópica. Resultados	80
4. Cotas más fuertes para una EUR en dimensión finita	87
III Entrelazamiento	95
5. Condiciones de separabilidad a partir de la relación de incertidumbre de Landau-Pollak	97
5.1. Condiciones de separabilidad para sistemas de dos qubits	98
5.1.1. Eficacia de las condiciones de separabilidad	102
5.2. Condiciones de separabilidad para sistemas de mayor dimensionalidad	106
6. Entrelazamiento y medidas de la correlación: Criterios de separabilidad a partir de la representación de Bloch	109
6.1. La representación de Bloch	111
6.2. Condiciones de separabilidad a partir de la representación de Bloch	113
6.3. Eficacia de los nuevos criterios	120
6.3.1. Ejemplos	120
6.3.2. Comparación con otros criterios importantes	123
7. Estimación de medidas de entrelazamiento a partir de criterios de separabilidad	131
7.1. Cuantificación del entrelazamiento y matriz de correlación	133
7.2. Cuantificación del entrelazamiento y relaciones de incertidumbre locales	140
IV Conclusiones y perspectivas	147
A. Descomposición en valores singulares y algunas normas matriciales	155
B. Una prueba alternativa de la identidad (3.86)	157

Prefacio

La Mecánica Cuántica se originó en 1900 para explicar la radiación del cuerpo negro y se desarrolló a lo largo del primer tercio del siglo XX con el objetivo de describir la realidad física en la escala atómica, para la cual se comprobó que las leyes de la física clásica dejaban de ser válidas. Desde entonces la teoría cuántica ha mostrado un rotundo éxito experimental en la descripción del microcosmos, que ha dado lugar a un sinfín de avances desde la perspectiva teórica hasta la tecnológica. La particular visión de la naturaleza que aporta esta teoría ha generado una revolución conceptual que sigue dando trabajo a los científicos y filósofos de hoy día, lo que, lejos de ser un obstáculo, ha constituido un caldo de cultivo idóneo para la propuesta de nuevas y aún más sorprendentes aplicaciones de los principios cuánticos. Esta efervescencia de ideas ha alcanzado momentos de gran esplendor en los últimos años dando lugar a la Teoría de la Información Cuántica, un emergente campo pluridisciplinar que utiliza las particularidades de la teoría cuántica para desarrollar formas revolucionarias para el procesado y transmisión de la información. Por otro lado, la Mecánica Cuántica tiene una estructura matemática de gran riqueza en la que tienen cabida la práctica totalidad de las ramas de las matemáticas. Por todo ello, creemos que esta teoría ofrece un marco de trabajo excepcional para el investigador en matemáticas con una orientación aplicada, tanto al nivel de la física teórica como al tecnológico.

En esta tesis se consideran problemas matemáticos relacionados con dos de los fenómenos más característicos y significativos de la Mecánica Cuántica: el principio de incertidumbre de Heisenberg y el entrelazamiento. El hilo conductor lo constituyen las medidas matemáticas de la incertidumbre asociada a distribuciones de probabilidad, que por un lado son indispensables para formular matemáticamente este principio y cuantificar la incertidumbre asociada a los distintos sistemas cuánticos y por otro lado proporcionan una herramienta útil para la caracterización del escurridizo fenómeno del entrelazamiento, paradigma central tanto dentro de los fundamentos de la Mecánica Cuántica como en las sorprendentes aplicaciones de la Teoría de la Información Cuántica. En esta memoria los dos primeros capítulos se han escrito a modo de introducción y los siguientes contienen nuestras aportaciones originales a estos problemas. En el primer capítulo presentamos cómo se construyen de una forma rigurosa desde un punto de vista matemático medidas de la incertidumbre y mostramos cómo la entropía de Shannon ocupa un lugar preferente entre ellas. En el segundo, después de introducir los postulados y la estructura matemática básica

de la Mecánica Cuántica, analizamos en detalle el principio de incertidumbre y repasamos las distintas relaciones de incertidumbre desarrolladas hasta la fecha, así como las argumentaciones que hacen a las relaciones de incertidumbre entrópicas la formulación más rigurosa del principio. En este capítulo se analizará también el fenómeno del entrelazamiento, su significado y relevancia. Aunque el problema de su caracterización sigue abierto (problema de la separabilidad), se presentarán los criterios previos más relevantes para su identificación y se mostrará que las relaciones de incertidumbre juegan un papel útil en esta tarea. Por último, se dará una brevísima introducción a la teoría de las medidas de entrelazamiento y su conexión con las medidas de incertidumbre. En el capítulo 3 se abordará el problema del cálculo de entropías de polinomios ortogonales, directamente relacionado con la expresión de la incertidumbre en sistemas atómicos y moleculares; se utilizarán principalmente técnicas analíticas, en particular de la teoría de funciones especiales. El capítulo 4 se relaciona con la derivación de desigualdades óptimas para la formulación de relaciones de incertidumbre entrópicas, apoyándonos sobre todo en técnicas de optimización. A continuación nos centraremos en el entrelazamiento y el problema de la separabilidad, para lo que principalmente utilizaremos álgebra lineal y análisis matricial. En el capítulo 5 se estudiará este problema utilizando una formulación particular del principio de incertidumbre. En el capítulo 6, motivados por la estructura interna de las condiciones de separabilidad expresadas a través de relaciones de incertidumbre, obtendremos condiciones para la presencia de entrelazamiento a través de medidas de la correlación. Por último, en el capítulo 7 estudiaremos cómo utilizar las condiciones del apartado anterior y las basadas en relaciones de incertidumbre para poder obtener información no sólo cualitativa sino también cuantitativa acerca del entrelazamiento.

La mayor parte de los resultados que se recogen en esta tesis han sido previamente publicados en distintas revistas científicas internacionales, tal y como señalamos a continuación (el número entre corchetes al final de cada referencia indica el capítulo de la tesis con el que se corresponde):

- J.I. de Vicente y J. Sánchez-Ruiz, *Separability conditions from the Landau-Pollak uncertainty relation*, Phys. Rev. A **71**, 052325 (2005). [5]
- J.I. de Vicente, S. Gandy y J. Sánchez-Ruiz, *Information entropy of Gegenbauer polynomials*, p. 579 en *Difference Equations, Special Functions and Orthogonal Polynomials: Proceedings of the International Conference*, eds. S. Elaydi, J. Cushing, R. Lasser, V. Papageorgiou, A. Ruffing y W. Van Assche (World Scientific, Singapur, 2007). [3]
- J.I. de Vicente, *Separability criteria based on the Bloch representation of density matrices*, Quantum Inf. Comput. **7**, 624 (2007). [6]
- J.I. de Vicente, *Lower bounds on concurrence and separability conditions*, Phys. Rev. A **75**, 052320 (2007). [7]
- J.I. de Vicente, S. Gandy y J. Sánchez-Ruiz, *Information entropy of*

Gegenbauer polynomials of integer parameter, J. Phys. A: Math. Theor. **40**, 8345 (2007). [3]

- J.I. de Vicente, *Further results on entanglement detection and quantification from the correlation matrix criterion*, J. Phys A: Math. Theor. **41**, 065309 (2008). [6,7]
- J.I. de Vicente y J. Sánchez-Ruiz, *Improved bounds on entropic uncertainty relations*, Phys. Rev. A **77**, 042110 (2008). [4]

Parte I

Introducción

Capítulo 1

Medidas de incertidumbre

Cuando nos enfrentamos a un proceso aleatorio no somos capaces de predecir con certeza el resultado del mismo. No obstante, esto no significa que nos encontremos en una situación de completa ignorancia y no podamos afirmar nada acerca del proceso. Podemos asociar a éste una distribución de probabilidad en la que queda codificado todo nuestro conocimiento y que podemos utilizar para inferir propiedades del sistema. Por ejemplo, consideremos un juego en el que se dispone de una urna que contiene bolas de distintos colores y en el que ganamos si adivinamos el color de la bola que extraerá una “mano inocente”. A menos que todas las bolas sean del mismo color no podemos afirmar a priori con completa certidumbre qué color tendrá la bola extraída. Sin embargo, según sea la situación podemos encontrarnos con distintos grados de predictabilidad. Sea el caso A en el que la urna contiene 90 bolas negras y 10 blancas y el caso B en el que la urna contiene 50 bolas negras y 50 bolas blancas. Cualquier persona dotada con un mínimo de sentido común elegiría jugar en la modalidad A puesto que en este caso es más probable la extracción de la bola negra y, por tanto, el resultado es más predecible y más probable la victoria. Por contra, en el caso B tenemos una completa incertidumbre acerca del resultado de la extracción. Este trivial ejemplo pone de manifiesto que existen procesos aleatorios con mayor grado de incertidumbre (es decir, en los que disponemos de menos información acerca del posible resultado) que otros. Consideremos ahora procesos que vengan dados por las siguientes distribuciones de probabilidad: $(0.6, 0.2, 0.2)$ y $(0.5, 0.4, 0.1)$ (60 bolas negras, 20 bolas blancas y 20 bolas rojas y 50 bolas negras, 40 bolas blancas y 10 bolas rojas respectivamente). ¿Acerca de cuál tenemos ahora más información? En el primer caso la extracción de un color (el negro) es bastante más probable que el resto, mientras que en el segundo caso la mayoría de la probabilidad se reparte entre el negro y el blanco dejando al rojo como suceso bastante improbable. Así pues, uno elegiría jugar de acuerdo al primer caso puesto que la extracción del color negro tiene más predictabilidad; sin embargo, si se alteraran las reglas del juego y la victoria se obtuviera adivinando algún color que no va a salir, uno elegiría jugar en la segunda modalidad. No queda, pues, claro para qué distribución

de probabilidad se tiene un menor grado de incertidumbre acerca del resultado del proceso aleatorio. Este sencillo ejemplo muestra que la mera intuición no sirve para responder a este tipo de preguntas; es más, el problema se complica si, más allá del nivel cualitativo, buscamos una forma de cuantificar nuestro conocimiento acerca del resultado de este tipo de procesos. En este capítulo se analiza el problema de encontrar expresiones matemáticas que midan con rigor la incertidumbre asociada a un proceso aleatorio. En él se dará cuenta de herramientas que, partiendo de una sólida base matemática, no sólo permiten afirmar si una distribución de probabilidad tiene asociada más ignorancia que otra, sino que cuantifican el grado de incertidumbre de cada una.

1.1. Mayorización

Una primera base sobre la que construir medidas de incertidumbre viene dada por el principio de transferencia, que fue desarrollado a principios del siglo XX en economía para el estudio de la distribución de la riqueza (véase la introducción de [1] y las referencias allí citadas). Traducido a nuestro contexto este principio se puede entender de la siguiente manera: Dada una distribución de probabilidad, si transferimos parte de la probabilidad de un suceso más probable a otro menos probable, de forma que la cantidad transferida es menor a la mitad de la diferencia de probabilidad entre estos dos sucesos, entonces la distribución de probabilidad obtenida está menos concentrada que la original y tiene asociada, por tanto, un mayor grado de incertidumbre. Consecuentemente, si una distribución de probabilidad puede ser obtenida a partir de otra por medio de una serie de transferencias como la anteriormente señalada, entonces concluiremos que la primera tiene asociada mayor incertidumbre que la segunda. Por ejemplo, si consideramos las distribuciones de probabilidad $(0.6, 0.2, 0.2)$ y $(0.4, 0.3, 0.3)$, la segunda puede ser obtenida a partir de la primera si transferimos 0.1 de la probabilidad del primer suceso al segundo y otro tanto al tercer suceso. Así, podemos afirmar que existe un mayor grado de incertidumbre en la segunda distribución.

Sean $\{p_1, p_2\}$ ($p_1 \geq p_2$) las dos probabilidades entre las que se va a establecer un proceso de transferencia para obtener $\{q_1, q_2\}$ ($p_1 \geq q_i \geq p_2$ para $i = 1, 2$). Esto significa que $q_1 = \lambda p_1 + (1 - \lambda)p_2$ con $\lambda \in [0, 1]$ y, puesto que $p_1 + p_2 = q_1 + q_2$, $q_2 = \lambda p_2 + (1 - \lambda)p_1$ ¹. O sea,

$$\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \lambda \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + (1 - \lambda) \begin{pmatrix} p_2 \\ p_1 \end{pmatrix}, \quad \lambda \in [0, 1]. \quad (1.1)$$

Como consecuencia se tiene que una distribución de probabilidad obtenida a partir de otra mediante una serie de procesos de transferencia se puede escribir como una combinación convexa de permutaciones de ésta.

¹Nótese la arbitrariedad en el etiquetado de la segunda distribución. Si $\lambda \in (1/2, 1]$, entonces $q_1 > q_2$, lo que significa que p_1 se convierte en q_1 en el proceso de transferencia. En caso contrario p_1 se transforma en q_2 .

Definición 1.1. Sean p y q los vectores n -dimensionales cuyas entradas contienen las distribuciones de probabilidad $\mathcal{P} = \{p_1, p_2, \dots, p_n\}$ y $\mathcal{Q} = \{q_1, q_2, \dots, q_n\}$. Si existen matrices de permutación² $\{\Pi_i\}$ de dimensión n tales que

$$q = \sum_i \lambda_i \Pi_i p, \quad (1.2)$$

donde $\lambda_i \geq 0 \forall i$ y $\sum_i \lambda_i = 1$, entonces \mathcal{Q} se puede obtener a partir de \mathcal{P} mediante una serie de procesos de transferencia y decimos que \mathcal{P} mayoriza a \mathcal{Q} , lo que denotamos como $\mathcal{Q} \prec \mathcal{P}$.

La mayorización captura, por tanto, el concepto del principio de transferencia y nos equipa con una herramienta cualitativa para comparar el grado de incertidumbre correspondiente a dos distribuciones de probabilidad. Lamentablemente, dadas dos distribuciones de probabilidad resulta muy complicado determinar si se verifica (1.2), por lo que resultarán útiles las caracterizaciones que damos a continuación.

Definición 1.2. Una matriz cuadrada se dice doblemente estocástica si todas sus entradas son no negativas y si la suma de todos los elementos de cualquiera de sus filas y de cualquiera de sus columnas es 1.

Teorema 1.1. $\mathcal{Q} \prec \mathcal{P}$ si y sólo si existe alguna matriz doblemente estocástica D tal que $q = Dp$.

Este teorema es una consecuencia directa del teorema de Birkhoff [2], que afirma que el conjunto de las matrices doblemente estocásticas constituye la envoltura convexa de las matrices de permutación. Las siguientes caracterizaciones se deben a Hardy, Littlewood y Pólya [3]:

Teorema 1.2. $\mathcal{Q} \prec \mathcal{P}$ si y sólo si $\sum_i \phi(q_i) \leq \sum_i \phi(p_i)$ para todas las funciones continuas convexas ϕ .

Teorema 1.3. Sean p y q ordenados en orden no creciente (i. e. $p_1 \geq p_2 \geq \dots \geq p_n$ y $q_1 \geq q_2 \geq \dots \geq q_n$). Entonces $\mathcal{Q} \prec \mathcal{P}$ si y sólo si

$$\sum_{i=1}^k p_i \geq \sum_{i=1}^k q_i \quad (1.3)$$

para $k = 1, 2, \dots, n$.

Este último resultado proporciona la forma más eficiente de verificar si una distribución de probabilidad mayoriza a otra. Además, nos permite dar cuenta de ejemplos sencillos que muestran la naturalidad del concepto de mayorización. Así, para toda distribución de probabilidad \mathcal{P} de n sucesos se tiene que

$$(1/n, \dots, 1/n) \prec \mathcal{P} \prec (1, 0, \dots, 0), \quad (1.4)$$

²El conjunto de las matrices de permutación está formado por la identidad y todas las posibles permutaciones de sus filas.

lo cual indica que, como cabía esperar, no existe distribución de probabilidad con más certeza que aquella en la que un suceso es seguro, y que la distribución uniforme es la de máxima ignorancia puesto que es mayorizada por todas las distribuciones de probabilidad. También se puede demostrar con facilidad a partir de (1.3) que

$$(1/n, \dots, 1/n) \prec (1/(n-1), \dots, 1/(n-1), 0) \prec \dots \prec (1/2, 1/2, 0, \dots, 0), \quad (1.5)$$

lo que representa que en la distribución uniforme la incertidumbre crece con el número de sucesos posibles.

El Teorema 1.3 también permite analizar con sencillez la mayorización como relación binaria en el conjunto de las distribuciones de probabilidad. En efecto, de él se deduce fácilmente que $\mathcal{P} \prec \mathcal{P}$ para toda distribución de probabilidad (reflexividad) y que $\mathcal{Q} \prec \mathcal{P}$ y $\mathcal{P} \prec \mathcal{R}$ implica $\mathcal{Q} \prec \mathcal{R}$ (transitividad), lo que significa que la mayorización es una relación de preorden. En cambio, no se puede decir que sea un orden parcial porque en rigor no se cumple la antisimetría. Es decir, $\mathcal{Q} \prec \mathcal{P}$ y $\mathcal{P} \prec \mathcal{Q}$ no implica $\mathcal{Q} = \mathcal{P}$ sino que existe una matriz de permutación Π tal que $q = \Pi p$, como se puede deducir de (1.2) teniendo en cuenta que la inversa de una matriz de permutación es ella misma. Esto es consecuencia directa de que desde el punto de vista de la incertidumbre el etiquetado de las probabilidades (es decir, el suceso que representa cada una) es indiferente. Más interesante resulta comprobar que la mayorización tampoco es un orden total ya que, dadas distribuciones de probabilidad cualesquiera \mathcal{P} y \mathcal{Q} , no tiene por qué satisfacerse que o bien $\mathcal{Q} \prec \mathcal{P}$ o bien $\mathcal{P} \prec \mathcal{Q}$, lo cual es fácilmente deducible del Teorema 1.3. Esta falta de totalidad implica que, pese a haber desarrollado una herramienta matemática rigurosa para estimar cualitativamente el grado de ignorancia asociado a un proceso aleatorio, no podemos aplicar la mayorización en todas las circunstancias. De hecho, si volvemos al ejemplo inicial en el que comparábamos las distribuciones $(0.6, 0.2, 0.2)$ y $(0.5, 0.4, 0.1)$, aplicando (1.3) llegamos a la conclusión de que ninguna mayoriza a la otra y seguimos sin poder responder a la pregunta de qué distribución tiene asociado un menor grado de incertidumbre. Este aspecto negativo de la mayorización pone de manifiesto, aún más si cabe, la necesidad de una medida cuantitativa de la incertidumbre, lo cual trataremos en las siguientes secciones. El lector interesado en más detalles acerca de la teoría de la mayorización es referido a [1].

1.2. Entropía de información de Shannon

A pesar de las limitaciones de la mayorización que acabamos de señalar, esta herramienta proporciona una de las primeras nociones a tener en cuenta a la hora de considerar una medida cuantitativa de la incertidumbre: la Schur-concavidad.

Definición 1.3. Una función real $f(\mathcal{P}) = f(p_1, \dots, p_n)$ se dice Schur-cóncava si

$$\mathcal{Q} \prec \mathcal{P} \implies f(\mathcal{Q}) \geq f(\mathcal{P}) \quad (1.6)$$

Efectivamente, puesto que $\mathcal{Q} \prec \mathcal{P}$ implica que hay una mayor ignorancia en el resultado de \mathcal{Q} que en el de \mathcal{P} , una medida de incertidumbre rigurosa debe asociar un valor mayor a la primera. Sin embargo, la sola imposición de la Schur-concavidad resulta ser una condición demasiado débil para definir una medida cuantitativa única de la incertidumbre. Basta considerar que, de acuerdo con el Teorema 1.2, $f(\mathcal{P}) = \sum_i \phi(p_i)$ es Schur-cóncava para cualquier función ϕ cóncava. Lógicamente, esto permite construir funciones Schur-cóncavas con características muy diversas (véase el capítulo 13.F de [1]). Es más, al no ser la mayorización un orden total, para los casos en los que ni $\mathcal{Q} \prec \mathcal{P}$ ni $\mathcal{P} \prec \mathcal{Q}$ el simple postulado de la Schur-concavidad deja una completa arbitrariedad en el orden que una medida de esta forma induce entre \mathcal{P} y \mathcal{Q} . Retomaremos estas ideas en la siguiente sección cuando hablemos de entropías generalizadas.

1.2.1. Derivación axiomática

La incapacidad de la Schur-concavidad como único ingrediente para construir una medida cuantitativa de la incertidumbre nos obliga a replantearnos el problema desde sus cimientos. ¿Qué axiomas debe satisfacer una medida rigurosa de la incertidumbre para que quede unívocamente definida? Este problema fue planteado y resuelto elegantemente por C. Shannon en 1948 a partir de tres sencillos y muy razonables axiomas [4] que mostramos a continuación. Sea un proceso aleatorio cuyo conjunto de sucesos posibles viene caracterizado por la distribución de probabilidad $\mathcal{P} = \{p_1, \dots, p_n\}$ y sea $H(\mathcal{P}) = H(p_1, \dots, p_n)$ ($H : \mathbb{R}^n \rightarrow \mathbb{R}$) una medida cuantitativa que exprese la incertidumbre acerca del resultado del proceso. Entonces:

1. H debe ser una función continua en todas sus variables.
2. $H(1/n, \dots, 1/n)$ es una función monótona creciente de n .
3. H debe satisfacer que

$$H(p_1, \dots, p_n) = H(p_1, \dots, p_{n-2}, p_{n-1} + p_n) + (p_{n-1} + p_n)H\left(\frac{p_{n-1}}{p_{n-1} + p_n}, \frac{p_n}{p_{n-1} + p_n}\right). \quad (1.7)$$

El primer axioma resulta bastante obvio, pues no sólo asegura que H tenga un buen comportamiento desde el punto de vista matemático sino que también implica que distribuciones de probabilidad próximas tengan un valor de H cercano. El segundo axioma establece que cuando todos los sucesos son equiprobables se tiene más incertidumbre si el número de sucesos posibles es mayor, noción que ya capturaba la mayorización (véase (1.5)). El tercer axioma puede entenderse fácilmente con un ejemplo. Consideremos un proceso aleatorio con distribución de probabilidad $(1/2, 1/3, 1/6)$. Este proceso puede dividirse en dos, en una primera instancia nos preguntamos si ha tenido lugar el primer suceso o si se ha dado uno de los otros dos posibles. El subproceso así construido tiene asociada la distribución de probabilidad $(1/2, 1/2)$. En caso de que se

diera la segunda opción consideramos entonces si se produce el segundo o tercer suceso, lo cual viene dado por la distribución de probabilidad $(2/3, 1/3)$. Puesto que la incertidumbre no puede depender del enfoque con que nos acerquemos al proceso aleatorio, se debe cumplir

$$H(1/2, 1/3, 1/6) = H(1/2, 1/2) + \frac{1}{2}H(2/3, 1/3),$$

donde el coeficiente $1/2$ del segundo término es un factor de peso debido a que la segunda opción (que se dé el segundo o tercer suceso) se produce con probabilidad $1/2$. Es decir, el tercer axioma establece que la cantidad de incertidumbre asociada a un proceso aleatorio es igual a la suma pesada de las incertidumbres asociadas a los subprocesos en los que éste se puede dividir.

Teorema 1.4. *La única medida que satisface los axiomas 1, 2 y 3 es de la forma*

$$H(\mathcal{P}) = -K \sum_{i=1}^n p_i \log p_i \quad (1.8)$$

donde K es una constante positiva.

La elección de la constante K simplemente fija una escala, por lo que, sin pérdida de generalidad, podemos considerar $K = 1$.

Definición 1.4. *La entropía de información de Shannon (o, simplemente, entropía) asociada a una distribución de probabilidad discreta $\mathcal{P} = \{p_1, p_2, \dots, p_n\}$ se define como³*

$$H(\mathcal{P}) = - \sum_{i=1}^n p_i \log p_i. \quad (1.9)$$

De acuerdo con el teorema de Shannon la entropía de información constituye la única medida cuantitativa rigurosa de la incertidumbre asociada a un proceso aleatorio regido por una determinada distribución de probabilidad. La base del logaritmo es arbitraria, siendo las elecciones más habituales 2 y e . La primera se suele dar en contextos donde las distribuciones de probabilidad binarias juegan un papel especialmente relevante. Como este no es nuestro caso y, además, la base natural hace más sencillas las manipulaciones matemáticas, en lo que sigue adoptaremos la segunda opción, usándose indistintamente \log y \ln para denotar el logaritmo neperiano. El término entropía se debe a la conexión de esta magnitud con la entropía de la termodinámica y la mecánica estadística. Como el propio Shannon señalaba H es, por ejemplo, la H del célebre teorema H de Boltzmann cuando \mathcal{P} representa la probabilidad de que un sistema de partículas se encuentre en cualquiera de sus microestados accesibles. El término información responde a una de las posibles interpretaciones de la entropía. Además de como una medida de incertidumbre (o falta de información) acerca de

³En el caso de que alguna probabilidad sea cero su contribución a la entropía se considera nula, lo que se justifica por continuidad dado que $x \log x \rightarrow 0$ cuando $x \rightarrow 0$.

un experimento aleatorio, H se puede entender como la cantidad de información que uno espera ganar en promedio al llevar a cabo el experimento. En este contexto es interesante señalar otra interpretación alternativa de la entropía como el número mínimo de preguntas binarias (i.e. cuya respuesta es sí o no) que es necesario hacer en promedio para determinar el resultado de un experimento aleatorio utilizando una estrategia óptima⁴, lo que se relaciona con la cantidad mínima de recursos físicos que se necesitan para almacenar la información producida por una fuente y que da lugar al teorema de codificación sin ruido (“noiseless coding”) de Shannon. De hecho, Shannon introdujo la entropía en el estudio de cómo transmitir y almacenar información. Ésta se mostró como una herramienta tan sumamente útil en este contexto⁵ que ha dado lugar a una rama propia de la investigación conocida como Teoría de la Información, siendo uno de sus primeros grandes resultados el teorema antes referido. Así, la entropía, además de constituir una medida rigurosa de la incertidumbre, que es el aspecto que nos interesa aquí, tiene multitud de aplicaciones en muy diversas ramas de la Ciencia.

Podemos volver ahora a considerar el ejemplo inicial en el que comparábamos la incertidumbre asociada a las distribuciones de probabilidad $(0.6, 0.2, 0.2)$ y $(0.5, 0.4, 0.1)$ y obtener una respuesta concreta. $H(0.6, 0.2, 0.2) = 0.9503$ y $H(0.5, 0.4, 0.1) = 0.9433$; por tanto, nuestra incertidumbre en el resultado del segundo proceso es ligeramente inferior.

1.2.2. Propiedades

En este apartado damos una lista de las propiedades matemáticas más relevantes de la entropía. Algunas se deducen más o menos directamente de resultados anteriormente citados. Para las demás, el lector interesado en la demostración es remitido a [5], referencia estándar actual en Teoría de la Información.

- $H(\mathcal{P})$ es Schur-cóncava. Este resultado se sigue del Teorema 1.2 y de la concavidad de la función $\phi(x) = -x \log x$. Como razonamos anteriormente, esta es una propiedad que debe exigirse a toda medida de la incertidumbre y no es sorprendente, por tanto, que sea satisfecha por la entropía.
- Como consecuencia directa de la propiedad anterior y de (1.4), se tiene que para toda distribución de probabilidad \mathcal{P} de n sucesos

$$H(1/n, \dots, 1/n) = \log n \geq H(\mathcal{P}) \geq 0 = H(1, 0, \dots, 0), \quad (1.10)$$

lo que particularmente implica la no negatividad de la entropía.

- $H(\mathcal{P})$ es una función cóncava de \mathcal{P} . Es decir,

$$H(\lambda\mathcal{P} + (1 - \lambda)\mathcal{Q}) \geq \lambda H(\mathcal{P}) + (1 - \lambda)H(\mathcal{Q}) \quad (1.11)$$

⁴Esto es así cuando el logaritmo se toma en base 2.

⁵De hecho, Shannon toma este éxito como una razón adicional para considerar acertados los tres axiomas de los que se deduce la entropía.

para $0 \leq \lambda \leq 1$. Para entender mejor esta propiedad, volvamos a considerar el juego de las bolas introducido al comienzo del capítulo en el que disponemos de dos urnas con un contenido diferente. Sea \mathcal{P} (\mathcal{Q}) la distribución de probabilidad correspondiente a extraer las bolas de la urna A (B). Supongamos que la urna con la que vamos a jugar viene determinada por un experimento aleatorio (extrayendo una bola etiquetada con A o B de la urna C, por ejemplo) en el que obtener A tiene probabilidad λ y obtener B tiene probabilidad $1 - \lambda$. Nuestra incertidumbre inicial viene dada por la entropía del miembro izquierdo de (1.11), mientras que el miembro derecho representa nuestra incertidumbre en promedio una vez conocido el resultado de la urna C. Así, (1.11) significa que nuestra incertidumbre después de llevar a cabo un experimento disminuye en promedio, o, lo que es lo mismo, se gana información en promedio al conocer el resultado de un experimento⁶.

- La entropía es subaditiva. Esto significa que la entropía conjunta asociada a dos procesos aleatorios es menor o igual que la suma de sus entropías,

$$H(\mathcal{P}, \mathcal{Q}) \leq H(\mathcal{P}) + H(\mathcal{Q}), \quad (1.12)$$

con igualdad (aditividad) si y sólo si \mathcal{P} y \mathcal{Q} son sucesos independientes, i.e. $\{\mathcal{P}, \mathcal{Q}\} = \{p_i, q_j\}$. Esta propiedad resulta de nuevo muy natural, pues nuestra incertidumbre conjunta acerca de procesos aleatorios que no tienen ninguna relación entre sí debe ser obviamente la suma de las incertidumbres individuales.

1.2.3. El caso continuo

Hasta ahora nos hemos limitado a considerar procesos aleatorios gobernados por una distribución de probabilidad discreta, pero esta no es la única posibilidad. En el caso de una variable aleatoria continua X soportada en $\Delta \subseteq \mathbb{R}$ con densidad de probabilidad dada por la función $p(x) \geq 0$ tal que

$$\int_{\Delta} p(x) dx = 1, \quad (1.13)$$

la entropía asociada a esta distribución de probabilidad viene dada por

$$H(X) = - \int_{\Delta} p(x) \log p(x) dx, \quad (1.14)$$

la cual a veces recibe el nombre de entropía diferencial. Su extensión a distribuciones de probabilidad multidimensionales es inmediata. La entropía diferencial fue propuesta también por Shannon en [4] como análogo natural de (1.9). Posteriormente, fue deducida a partir de un conjunto de axiomas

⁶Nótese que en un experimento individual nuestra incertidumbre puede incrementarse después de conocer el resultado de C (por ejemplo si $H(\mathcal{P}) \geq H(\lambda\mathcal{P} + (1 - \lambda)\mathcal{Q})$ y la urna C dictamina que juguemos con la A).

análogo al del caso discreto [6]. También se puede obtener discretizando una distribución continua mediante la división de Δ en intervalos, considerando la entropía discreta y tomando el límite de la misma cuando la anchura de los intervalos tiende a cero, si bien en este caso hay que renormalizar⁷ el resultado obtenido (véase la sección 9.3 en [5]).

La entropía diferencial hereda muchas de las propiedades de la discreta. Por ejemplo, se puede establecer la Schur-concavidad considerando el análogo a una transformación doblemente estocástica en el caso continuo. Una diferencia importante es que la entropía diferencial puede tomar valores negativos, lo cual no plantea ningún inconveniente. Se sigue manteniendo que a menor incertidumbre menor es la entropía, pero en este caso cuanto más concentrada sea la densidad de probabilidad (i.e. menos incertidumbre) más se acerca a $-\infty$ su entropía.

1.3. Otras medidas

1.3.1. Entropías generalizadas

Existen autores que han relajado la axiomática de Shannon para una medida de la incertidumbre, lo que ha dado lugar a un abanico más amplio de medidas que se conocen como entropías generalizadas. Básicamente, estas medidas explotan la libertad a la que da lugar la Schur-concavidad como única restricción para una medida de incertidumbre. Si bien las entropías generalizadas, como discutiremos en más detalle a continuación, no poseen todas las buenas propiedades matemáticas que caracterizan a la entropía de Shannon y que le otorgan un lugar central en Teoría de la Información, su interés y utilidad resultan incuestionables.

Definición 1.5. *Las entropías de Rényi de una distribución de probabilidad discreta $\mathcal{P} = \{p_1, p_2, \dots, p_n\}$ se definen como [7]*

$$H_q(\mathcal{P}) = \frac{1}{1-q} \log \left(\sum_{i=1}^n (p_i)^q \right) \quad (q \geq 0). \quad (1.15)$$

Básicamente, la axiomática de Rényi se puede entender como el resultado de sustituir el axioma 3 de Shannon por la condición más débil de aditividad. Alternativamente, estas entropías también se pueden derivar exigiendo la Schur-concavidad junto con alguna condición más de índole técnica [8]. Puesto que proviene de una relajación de los axiomas originales de Shannon y no de una axiomática diferente, la entropía de Rényi contiene como caso particular a la de Shannon; concretamente cuando $q = 1$, lo cual puede comprobarse tomando el correspondiente límite en (1.15). Otro caso particular destacable que será útil más adelante es

$$H_\infty(\mathcal{P}) = -\log \max_i p_i, \quad (1.16)$$

⁷En el sentido de la física de partículas, es decir, extraer partes divergentes.

que también se demuestra con el correspondiente límite.

Como función del parámetro q , $H_q(\mathcal{P})$ es no creciente, lo cual es útil para acotar la entropía de Shannon. Como ya hemos mencionado, las entropías de Rényi son siempre aditivas. Sin embargo, la subaditividad se da si y sólo si $q = 1$ (i.e. en el caso Shannon). Además, $H_q(\mathcal{P})$ es cóncava si $0 < q \leq 1$ (si $n = 2$ la concavidad también se cumple para $0 < q \leq 2$), mientras que para $q > 2$ $H_q(\mathcal{P})$ no es ni cóncava ni convexa [9]. Como rasgo positivo de estas entropías, cabe señalar que para $q > 1$ están libres de problemas de continuidad que sufre la entropía de Shannon cuando se permite un conjunto infinito de posibles sucesos [8, 10].

Definición 1.6. *Las entropías de Tsallis de una distribución de probabilidad discreta $\mathcal{P} = \{p_1, p_2, \dots, p_n\}$ se definen como [11]*

$$\tilde{H}_q(\mathcal{P}) = \frac{1}{1-q} \left(\sum_{i=1}^n (p_i)^q - 1 \right) \quad (q \geq 0). \quad (1.17)$$

Nótese que estas entropías también se reducen a la de Shannon en el caso $q = 1$. Si bien las entropías de Tsallis son cóncavas para todo q , no son ni subaditivas ni aditivas dejando aparte el caso especial $q = 1$. A pesar de que la falta de aditividad puede resultar sorprendente e inapropiada para una medida de la incertidumbre, este es el ingrediente básico de su axiomática [12] y la razón principal de su interés en termodinámica y estadística de sistemas complejos (véase por ejemplo [13]).

Análogamente al caso de Shannon, para distribuciones de probabilidad continuas, siguiendo la notación de la sección 1.2.3, las correspondientes entropías de Rényi y Tsallis se definen como

$$H_q(X) = \frac{1}{1-q} \log \left(\int_{\Delta} (p(x))^q dx \right), \quad (1.18)$$

$$\tilde{H}_q(X) = \frac{1}{1-q} \left(\int_{\Delta} (p(x))^q dx - 1 \right). \quad (1.19)$$

1.3.2. Desviación estándar

Como ya hemos señalado, el seminal trabajo de Shannon que dio lugar a la entropía data de 1948. La derivación de entropías generalizadas ha sido, lógicamente, posterior. Sin embargo, cuantificar la incertidumbre de procesos aleatorios es un problema fundamental con el que directa o indirectamente han tenido que tratar los científicos desde antes del siglo XX, por lo que, a falta de un edificio axiomático riguroso sobre el que apoyarse, con anterioridad al trabajo de Shannon ya se consideraban una serie de medidas de la incertidumbre construidas sobre la base de la mera intuición. Sin duda, la medida más habitual ha sido la desviación estándar (véase [8] para otras medidas de este tipo). Dada una variable aleatoria X , la desviación estándar ΔX se define como

$$\Delta X = \sqrt{\langle (X - \langle X \rangle)^2 \rangle} = \sqrt{\langle X^2 \rangle - \langle X \rangle^2}, \quad (1.20)$$

donde $\langle X \rangle$ es el valor esperado de X , i.e.

$$\langle X \rangle = \sum_{i=1}^n x_i p_i, \quad (1.21)$$

$$\langle X \rangle = \int_{\Delta} xp(x)dx, \quad (1.22)$$

según la distribución de probabilidad para los posibles valores de la variable aleatoria sea discreta o continua y siguiendo la notación de las secciones anteriores. La desviación estándar se anula para la distribución de probabilidad delta y va aumentando su valor según nos alejamos de esta distribución. En el caso de la distribución de probabilidad gaussiana la desviación estándar está directamente relacionada con el ancho de la campana. Consideraciones como estas son las que han dado un respaldo intuitivo al uso a lo largo de los años de la desviación estándar como medida de incertidumbre. Este uso ha podido ser sin lugar a dudas útil, especialmente antes de la introducción de la entropía, pero, como veremos a continuación, el considerar la desviación estándar como una medida de la incertidumbre no sólo puede ser arbitrario (puesto que, a diferencia de la entropía, no se deriva de una axiomática) sino también inadecuado. En primer lugar, el valor de ΔX depende no sólo de la correspondiente distribución de probabilidad sino también de los posibles valores que toma la variable aleatoria (x_i o x según el caso). Esta situación no resulta satisfactoria, pues dos variables aleatorias regidas por la misma distribución de probabilidad pero a las que cada suceso asocia diferentes valores tendrán diferente desviación estándar cuando claramente la incertidumbre es la misma. Como ya discutimos en otras secciones, una medida de incertidumbre debe ser función únicamente de la distribución de probabilidad. En segundo lugar, aunque, como ya hemos mencionado, la desviación estándar da una estimación razonable de la incertidumbre en el caso de distribuciones como la gaussiana, en otros casos esta estimación puede ser francamente mala. Un ejemplo clásico es el de la distribución de Cauchy, i.e.

$$p(x) = \frac{\gamma}{\pi((x-a)^2 + \gamma^2)}, \quad (1.23)$$

para la que ΔX diverge para todos los valores del parámetro γ a pesar de que $p(x)$ tiende a la distribución delta cuando $\gamma \rightarrow 0$. De hecho, los valores de la desviación estándar pueden diferir completamente de los de la entropía. Por ejemplo, sea $\Delta = [0, a]$. De acuerdo con la intuición, el máximo valor posible de $H(X)$ en este caso se da para la distribución de probabilidad uniforme

$$p(x) = \frac{1}{a}\chi_{\Delta} \quad (1.24)$$

donde χ es la función característica. Sin embargo, no es difícil demostrar que el máximo de ΔX se produce cuando la probabilidad se concentra en los extremos del intervalo, o sea, cuando

$$p(x) = \frac{1}{2}\delta(x) + \frac{1}{2}\delta(x-a). \quad (1.25)$$

No obstante, se puede establecer una relación entre la entropía y la desviación estándar para el caso de distribuciones de probabilidad continuas [4],

$$H(X) \leq \log(\sqrt{2\pi e}\Delta X), \quad (1.26)$$

con igualdad si y sólo si la distribución es gaussiana, lo que significa que para un valor de la desviación estándar dado esta distribución es la que maximiza la entropía.

Capítulo 2

Mecánica Cuántica. Principio de incertidumbre y entrelazamiento

Este capítulo está dedicado a la Mecánica Cuántica haciendo hincapié en los dos conceptos principales sobre los que versa esta tesis: el principio de incertidumbre y su formulación matemática y el fenómeno del entrelazamiento. Antes de entrar de lleno en estas cuestiones se hará una introducción a los conceptos básicos y postulados de la teoría cuántica con el objeto de que esta memoria sea lo más autocontenida posible, de forma que el lector no familiarizado con esta teoría pueda seguir los capítulos siguientes. Nos limitaremos a una exposición de la estructura matemática de la teoría, el lector interesado en un análisis más detallado es remitido a las referencias clásicas (por ejemplo [14]) y a las más modernas [15, 16, 17], especialmente para una fundamentación física de dicha estructura.

2.1. Conceptos básicos de la Mecánica Cuántica

2.1.1. Estados

Uno de los conceptos centrales de la teoría cuántica es el de estado. El estado cuántico codifica toda la información disponible para un observador acerca de un determinado sistema físico. Cada sistema cuántico tiene asociado un espacio de Hilbert H , es decir, un espacio vectorial complejo con producto escalar $\langle \cdot | \cdot \rangle$. Cuando disponemos de máxima información acerca de un sistema cuántico, el estado correspondiente, llamado estado puro, se representa por un rayo unidad de H , o sea, por un vector normalizado de H en el que la fase global es indiferente. Este vector se denota normalmente por ψ y, a veces, recibe también el nombre de función de onda. El opuesto también es cierto, cualquier rayo unidad de H representa un estado cuántico (puro).

El sistema cuántico más sencillo que se puede concebir es el de dos niveles, cuyos estados reciben el nombre de qubits (por ser el análogo cuántico del bit clásico) y para el que $H \simeq \mathbb{C}^2$. Este sería el caso, por ejemplo, de una partícula que puede acceder a dos niveles energéticos (el estado fundamental o el estado excitado), del espín de un electrón a lo largo de una determinada dirección (arriba o abajo) o de la polarización de un fotón (girando hacia la izquierda o hacia la derecha si es circular). Sea $\{|0\rangle, |1\rangle\}$ una base ortonormal de H donde cada elemento representa uno de los dos niveles posibles (por ejemplo, $|0\rangle \equiv$ estado fundamental y $|1\rangle \equiv$ estado excitado). Entonces todos los estados posibles son de la forma

$$|\psi\rangle = a|0\rangle + b|1\rangle \quad (2.1)$$

para $a, b \in \mathbb{C}$ tales que $|a|^2 + |b|^2 = 1$. Sorprendentemente, los únicos estados posibles no son sólo $|0\rangle$ y $|1\rangle$ sino que también puede darse cualquier superposición coherente de ambos, lo que se conoce como principio de superposición. También puede darse el caso de sistemas cuánticos cuyo espacio de Hilbert correspondiente tenga dimensión infinita. Tal sería el caso de la posición de una partícula confinada a una región unidimensional del espacio entre los puntos a y b . En esta situación los posibles estados vendrían representados por una función normalizada de $L^2[a, b]$.

Hasta ahora nos hemos limitado a describir estados de máxima información, es decir, estados puros. La situación más general debe incluir la posibilidad de una descripción del estado en la que la información es parcial, como en el caso de un estado que se prepara en una serie de estados puros diferentes con cierta probabilidad ¿Cómo describir estados que sean mezclas estadísticas de estados puros? En esta descripción lo más general posible resulta que los estados se representan por operadores lineales acotados que actúan en H , denominados operadores densidad y que habitualmente se denotan por ρ . Estos operadores deben ser autoadjuntos ($\rho = \rho^\dagger$), semidefinidos positivos ($\rho \geq 0$) y de traza unidad ($\text{Tr}\rho = 1$). Cuando ρ es de rango 1, es decir, cuando ρ es un proyector ($\rho = |\psi\rangle\langle\psi|$), el operador densidad representa a un estado puro (con vector de estado $|\psi\rangle$). Cuando ρ no es puro se dice mezcla. Debido a las propiedades arriba señaladas, cualquier operador densidad se puede escribir como una combinación convexa de estados puros, i.e.

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \quad (2.2)$$

con $0 \leq p_i \leq 1 \forall i$ y $\sum_i p_i = 1$, lo cual justifica la visión de los estados mezcla como mezclas estadísticas de estados puros. Sin embargo, debe tenerse en cuenta que la descomposición (2.2) no es única por lo que diferentes mezclas estadísticas pueden dar lugar a un mismo estado cuántico. En resumen, para un espacio de Hilbert dado H , el conjunto de todos los estados, que denotaremos $\mathcal{S}(H)$, está formado por los operadores densidad que actúan en H . Éste es un conjunto convexo cuyos puntos extremos son los estados puros. Obviamente $\mathcal{S}(H) \subset \mathcal{B}(H) \subset \mathcal{L}(H)$, donde $\mathcal{B}(H)$ denota el espacio vectorial de los operadores autoadjuntos que actúan en H y $\mathcal{L}(H)$ el espacio vectorial de los operadores

(lineales) que actúan en H . A su vez, ambos espacios constituyen también un espacio de Hilbert, conocido como espacio de Hilbert-Schmidt, i.e. con producto escalar

$$\langle X, Y \rangle_{HS} = \text{Tr}(X^\dagger Y). \quad (2.3)$$

En lo que sigue siempre consideraremos que tanto $\mathcal{B}(H)$ como $\mathcal{L}(H)$ llevan asociado este producto escalar.

La interpretación de los estados puros como estados de máxima información plantea la posibilidad de asociar una medida entrópica a los estados cuánticos que cuantifique el correspondiente grado de mezcla. Esta medida se conoce como entropía de Von Neumann $S(\rho)$ y se define como generalización directa de la entropía de Shannon a los estados cuánticos,

$$S(\rho) = -\text{Tr}(\rho \log \rho) = H(\{\lambda_i(\rho)\}), \quad (2.4)$$

donde $\{\lambda_i(\rho)\}$ son los autovalores de ρ , los cuales, dado que ρ es semidefinido positivo y de traza unidad, constituyen una distribución de probabilidad. De esta forma la entropía de Von Neumann alcanza su valor mínimo, $S(\rho) = 0$, cuando ρ es puro. Otra caracterización habitual de los estados puros es $\text{Tr}\rho^2 = 1$, mientras que los estados mezcla satisfacen $\text{Tr}\rho^2 < 1$.

2.1.2. Estados compuestos

Supongamos que tenemos un sistema físico compuesto por varios subsistemas $A, B, C \dots$ (por ejemplo, varias partículas a repartir en niveles energéticos, el espín de un conjunto de electrones, etc). Sabemos que a cada subsistema le corresponde un determinado espacio de Hilbert (al subsistema A H_A , al subsistema B H_B , ...). Cabe preguntarse, entonces, cuál es la relación entre el espacio de Hilbert H correspondiente al sistema total y los espacios de Hilbert asociados a los subsistemas. La teoría cuántica afirma que el espacio de Hilbert correspondiente al sistema total es el producto tensorial (o de Kronecker) de los espacios de Hilbert de los subsistemas, i.e. $H = H_A \otimes H_B \otimes H_C \otimes \dots$. Nótese que esto implica que $\dim H = \dim H_A \cdot \dim H_B \cdot \dim H_C \cdot \dots$. Por ejemplo, para un sistema de dos qubits ($H \simeq \mathbb{C}^2 \otimes \mathbb{C}^2$) los estados puros son de la forma

$$|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle, \quad (2.5)$$

con $|a|^2 + |b|^2 + |c|^2 + |d|^2 = 1$ y donde $|ij\rangle$ es una abreviación para $|i\rangle_A \otimes |j\rangle_B$ (a veces también se escribe $|i\rangle_A |j\rangle_B$). De igual forma, los estados en general (incluyendo mezclas) vendrán dados por operadores densidad actuando en $H \simeq \mathbb{C}^2 \otimes \mathbb{C}^2$.

Por simplicidad supongamos un sistema compuesto por dos subsistemas A y B . Si cada subsistema X es preparado independiente en el estado ρ_X entonces el estado conjunto es $\rho = \rho_A \otimes \rho_B$. Sin embargo, cuando el estado $\rho \in \mathcal{S}(H_A \otimes H_B)$ del sistema compuesto es preparado actuando conjuntamente sobre ambos subsistemas resulta natural preguntarse cómo es el operador densidad que caracteriza la información que tiene cada observador de su correspondiente

subsistema. Tal operador se obtiene ignorando la información correspondiente a otros subsistemas mediante la traza parcial,

$$\rho_A = \text{Tr}_B \rho \quad (2.6)$$

donde ρ_A es el operador densidad que caracteriza el estado del subsistema A . La traza parcial con respecto a B se define como la operación lineal de $\mathcal{L}(H_A \otimes H_B)$ a $\mathcal{L}(H_A)$ que satisface

$$\text{Tr}_B(X \otimes Y) = X \text{Tr}(Y) \quad (2.7)$$

para $X \in \mathcal{L}(H_A)$ e $Y \in \mathcal{L}(H_B)$. La linealidad de esta operación permite obtener la traza parcial de cualquier operador expresándolo en una base producto. Así si $\{|i\rangle_A\}$ e $\{|i\rangle_B\}$ son bases ortonormales de H_A y H_B respectivamente y

$$\rho = \sum_{ijkl} \rho_{ij,kl} |i\rangle_A \langle j| \otimes |k\rangle_B \langle l|, \quad (2.8)$$

entonces

$$\rho_A = \sum_{ij} \left(\sum_k \rho_{ij,kk} \right) |i\rangle_A \langle j|, \quad (2.9)$$

donde hemos usado que $\text{Tr}(|k\rangle_B \langle l|) = \delta_{kl}$. Análogamente $\rho_B = \text{Tr}_A \rho$. Los operadores densidad que caracterizan el estado de los subsistemas se denominan reducciones de ρ . Nótese que en general no tiene por qué satisfacerse $\rho = \rho_A \otimes \rho_B$.

Cuando tratamos con estados puros de sistemas cuánticos compuestos bipartitos la descomposición de Schmidt constituye una herramienta muy útil que utilizaremos en capítulos posteriores. Se trata de que cuando ψ denota un vector de estado de $H = H_A \otimes H_B$, entonces existen bases ortonormales de H_A y H_B , $\{|i\rangle_A\}$ e $\{|i\rangle_B\}$, tales que ψ admite una descomposición de la forma

$$|\psi\rangle = \sum_i \sqrt{\mu_i} |i\rangle_A |i\rangle_B, \quad (2.10)$$

donde $0 \leq \mu_i \leq 1$ y $\sum_i \mu_i = 1$. La ventaja de considerar la expansión en la base de Schmidt consiste en que muchos de los coeficientes son nulos (todos los correspondientes a $|i\rangle_A |j\rangle_B$ cuando $i \neq j$). La demostración de la existencia de esta descomposición se basa en la descomposición en valores singulares de matrices (véase por ejemplo [17]). Los $\{\mu_i\}$ están unívocamente determinados y reciben el nombre de coeficientes de Schmidt y el número de coeficientes de Schmidt no nulos, que puede ir de uno a $\min\{\dim H_A, \dim H_B\}$, se denomina número de Schmidt. Como consecuencia directa de la descomposición de Schmidt se tiene la isoespectralidad de las reducciones de un estado puro bipartito ($\rho_A = \sum_i \mu_i |i\rangle_A \langle i|$, $\rho_B = \sum_i \mu_i |i\rangle_B \langle i|$).

2.1.3. Medidas

Por el momento nos hemos limitado a analizar la representación matemática del estado cuántico, que, como ya se ha señalado, contiene toda la información

disponible para un observador acerca de un determinado sistema físico. Las medidas sobre el sistema físico son la forma mediante la cual el observador extrae esta información. Las magnitudes medibles, llamadas observables, se representan en la teoría cuántica mediante operadores autoadjuntos que actúan en el espacio de Hilbert H correspondiente al sistema. De igual forma, cualquier operador autoadjunto representa un observable. Sea el observable $M \in \mathcal{B}(H)$ con descomposición espectral

$$M = \sum_i \lambda_i P_i, \quad (2.11)$$

donde P_i denota el proyector sobre el subespacio propio de M correspondiente al autovalor λ_i . Entonces, los posibles resultados de una medida de M sobre un estado cuántico caracterizado por ρ se corresponden con los autovalores de M , $\{\lambda_i\}$, dándose el i -ésimo posible resultado con probabilidad

$$p_i = \text{Tr}(P_i \rho). \quad (2.12)$$

Además, inmediatamente después de la medida el estado del sistema es proyectado al estado correspondiente al resultado de la medida. O sea, asumamos que se ha obtenido el i -ésimo posible resultado, entonces el estado del sistema pasa a ser

$$\rho_i = \frac{P_i \rho P_i}{\text{Tr}(P_i \rho)}. \quad (2.13)$$

Quedan aquí presentadas otras dos de las características sorprendentes de la Mecánica Cuántica. La primera, el carácter probabilístico de la teoría, ya que es imposible predecir con total certeza el resultado de una medida (a menos que el sistema se encuentre en uno de los autoestados del observable) y la segunda, el hecho de que realizar una medida sobre un determinado sistema altera irrevocablemente su estado.

Consideremos, por ejemplo, un qubit en un estado puro dado por (2.1) y una medida M que distinga $\{|0\rangle, |1\rangle\}$, o sea, $M = \lambda_0 |0\rangle\langle 0| + \lambda_1 |1\rangle\langle 1|$ (los autovalores $\lambda_0 \neq \lambda_1$ nos son indiferentes). Entonces,

$$p_0 = \text{Tr}(|0\rangle\langle 0|\rho) = |\langle 0|\psi\rangle|^2 = |a|^2, \quad (2.14)$$

$$p_1 = \text{Tr}(|1\rangle\langle 1|\rho) = |\langle 1|\psi\rangle|^2 = |b|^2. \quad (2.15)$$

Si se obtuviera el resultado 0, de acuerdo con (2.13), el estado del sistema pasaría ser

$$|\psi_0\rangle = \frac{a}{|a|} |0\rangle, \quad (2.16)$$

que es equivalente a $|0\rangle$ puesto que, como ya señalamos, los estados se caracterizan por rayos de H y las fases globales son indiferentes (de hecho ambos casos se corresponden al operador densidad $|0\rangle\langle 0|$).

En el caso de sistemas que tienen asociado un espacio de Hilbert de dimensión infinita este formalismo puede dar lugar a inconsistencias. Por ejemplo, en esta situación se pueden dar operadores autoadjuntos (y, por

tanto, observables) que no tengan autovalores, los cuales se conocen como observables de espectro continuo. Tal es el caso del operador posición X para el sistema antes mencionado de la partícula unidimensional cuyo estado (puro) se representa por una función de onda $\psi(x) \in L^2[a, b]$. Este observable se representa por el operador multiplicación por x , que efectivamente no tiene autovalores. Superar este problema formal pasa por discretizar el espectro continuo y considerar resoluciones espectrales del operador (véase el capítulo 4 de [15] para un tratamiento detallado de este problema, habitualmente ignorado en los textos clásicos sobre Mecánica Cuántica). El resultado final muestra que la probabilidad de que la partícula se encuentre entre los puntos x_1 y x_2 viene dada por

$$p(x_1, x_2) = \int_{x_1}^{x_2} |\psi(x)|^2 dx, \quad (2.17)$$

lo que da lugar en estos casos a la interpretación del cuadrado de la función de onda como densidad de probabilidad.

Las medidas que hemos introducido hasta ahora reciben el nombre de medidas proyectivas o de Von Neumann. Si bien estas son las únicas medidas que físicamente se pueden realizar, el formalismo presentado no agota todas las posibilidades que pueden darse en la teoría. La solución a esta aparente paradoja consiste en que el observador puede unir a su sistema (caracterizado por $\rho \in \mathcal{S}(H)$) un sistema auxiliar (caracterizado por $\rho' \in \mathcal{S}(H')$) y realizar medidas en el sistema conjunto (caracterizado por $\rho \otimes \rho' \in \mathcal{S}(H \otimes H')$), de forma que, aunque la medida en el sistema total sigue el formalismo hasta ahora introducido, éste no basta para explicar lo sucedido en el sistema objeto de nuestro interés una vez descartado el sistema auxiliar. Con esto en mente, la forma más general de describir las medidas en Mecánica Cuántica resulta ser mediante medidas generalizadas o POVM¹, que presentamos a continuación. Estas medidas se describen por una colección de operadores $\{A_i\}$ de $\mathcal{L}(H)$, denominados operadores de Kraus, que satisfagan $\sum_i A_i^\dagger A_i = I$, donde I es la identidad. La probabilidad de obtener el i -ésimo resultado en la medida viene dada por

$$p_i = \text{Tr}(A_i \rho A_i^\dagger). \quad (2.18)$$

y el estado del sistema después de la medida si se da el i -ésimo resultado es

$$\rho_i = \frac{A_i \rho A_i^\dagger}{\text{Tr}(A_i \rho A_i^\dagger)}. \quad (2.19)$$

Nótese que las medidas generalizadas incluyen el caso particular de las medidas proyectivas cuando los $\{A_i\}$ son un conjunto completo de proyectores. Por último, mencionamos que el hecho de que las POVM puedan ser realizadas como medidas proyectivas en un espacio de Hilbert ampliado se demuestra en el teorema de Neumark [18]. Este hecho, unido a razones históricas, justifica que en los próximos capítulos se dé un papel principal a las medidas proyectivas. En muchas ocasiones consideraremos exclusivamente este tipo de

¹Del inglés, *positive operator valued measure*.

medidas, pudiéndose extender de manera relativamente sencilla los resultados considerados a medidas generalizadas.

2.1.4. Evolución

Una vez descritos los estados de los sistemas cuánticos y cómo los observadores extraen información de ellos mediante medidas, queda por aclarar como evolucionan con el tiempo estos sistemas en ausencia de interacciones con observadores, es decir, cuando los sistemas son cerrados. La evolución de un sistema cuántico cerrado viene descrita por un operador unitario de H , de manera que el estado ρ' de un sistema en el instante t' se relaciona con el estado ρ del sistema en el instante previo t mediante el operador unitario U que depende sólo de t y t' como

$$\rho' = U\rho U^\dagger. \quad (2.20)$$

Este postulado puede darse de una forma más refinada, describiendo la evolución en el tiempo como continuo mediante la ecuación de Schrödinger,

$$i\hbar \frac{d\rho}{dt} = [\mathcal{H}, \rho], \quad (2.21)$$

donde $[\cdot, \cdot]$ denota el conmutador, \mathcal{H} es el hamiltoniano del sistema, un operador autoadjunto que actúa en H (observable por consiguiente), y \hbar es la constante de Planck que se puede absorber dentro del hamiltoniano y, por tanto, en lo que sigue tomaremos $\hbar = 1$. La forma explícita del hamiltoniano (que está relacionada con la de U ya que $U = \exp(-iH(t' - t))$), la cual gobierna completamente la dinámica del estado cuántico, debe determinarse de acuerdo con las propiedades físicas del sistema. Más adelante consideraremos algunos hamiltonianos en particular.

Finalmente, merece la pena mencionarse que en la práctica un sistema cuántico difícilmente se encuentra completamente aislado de su entorno como para considerarlo cerrado. Por ello, es unitaria la evolución conjunta del estado con el entorno $\rho \otimes \rho_{\text{entorno}}$, mientras que la evolución reducida del sistema puede tener una descripción más compleja, lo que ha sido estudiado en la teoría de sistemas cuánticos abiertos.

2.1.5. Los postulados de la Mecánica Cuántica

En esta sección, para comodidad del lector, enunciamos los postulados básicos de la teoría cuántica que se han desarrollado anteriormente. Debe decirse que existen estudios que consideran cómo desarrollar la Mecánica Cuántica a partir de un conjunto minimal de axiomas o a partir de axiomáticas alternativas con un mayor sentido físico. Estos trabajos, sin duda interesantes, están fuera del alcance de esta tesis, por lo que aquí nos limitamos a adoptar los postulados estándar.

Postulado 1. *Cada sistema físico tiene asociado un espacio de Hilbert H . Cada posible estado del sistema viene descrito por un operador ρ , denominado*

operador densidad, que actúa en H y que es autoadjunto, semidefinido positivo y de traza unidad.

Postulado 2. El espacio de Hilbert correspondiente a un sistema cuántico multicomponente es el producto tensorial de los espacios de Hilbert correspondientes a los subsistemas.

Postulado 3. Todas las medidas posibles se describen por una colección de operadores $\{A_i\}$ de $\mathcal{L}(H)$ que satisfagan $\sum_i A_i^\dagger A_i = I$. Cada operador representa un posible resultado de la medición. La probabilidad de obtener el i -ésimo resultado viene dada por

$$p_i = \text{Tr}(A_i \rho A_i^\dagger).$$

Inmediatamente después de la medida el estado del sistema pasa a ser

$$\rho_i = \frac{A_i \rho A_i^\dagger}{\text{Tr}(A_i \rho A_i^\dagger)}$$

en el supuesto de que se haya obtenido el i -ésimo posible resultado.

Postulado 4. La evolución temporal de un sistema cuántico cerrado se describe por un operador unitario. En particular, ésta viene dada por la ecuación de Schrödinger,

$$i \frac{d\rho}{dt} = [\mathcal{H}, \rho],$$

donde \mathcal{H} es un observable de H , llamado hamiltoniano, que viene determinado por las características físicas del sistema.

2.2. Principio de incertidumbre y relaciones de incertidumbre

Como acabamos de ver, en Mecánica Cuántica la medida de un observable lleva asociada cierta incertidumbre puesto que la teoría simplemente proporciona probabilidades de que se den determinados resultados. No obstante, esta incertidumbre asociada al resultado de la medida puede ser arbitrariamente pequeña ya que el sistema puede haber sido preparado en un estado arbitrariamente cercano a alguno de los autoestados del correspondiente observable. Es más, puede existir completa certeza acerca del resultado de la medida si el estado del sistema viene representado por uno de estos autoestados. Sin embargo, esta situación no puede mantenerse cuando consideramos la medida de dos observables, lo que da lugar al principio de incertidumbre que enunciamos a continuación. Existe una cota inferior irreducible en la incertidumbre relativa al resultado de una medida simultánea de un par arbitrario de observables que no conmutan en un sistema cuántico. De forma equivalente, puede decirse que existe una cota superior a la precisión con la que se pueden especificar los valores que tomarán un par arbitrario de observables

que no conmutan al preparar un estado cuántico. El principio de incertidumbre es uno de los puntos fundamentales en los que la teoría cuántica se diferencia de las teorías clásicas por lo que desde su formulación ha atraído considerable atención, incluso desde ámbitos habitualmente ajenos al desarrollo científico. Esto ha propiciado en cierta forma la creación de malentendidos en torno al contenido del principio que precisamos brevemente a continuación. En primer lugar cuando se afirma “existe una cota inferior irreducible en la incertidumbre relativa al resultado de una medida” o “existe una cota superior a la precisión con la que se pueden especificar los valores”, en ningún modo se puede achacar esta limitación a errores experimentales o a la dificultad de construir en la práctica un laboratorio con la suficiente precisión experimental. La limitación a la que se refiere el principio de incertidumbre es de carácter fundamental e inherente a la teoría y se basa en consideraciones teóricas en las que (como suelen hacer los físicos teóricos) se asume que todas las observaciones son ideales y perfectamente precisas. Otra desafortunada interpretación del contenido del principio expresa que la medida de un observable con una determinada precisión altera el estado del sistema de forma que queda restringida la precisión con la que puede medirse el otro observable. Si bien es cierto que, como ya hemos visto, la medida de un observable altera el estado del sistema, esto no tiene nada que ver con el principio de incertidumbre. Esta confusión enlaza con otra, que es la del significado de la simultaneidad de las medidas de los dos observables. Precisamente debido a que una medida perturba el estado de un sistema, en la teoría cuántica no son posibles las medidas conjuntas sobre un mismo sistema de dos observables que no conmuten (si los observables conmutan entonces un conocido resultado del álgebra lineal afirma que ambos tienen los mismos autovectores de forma que la medida de uno es equivalente a la medida del otro siendo posible, por tanto, la medida conjunta). La única posible interpretación del principio es entonces considerando la preparación de un gran número de sistemas cuánticos en estados idénticos. En unos se mide un observable y en el resto el otro. Las dos medidas podrían tener en un principio un grado de predictabilidad arbitrario (una incertidumbre asociada al resultado de la medida arbitraria), que es, sin embargo, lo que viene a prohibir el principio de incertidumbre.

Aunque el principio de incertidumbre es una consecuencia directa de los postulados de la Mecánica Cuántica, éste fue desarrollado en 1927, cuando la teoría estaba construyéndose, por W. Heisenberg [19] considerando varios experimentos imaginarios (*gedankenexperiments*) en los que había que determinar la posición y el momento de un electrón. Heisenberg llegó a la conclusión de que el producto de las imprecisiones en la posición y el momento debían ser del orden de la unidad (recuérdese que hemos fijado $\hbar = 1$), de manera que cuanto menor es la imprecisión en la posición mayor es en el momento y viceversa. Otra forma intuitiva de entender el principio de incertidumbre es considerar la medida de dos observables en un sistema cuántico de dimensión finita en un estado puro. Al medir un observable, la única forma de tener una certidumbre absoluta en el resultado es que el estado del sistema sea uno de los autoestados del observable (véase (2.12)). Ahora bien, como los observables no conmutan tienen autoestados diferentes, por lo que resulta imposible que el

estado del sistema sea simultáneamente un autoestado de los dos observables y, por consiguiente, no se puede tener certeza absoluta en el resultado de las medidas de ambos observables. Fijado un autovector de cada observable, el vector de estado del sistema se puede acercar arbitrariamente a uno de ellos a costa de alejarlo del otro, de forma que acercar a uno la probabilidad de obtener un particular resultado al medir un observable implica alejar de uno la probabilidad de obtener un particular resultado al medir el otro observable.

Una vez formulado el principio de incertidumbre y analizado su contenido, resulta natural plantearse dos preguntas. Primero, ¿cómo deducir rigurosamente el principio a partir de los postulados de la teoría cuántica? y, segundo, dada una cantidad de incertidumbre asociada a la medida de un observable, ¿cuál es la menor cantidad de incertidumbre posible para el segundo observable? La respuesta a estas dos preguntas es la misma: mediante relaciones de incertidumbre. Una relación de incertidumbre es una formulación matemática del principio de incertidumbre en forma cuantitativa. Así, estas relaciones pueden deducirse de los axiomas de la teoría cuántica para dar un fundamento riguroso al principio y sirven para hacer estimaciones cuantitativas acerca del mismo. Como resulta obvio, esta empresa precisa de formas de medir la incertidumbre asociada a una distribución de probabilidad, lo que de forma natural lleva a considerar la entropía introducida en el primer capítulo de esta memoria.

2.2.1. Formulación estándar. Relación de incertidumbre de Heisenberg-Robertson

En su análisis Heisenberg, ya ofrece una versión rudimentaria de una relación de incertidumbre al constreñir el producto de las imprecisiones en el conocimiento de la posición y el momento. Sin embargo, no da una definición matemática concreta de estas imprecisiones, sino que se limita a estimarlas separadamente mediante argumentos plausibles. Fue Kennard [20], poco tiempo después, quien demostró la primera formulación matemática del principio de incertidumbre. Concretamente, para un estado puro ψ se tiene que

$$\Delta_\psi X \Delta_\psi P \geq \frac{1}{2}, \quad (2.22)$$

donde X es el observable correspondiente a la posición y P el correspondiente al momento. Recordamos que en la representación de posiciones x , el operador X se representa por multiplicación por x y el operador P por $-id/dx$. Alternativamente, en la representación de momentos p , X se representa por id/dp y P por p . Todo esto implica que la función de onda en la representación de momentos $\phi(p)$ es la transformada de Fourier de la función de onda en la representación de posiciones (i.e. el estado del sistema para la posición condiciona completamente su estado para el momento y viceversa), lo que será relevante más adelante. $\Delta_\psi X$ y $\Delta_\psi P$ denotan las desviaciones estándar de las distribuciones de probabilidad asociadas a las medidas de X y P en el sistema cuántico caracterizado por ψ , que se definen como en (1.20). Los valores

esperados en estos casos se obtienen adaptando directamente (1.22) al lenguaje cuántico vía (2.17),

$$\langle X \rangle_\psi = \int x |\psi(x)|^2 dx, \quad (2.23)$$

$$\langle P \rangle_\psi = \int p |\phi(p)|^2 dp = -i \int \overline{\psi(x)} \psi'(x) dx. \quad (2.24)$$

La desigualdad de Kennard es óptima, en el sentido de que la cota inferior se alcanza para algún estado. Este es el caso cuando las distribuciones de probabilidad para la posición y el momento vienen dadas por sendas gaussianas, lo cual sucede para sistemas físicos reales como el estado fundamental del oscilador armónico unidimensional, que analizaremos más adelante.

No mucho tiempo después, Robertson [21] generalizó la relación de incertidumbre (2.22) a un par arbitrario de observables A y B ,

$$\Delta_\psi A \Delta_\psi B \geq \frac{1}{2} |\langle [A, B] \rangle_\psi|. \quad (2.25)$$

Nótese que $[X, P] = iI$, por lo que (2.22) se puede deducir de (2.25). Cuando A represente un observable para un sistema cuántico de dimensión finita, ΔA se sigue definiendo por (1.20) y los valores esperados se obtienen ahora traduciendo (1.21) al formalismo cuántico mediante (2.12), o sea,

$$\langle A \rangle_\rho = \text{Tr}(A\rho), \quad (2.26)$$

que se reduce a $\langle \psi | A | \psi \rangle$ para estados puros. La demostración de (2.25) se basa únicamente en la desigualdad de Cauchy-Schwarz. Para que las manipulaciones matemáticas de la demostración así como la relación de incertidumbre en sí tengan sentido debe cumplirse que ψ pertenezca a los dominios² de AB , BA , A^2 y B^2 (además de, por supuesto, pertenecer a los dominios de A y B). Obviamente, esto no representa ningún problema para operadores de espacios de Hilbert de dimensión finita; sin embargo, se debe ser cuidadoso en el caso infinito con los operadores no acotados (como X y P , por ejemplo). A Heisenberg le satisfizo sobremana esta formulación porque conectaba dos de los aspectos revolucionarios de la por entonces nueva teoría: el principio de incertidumbre y la no conmutatividad. Por ello, fue adoptada por él como la auténtica expresión matemática del principio de incertidumbre, por lo que se conoce como relación de incertidumbre de Heisenberg-Robertson.

2.2.2. Crítica a la formulación estándar. Formulaciones entrópicas

La relación de incertidumbre de Heisenberg-Robertson también se denomina relación de incertidumbre estándar ya que, siguiendo la opinión de Heisenberg, esta expresión ha sido tradicionalmente considerada como la formulación

² $\psi \in H$ pertenece al dominio del operador A si $A(\psi) \in H$.

matemática del principio de incertidumbre en prácticamente todos los textos clásicos de Mecánica Cuántica. Sin embargo, como analizaremos a continuación esta relación padece serias deficiencias que impiden considerarla una formulación matemática satisfactoria del principio de incertidumbre. Sorprendentemente, esta ineficacia no fue puesta claramente de manifiesto hasta la década de los 80 gracias a los trabajos de Hilgevoord y Uffink [22, 8] y, principalmente, de Deutsch [23], los cuales seguimos de cerca en esta sección.

De acuerdo al enunciado del principio de incertidumbre cualquier formulación matemática cuantitativa del principio debe ser de la forma

$$\left[\begin{array}{c} \text{Incertidumbre en el resultado} \\ \text{de la medida de } A \text{ y } B \end{array} \right] \geq [\text{Cota inferior irreducible}], \quad (2.27)$$

o sea,

$$M(A, B, \rho) \geq M_0(A, B) = \inf_{\rho} M(A, B, \rho), \quad (2.28)$$

donde M será una medida de la incertidumbre asociada a la medición de A y B en un sistema cuántico arbitrario con estado dado por ρ y la cota M_0 es un número que queda fijado por el par de observables. Este número será estrictamente positivo salvo que A y B tengan autoestados comunes, en cuyo caso valdrá cero (que es el valor que se asume tomará M en caso de completa certidumbre).

En primer lugar, por tanto, para expresar el principio de incertidumbre se debe elegir la medida de incertidumbre M . Teniendo en cuenta lo desarrollado en el primer capítulo la elección natural es la entropía de información de Shannon. Consecuentemente, llama la atención la inclinación por la desviación estándar por parte de los pioneros en el estudio del principio de incertidumbre. Es más, esta elección resulta ser totalmente arbitraria puesto que ni Heisenberg ni Kennard dan explicación alguna y Robertson se limita a señalar que se produce “in accordance with statistical usage” [21]. Uffink y Hilgevoord [22, 8] han sugerido que esta elección se puede deber a una confusión entre la incertidumbre cuántica y los errores en las medidas de las teorías clásicas. En el segundo caso el observador mide una magnitud para la cual se asume que existe un único valor que se desconoce. Debido a complicaciones experimentales el valor medido difiere del valor verdadero. Esta imprecisión se debe a varios factores independientes y se puede demostrar que, bajo condiciones generales, estos errores siguen una distribución gaussiana, para la cual, como ya vimos, la desviación estándar es una medida razonable de la incertidumbre. Sin embargo, ya se ha analizado que la incertidumbre de la teoría cuántica es de una naturaleza completamente diferente. Pero, además, la elección de la desviación estándar como medida de incertidumbre resulta ser no sólo arbitraria sino inadecuada. Las razones ya han sido señaladas básicamente en la sección 1.3.2; en primer lugar esta medida es incapaz de estimar la incertidumbre para muchas distribuciones y, en segundo lugar, su valor depende de los valores que tome la variable aleatoria. Esto último significa en el caso cuántico que el valor de la desviación estándar depende de los autovalores de los observables, los cuales son totalmente irrelevantes de cara a la incertidumbre en el resultado de la medida ya que, de acuerdo con (2.12),

la distribución de probabilidad que rige este proceso aleatorio sólo depende de los autovectores. Así, los únicos valores significativos de la desviación estándar son nula o no nula porque un mero reetiquetado de los autovalores, el cual no tiene ningún significado físico, produce un cambio arbitrario en el valor de la desviación estándar.

En segundo lugar, a diferencia de (2.28) el miembro derecho de la relación de incertidumbre de Heisenberg-Robertson no es una cota fija dado el par A, B sino que depende del estado ψ del sistema. La única excepción se da para el par X, P para el que se obtiene la relación de Kennard (2.22) (esta situación no puede darse en el caso de espacios de Hilbert de dimensión finita puesto que $\text{Tr}[A, B] = 0$ y, por tanto, el conmutador no puede ser proporcional a la identidad). Esta deficiencia no se reduce a un reproche meramente formal, nótese que para algunas elecciones de ψ esta cota se anula aun en el caso de que A y B no tengan autoestados simultáneos de forma que para estos casos (2.25) falla a la hora de expresar el principio de incertidumbre. Por ejemplo, para un autoestado de A (que no es autoestado de B) tanto ΔA como la cota de Robertson se anulan, de forma que la relación de incertidumbre (2.25), trivial en este caso, no expresa ninguna restricción en la incertidumbre de B cuando ésta está claramente constreñida.

Por último, recuérdese que (2.25) sólo tiene sentido cuando los estados pertenecen también a los dominios de AB y BA , lo que resulta problemático en el caso de operadores no acotados de espectro continuo. Esto produce que para aquellos estados que pertenecen a los dominios de A y B pero no a los de AB y/o BA no existe una relación de incertidumbre si sólo consideramos la formulación estándar. Un ejemplo particularmente traumático viene dado por el par ángulo Φ (que se representa por $\phi \in [0, 2\pi]$) y momento angular $P_\phi = -i\hbar d/d\phi$, ambos operadores autoadjuntos de $L^2[0, 2\pi]$ si el dominio de P_ϕ se restringe a funciones periódicas $\psi(0) = \psi(2\pi)$, para los que la relación de Heisenberg-Robertson tomaría la forma $\Delta\Phi\Delta P_\phi \geq 1/2$ ya que $[\Phi, P_\phi] = i\hbar$. Sin embargo, existen muchos estados relevantes que no pertenecen al dominio de $P_\phi\Phi$ para los que la relación de incertidumbre no tiene sentido, como por ejemplo los autoestados de P_ϕ ($\psi_m(\phi) = \exp(im\phi)/\sqrt{2\pi}$) para los que se tiene $\Delta\Phi\Delta P_\phi = 0$ (aparente contradicción con la relación de Heisenberg-Robertson que ha llevado a confusión a muchos autores).

Como conclusión a este análisis se llega al previamente anunciado resultado de que la relación de incertidumbre estándar (2.25) no es una formulación matemática satisfactoria del principio de incertidumbre. No obstante, no se debe ser demasiado riguroso con los padres de la teoría cuántica, nótese que los trabajos de Shannon dirigidos a la construcción de medidas rigurosas de la incertidumbre no fueron desarrollados hasta cerca de 20 años después de los trabajos de Heisenberg, Kennard y Robertson³. Sin embargo, sí llama la atención

³Así mismo, el concepto más general de estados representados por un operador densidad también es considerablemente posterior. El lector moderno puede preguntarse si la desigualdad de Heisenberg-Robertson (2.25) sigue siendo válida para estados mezcla. La respuesta es afirmativa. Esto se sigue de la demostración alternativa de la desigualdad dada también por Robertson en [24] que se debe a la no negatividad de la matriz de covarianza. Puesto que esta

que los científicos posteriores, salvo contadas excepciones, no se percataran de las limitaciones de la formulación estándar hasta bien entrado el siglo XX en los trabajos anteriormente citados. En cualquier caso, debe subrayarse que a pesar de sus limitaciones para expresar el principio de incertidumbre, la desigualdad de Heisenberg-Robertson es, por supuesto, cierta como relación matemática, y como tal puede ser útil en numerosas aplicaciones (de hecho, se utilizarán desigualdades de este tipo en capítulos posteriores).

Así mismo, de este análisis también se deduce que una formulación matemática satisfactoria del principio de incertidumbre debe ser de la forma [23]

$$H_\rho(A) + H_\rho(B) \geq H_{AB}, \quad (2.29)$$

donde $H_\rho(A)$ y $H_\rho(B)$ son las entropías de Shannon asociadas a la distribución de probabilidad correspondiente al proceso aleatorio que representa la medida de A y B en el sistema cuántico caracterizado por ρ , o sea,

$$H_\rho(A) = - \sum_i \text{Tr}(P_i(A)\rho) \log \text{Tr}(P_i(A)\rho) \quad (2.30)$$

cuando A es un observable de dimensión finita y $P_i(A)$ denota el proyector en el i -ésimo subespacio propio de A . Puesto que esta relación de incertidumbre se basa en la entropía como medida de la incertidumbre, una desigualdad del tipo (2.29) recibe el nombre de relación de incertidumbre entrópica (EUR). Dado que las entropías son no negativas se tiene que $H_{AB} \geq 0$. Además, como las entropías son nulas solamente para la distribución de probabilidad de máxima certeza $(1, 0, 0, \dots)$ se obtiene el satisfactorio resultado de que $H_{AB} = 0$ si y sólo si el estado del sistema es un autoestado común de A y B . Así mismo, una EUR abarca desde un principio el caso más general en el que los estados cuánticos vienen descritos por un operador densidad ρ . Pero, además, sea $\rho = \sum_i |\psi_i\rangle\langle\psi_i|$, debido a la concavidad de la entropía se cumple que $H_\rho(A) + H_\rho(B) \geq \sum_i p_i(H_{\psi_i}(A) + H_{\psi_i}(B))$, por lo que si H_{AB} es una cota válida para estados puros entonces también lo es para estados mezcla. Este hecho simplifica considerablemente la situación al nivel matemático ya que a la hora de buscar la forma explícita de H_{AB} , uno se puede restringir al caso de los estados puros con la tranquilidad de saber que la cota también será válida en el caso general. Por esta razón y para aligerar la notación, en lo que sigue no explicitaremos el ρ o ψ que denota la dependencia en el estado de entropías, probabilidades, valores esperados, etc. Ocasionalmente incumpliremos esta regla y escribiremos ψ para aclarar que el resultado se refiere únicamente a estados puros o ρ para subrayar que el resultado es válido también para estados mezcla. Por supuesto, también se pueden formular EURs para el caso en el que los observables actúan en un espacio de Hilbert de dimensión infinita. En tal

matriz sigue siendo semidefinida positiva para estados mezcla (ver por ejemplo [25, 26]) se tiene que (2.25) también se cumple para estos estados.

situación las entropías toman la forma

$$H(X) = - \int |\psi(x)|^2 \log |\psi(x)|^2 dx, \quad (2.31)$$

$$H(P) = - \int |\phi(p)|^2 \log |\phi(p)|^2 dp, \quad (2.32)$$

para, por ejemplo, el caso de los operadores posición y momento. Evidentemente, en este caso, en un principio H_{AB} puede tomar valores negativos puesto que a la distribución de máxima certidumbre le corresponde el valor de la entropía $-\infty$.

Dedicamos el resto de la sección a dar cuenta de las EURs conocidas hasta la fecha, es decir, a analizar la forma explícita que se puede deducir para H_{AB} en (2.29) según sean A y B . El lector de inclinaciones más matemáticas encontrará demostraciones rigurosas y con un lenguaje de su gusto de los resultados que revisamos a continuación en [27]. El interés de las EURs no se limita a su expresión rigurosa del principio de incertidumbre. Así mismo, constituyen una herramienta muy útil en la denominada Teoría Cuántica de la Información [17, 28, 29], que estudia la transmisión y manipulación de información almacenada en sistemas regidos por las leyes de la Mecánica Cuántica. Las EURs proporcionan una conexión natural entre el principio de incertidumbre y esta rama de la física cuántica en la que los conceptos de información y entropía también aparecen. De esta forma, por ejemplo, las EURs permiten acotar la información que se puede extraer de sistemas para comunicación cuántica [30, 31], demostrar la seguridad de protocolos de criptografía cuántica [32, 33, 34, 35] y demostrar la existencia de fenómenos particulares como el de “locking” de correlaciones clásicas [36, 37]. Además, las EURs, al igual que otras relaciones de incertidumbre, permiten obtener condiciones para verificar la presencia de entrelazamiento (un fenómeno muy relevante en Teoría Cuántica de la Información que introduciremos en la sección 2.4), lo cual estudiaremos en detalle en otros capítulos de esta tesis.

Relación de incertidumbre de Białynicki-Birula-Mycielski

Con anterioridad a los trabajos antes señalados, algunos autores ya consideraron el uso de la entropía para expresar el principio de incertidumbre. Este es el caso de Everett (en su famosa tesis [38], que dio lugar a la interpretación de los muchos mundos de la Mecánica Cuántica) y de Hirschman [39], los cuales demostraron independientemente en 1957 la primera EUR:

$$H(X) + H(P) \geq \log(2\pi), \quad (2.33)$$

que constituye una relación de incertidumbre para el par posición y momento. Su demostración se basa en la desigualdad de Hausdorff-Young (recuérdese que $\phi(p)$ es la transformada de Fourier de $\psi(x)$). Ambos autores conjeturaron que la cota de (2.33) no era óptima, sugiriendo que podría ser elevada a $1 + \log \pi$, puesto que este es el valor que toma $H(X) + H(P)$ cuando las distribuciones de probabilidad de X y P son gaussianas, que era el caso de mínima incertidumbre según la

relación de incertidumbre de Kennard (2.22). Efectivamente, esta conjetura fue probada unos 20 años después por Białynicki-Birula y Mycielski [40],

$$H(X) + H(P) \geq 1 + \log \pi, \quad (2.34)$$

que es, por tanto, una desigualdad óptima. Su demostración se sigue de la desigualdad de Beckner [41], que es una mejora de la desigualdad de Hausdorff-Young. Además de ser una expresión más adecuada del principio de incertidumbre, la EUR de Białynicki-Birula-Mycielski es más fuerte que la relación de incertidumbre de Kennard ya que ésta última se puede deducir de la primera utilizando (1.26).

Así mismo, en [39] también se dedujo una EUR para los observables de un espacio de Hilbert de dimensión infinita Φ y P_ϕ a partir de la desigualdad de Hausdorff-Young (pero expresada para series de Fourier en vez de integrales ya que el operador P_ϕ es de espectro discreto)

$$H(\Phi) + H(P_\phi) \geq \log(2\pi), \quad (2.35)$$

que es óptima ya que la cota se alcanza para los autoestados de P_ϕ . Esto muestra que la EUR para Φ y P_ϕ está libre de los problemas antes señalados que padecía la correspondiente relación de incertidumbre en la formulación estándar, lo cual se extiende, en general, a todas las EURs de observables del caso infinito-dimensional. Esto se debe a que, a diferencia de la formulación estándar, para que la EUR tenga sentido no se necesita exigir que las funciones de onda sean del dominio del producto de los operadores. Sin embargo, sí debe exigirse que la integral entrópica esté definida, lo cual no siempre ocurre. No obstante, en contraste con los casos patológicos de la relación de incertidumbre estándar, este fenómeno no se aprecia en distribuciones de probabilidad con relevancia física.

Merece la pena señalarse que a partir de las desigualdades de las que se deducen estas EURs, también pueden obtenerse relaciones de incertidumbre para entropías generalizadas, por ejemplo [42],

$$H_{2-2/q}(X) + H_{q/2}(P) \geq \log\left(\frac{2\pi}{q}\right) + \frac{(q-1)\log(q-1)}{q-2}, \quad q \geq 1. \quad (2.36)$$

Nótese que tomando el límite $q \rightarrow 2$ se recupera (2.34).

Relación de incertidumbre de Deutsch

La primera EUR para un par de observables de un espacio de Hilbert de dimensión finita N fue deducida por Deutsch en [23], donde se demuestra que

$$H(A) + H(B) \geq -2 \log\left(\frac{1+c}{2}\right), \quad (2.37)$$

donde, si $\{|a_i\rangle\}$ y $\{|b_i\rangle\}$ ($i = 1, \dots, N$) denotan conjuntos completos de autovectores ortonormales de A y B ,

$$c = c(A, B) \equiv \max_{i,j} |\langle a_i | b_j \rangle|, \quad (2.38)$$

que se conoce como solapamiento de los observables A y B (nótese que $1/\sqrt{N} \leq c \leq 1$). Por simplicidad, se asume que los observables A y B tienen espectros no degenerados (todos los autovalores son de multiplicidad 1), de forma que tienen N posibles resultados en sus medidas siendo $p_i(A) = \text{Tr}(|a_i\rangle\langle a_i|\rho) = \langle a_i|\rho|a_i\rangle$ y análogamente para B sus correspondientes probabilidades según (2.12).

La cota inferior de (2.37) es no trivial para cualquier par de observables sin autoestados comunes; sin embargo, no es óptima en el sentido de que no se alcanza para ningún estado cuántico. Aunque la desigualdad (2.37) resulte por tanto bastante débil, su principal mérito reside en que demostraba la posibilidad de formular una EUR para observables de espectro discreto.

Relación de incertidumbre de Maassen-Uffink

Cinco años después del trabajo de Deutsch, Maassen y Uffink [10] mejoraron la anterior relación estableciendo

$$H(A) + H(B) \geq -2 \log c, \quad (2.39)$$

lo cual había sido conjeturado con anterioridad en [43]. Esta desigualdad es la EUR más fuerte conocida hasta la fecha para un par de observables arbitrarios en el caso finito-dimensional. La desigualdad (2.39) se sigue de una versión adecuada del teorema de convexidad de Riesz-Thorin. Así mismo, su demostración también asume la no degeneración de los espectros de A y B , aunque se puede dar una versión para el caso de espectros degenerados y medidas POVM [44].

En el caso particular de que A y B sean observables complementarios [45, 43], i.e. $|\langle a_i|b_j\rangle| = 1/\sqrt{N}$ para todo $i, j = 1, \dots, N$, la cota inferior $\log N$ dada por (2.39) es óptima ya que es alcanzada cuando el sistema se encuentra en algún autoestado de A o B [43]. No obstante, dejando aparte este caso especial la EUR de Maassen-Uffink no es óptima. El problema de encontrar una cota óptima para un par arbitrario de observables resulta ser muy difícil. De hecho, hasta ahora sólo se ha resuelto restringiéndose a un espacio de Hilbert de dimensión dos [46] (véase también [47]).

Al igual que en casos anteriores el proceso de demostración de la EUR de Maassen-Uffink permite derivar también una relación de incertidumbre con entropías generalizadas

$$H_{p/2}(A) + H_{q/2}(B) \geq -2 \log c, \quad \frac{1}{p} + \frac{1}{q} = 1. \quad (2.40)$$

Relación de incertidumbre de Landau-Pollak

Como consecuencia de la desigualdad triangular en la esfera unidad,

$$\theta_{|a_i\rangle, |\psi\rangle} + \theta_{|b_i\rangle, |\psi\rangle} \geq \theta_{|a_i\rangle, |b_i\rangle}, \quad (2.41)$$

siendo $\theta_{x,y}$ el ángulo que forman los vectores x e y , se puede deducir

$$\arccos \sqrt{M_\infty(A, \psi)} + \arccos \sqrt{M_\infty(B, \psi)} \geq \arccos c, \quad (2.42)$$

donde $M_\infty(A, \psi) \equiv \max_i p_i(A, \psi)$ siguiendo la notación de Uffink [8, 10], que demostró (2.42) adaptando al lenguaje cuántico el trabajo de Landau y Pollak en incertidumbre en Teoría de la Señal [48]. Por ello, la desigualdad (2.42) se conoce como relación de incertidumbre de Landau-Pollak. A pesar de su apariencia enrevesada, esta desigualdad captura la esencia del principio de incertidumbre ya que proporciona una cota no trivial a las distribuciones de probabilidad para A y B cuando $c < 1$ (i.e. cuando no tienen autoestados comunes) impidiendo que $\max_i p_i(A, \psi)$ y $\max_i p_i(B, \psi)$ estén arbitrariamente cerca de 1 de forma simultánea. Es más, la relación de incertidumbre de Landau-Pollak no es más débil que la EUR de Maassen-Uffink, aunque tampoco más fuerte, dado que se pueden encontrar ejemplos de distribuciones de probabilidad admitidas por una pero prohibidas por la otra y viceversa [8]. Lo mismo ocurre si la comparamos con la desigualdad de Maassen-Uffink en su versión para entropías generalizadas (2.40). Sin embargo, aunque la medida de incertidumbre en (2.42) está directamente relacionada con las entropías generalizadas ($M_\infty(A) = \exp(-H_\infty(A))$ según (1.16)), la relación de incertidumbre de Landau-Pollak no depende de la medida deseada que es la entropía de Shannon. Además, lo que es peor, esta relación se demuestra únicamente para estados puros y, a diferencia de las EURs, la complicada forma de la medida de incertidumbre en este caso impide extender trivialmente su validez a estados generales (mezcla) a excepción del caso bidimensional como mostramos a continuación en el Teorema 2.1. Por ello, en muchas ocasiones resulta útil considerar la siguiente versión más débil pero de validez general

$$M_\infty(A, \rho) + M_\infty(B, \rho) \leq 1 + c, \quad (2.43)$$

que se puede demostrar maximizando $M_\infty(A) + M_\infty(B)$ sujeto a la condición (2.42) y extendiendo para estados mezcla por la convexidad de M_∞ , i.e. si $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, $M_\infty(A, \rho) \leq \sum_i p_i M_\infty(A, \psi_i)$ (con igualdad si, y sólo si, el resultado de probabilidad máxima es el mismo para todos los ψ_i). Nos referiremos a (2.43) como relación de incertidumbre de Landau-Pollak débil. Como en los casos anteriores la desigualdad de Landau-Pollak sólo es válida para medidas proyectivas no degeneradas.

Teorema 2.1. *En el caso de que el sistema cuántico tenga asociado un espacio de Hilbert bidimensional, la relación de incertidumbre de Landau-Pollak (2.42) es válida para estados generales (mezcla).*

Demostración. Sea $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ la descomposición espectral de la matriz densidad. Puesto que $f(x) = \arccos \sqrt{x}$ es una función decreciente y M_∞ es convexo se tiene que

$$\arccos \sqrt{M_\infty(A, \rho)} \geq \arccos \sqrt{\sum_i p_i M_\infty(A, \psi_i)}.$$

Teniendo ahora en cuenta la concavidad de $f(x)$ en $[1/2, 1]$ (nótese que en el

caso bidimensional $1/2 \leq M_\infty \leq 1$) se llega a

$$\arccos \sqrt{M_\infty(A, \rho)} \geq \sum_i p_i \arccos \sqrt{M_\infty(A, \psi_i)}.$$

Por último, considerando la misma desigualdad para $\arccos \sqrt{M_\infty(B, \rho)}$ y aplicando la relación de incertidumbre de Landau-Pollak para estados puros (2.42) se llega al resultado deseado

$$\arccos \sqrt{M_\infty(A, \rho)} + \arccos \sqrt{M_\infty(B, \rho)} \geq \arccos c.$$

□

2.3. Incertidumbre en sistemas atómicos y moleculares

Los sistemas atómicos y moleculares son un ejemplo clásico de sistema mecanocuántico. Los estados de posición y momento en estos casos se caracterizan por tener unos valores específicos permitidos de la energía, por lo que la descripción adecuada de estos sistemas pasa por conocer las autofunciones del hamiltoniano correspondiente, que, además de ser el operador responsable de la evolución temporal de los estados como ya vimos en la sección 2.1, es el observable asociado a esta magnitud. Es decir, el hamiltoniano \mathcal{H} en este caso es un observable de un espacio de Hilbert de dimensión infinita pero de espectro discreto cuyos autovalores son la energía del estado $E \in \mathbb{R}$: $\mathcal{H}\psi_{\{n\}}(x) = E_{\{n\}}\psi_{\{n\}}(x)$, donde $\{n\}$ es un conjunto de etiquetas que caracteriza a las autofunciones del operador. El hamiltoniano se representa por

$$\mathcal{H} = -\frac{1}{2m} \frac{d^2}{dx^2} + V(x), \quad (2.44)$$

donde m es la masa del sistema y V el potencial que describe su comportamiento. Así, resolver el problema de autovalores anterior significa resolver la ecuación diferencial

$$\left[-\frac{1}{2m} \frac{d^2}{dx^2} + V(x) \right] \psi(x) = E\psi(x), \quad (2.45)$$

la cual se conoce como ecuación de Schrödinger estacionaria. El potencial V modeliza el comportamiento del sistema y su forma explícita se deduce en base a argumentaciones físicas. El potencial por antonomasia y, posiblemente, el más utilizado en la práctica es el del oscilador armónico en el que $V(x) \propto x^2$ (asumimos el caso unidimensional por simplicidad). Resolviendo la ecuación de Schrödinger correspondiente a este potencial se halla que los autoestados del hamiltoniano para la posición son de la forma

$$\psi_n(x) = \frac{1}{\sqrt{\alpha\sqrt{\pi}2^n n!}} e^{-\frac{x^2}{2\alpha^2}} H_n(x/\alpha) \quad (n = 0, 1, 2, \dots), \quad (2.46)$$

donde α es un parámetro del sistema y $H_n(x)$ denota los polinomios de Hermite de grado n , ortogonales en $(-\infty, \infty)$ con respecto a la función peso $\exp(-x^2)$

$$\int_{-\infty}^{\infty} H_n(x)H_m(x)e^{-x^2} dx = \sqrt{\pi}2^n n! \delta_{nm}. \quad (2.47)$$

Otro potencial muy importante es el de Coulomb, $V(x) \propto -1/|x|$, que en el caso tridimensional modeliza al átomo de hidrógeno y sirve como punto de partida aproximado para los átomos más complejos. La ecuación de Schrödinger correspondiente se resuelve utilizando coordenadas esféricas r, θ, ϕ ya que así la solución factoriza

$$\psi_{nlm}(r, \theta, \phi) = R_{nl}(r)Y_{lm}(\theta, \phi) \quad (n \geq 1; l = 0, 1, \dots, n-1; m = -l, \dots, l). \quad (2.48)$$

La parte radial viene dada por

$$R_{nl}(r) = \frac{2^{l+1}}{n^{l+2}} \left(\frac{(n-l-1)!}{(n+l)!} \right)^{1/2} r^l e^{-r/n} L_{n-l-1}^{(2l+1)} \left(\frac{2r}{n} \right), \quad (2.49)$$

donde $L_n^{(\alpha)}(x)$ son los polinomios de Laguerre de grado n y parámetro α , ortogonales si $\alpha > -1$ en $[0, \infty)$ con respecto a la función peso $x^\alpha \exp(-x)$

$$\int_0^\infty L_n^{(\alpha)}(x)L_m^{(\alpha)}(x)x^\alpha e^{-x} dx = \frac{\Gamma(n+\alpha+1)}{n!} \delta_{nm}. \quad (2.50)$$

La parte angular viene dada por los armónicos esféricos $Y_{lm}(\theta, \phi)$,

$$Y_{lm}(\theta, \phi) = \frac{\sqrt{(2l+1)(l-m)!(l+m)!}}{\sqrt{\pi}2^{m+1}l!} (\sin \theta)^m e^{im\phi} C_{l-m}^{(m+1/2)}(\cos \theta), \quad (2.51)$$

donde $C_n^{(\lambda)}(x)$ son los polinomios de Gegenbauer o ultrasféricos de grado n y parámetro λ , ortogonales si $\lambda > -1/2$ en $[-1, 1]$ con respecto a la función peso $(1-x^2)^{\lambda-1/2}$

$$\int_{-1}^1 C_n^{(\lambda)}(x)C_m^{(\lambda)}(x)(1-x^2)^{\lambda-1/2} dx = \frac{2^{1-2\lambda}\pi\Gamma(n+2\lambda)}{(n+\lambda)n![\Gamma(\lambda)]^2} \delta_{nm}. \quad (2.52)$$

De hecho, la propiedad de separación de variables (2.48) no es exclusiva del potencial de Coulomb. Esto también es posible para cualquier potencial central, es decir, cuando la interacción depende únicamente de la distancia (i.e. $V = V(r)$). En tal caso la parte angular también viene dada por los armónicos esféricos ya que todos los potenciales centrales tienen la misma función de onda angular y se diferencian sólo en la parte radial.

Estos ejemplos sirven para ilustrar el hecho de que, en general, las autofunciones del hamiltoniano en multitud de sistemas atómicos y moleculares vienen dadas en términos de polinomios ortogonales y otras funciones especiales. De igual forma, las correspondientes funciones de onda en el espacio de

momentos (que, recuérdese, son la transformada de Fourier de $\psi(x)$) también se representan por este tipo de funciones. [49] constituye una referencia muy completa como introducción a los polinomios ortogonales y otras funciones especiales, sus propiedades y sus aplicaciones en la física cuántica.

Así pues, el cálculo de las entropías de la distribución de probabilidad para posición y momento (2.31)-(2.32) para los estados de estos importantes sistemas cuánticos involucra el cálculo de complicadas integrales de funciones especiales; problema que se reduce a la evaluación de la denominada entropía de polinomios ortogonales, el cual explicamos en la siguiente sección. El interés en el cálculo de estas entropías es doble. Por un lado, de acuerdo a lo desarrollado previamente, proporciona una medida cuantitativa rigurosa de la incertidumbre en la posición y momento de estos sistemas, lo que, por ejemplo, permite analizar cómo se distribuye la incertidumbre en la relación de Bialynicki-Birula-Mycielski (2.34) para estos sistemas. Por otro lado, un conocimiento riguroso de la incertidumbre en estos sistemas da en cierto modo la llave para conocer otras de sus propiedades. De acuerdo a la moderna teoría del funcional de densidad [50], las propiedades físicas y químicas de sistemas multielectrónicos quedan completamente descritas por la densidad de probabilidad en posición y momento de una única partícula. De esta forma las entropías (2.31)-(2.32) se pueden utilizar para estudiar la dinámica y la estructura de sistemas atómicos y moleculares así como para estimar magnitudes físicamente relevantes como la energía cinética y la susceptibilidad magnética [51, 52, 53, 54, 55, 56] (véase también [57]).

2.3.1. Entropía de polinomios ortogonales

Sea $\{p_n(x)\}$ una secuencia de polinomios ortogonales en $\Delta \subseteq \mathbb{R}$ con respecto a la función peso $\omega(x)$,

$$\int_{\Delta} p_n(x)p_m(x)\omega(x)dx = \varkappa_n\delta_{nm}. \quad (2.53)$$

Podemos, entonces, considerar las distribuciones de probabilidad con función densidad normalizada dada por

$$\rho_n(x) = \frac{1}{\varkappa_n} p_n^2(x)\omega(x). \quad (2.54)$$

Estas distribuciones de probabilidad son grosso modo las que aparecen en muchos sistemas cuánticos como los expuestos más arriba. La entropía asociada a estas distribuciones es por tanto

$$H(X) = - \int_{\Delta} \rho_n(x) \log \rho_n(x) dx = \log \varkappa_n + \frac{1}{\varkappa_n} (E(p_n) + I(p_n)), \quad (2.55)$$

donde

$$E(p_n) = - \int_{\Delta} p_n^2(x) \omega(x) \log p_n^2(x) dx, \quad (2.56)$$

$$I(p_n) = - \int_{\Delta} p_n^2(x) \omega(x) \log \omega(x) dx. \quad (2.57)$$

Las integrales $I(p_n)$ son relativamente sencillas y se han evaluado en forma analítica para todas las familias clásicas (Hermite, Laguerre y Jacobi, de los que los antes mencionados polinomios de Gegenbauer son un caso particular) [58]. Así el problema del cálculo de las entropías de información en posición y momento de los sistemas cuánticos mencionados se reduce a la evaluación de las integrales $E(p_n)$, que se conocen como entropías de información de los polinomios p_n . Éstas son además un objeto matemático de interés puesto que están relacionadas con los potenciales logarítmicos y las normas L^p de los polinomios ortogonales (véase por ejemplo [59]). El cálculo de estas integrales es un problema sumamente difícil, como ponen de manifiesto los resultados obtenidos desde que se comenzó su estudio a principios de los 90. De hecho, como se analizará más en detalle en el capítulo 3, sólo se tienen resultados analíticos para unos pocos casos particulares de los polinomios de Gegenbauer. En la mayoría de los casos tan sólo se dispone de desarrollos asintóticos para $n \rightarrow \infty$ [60, 61, 62, 63]. Además como todas los ceros de p_n son simples y están contenidos en Δ [49], a menos que n sea muy pequeño incluso la evaluación numérica de $E(p_n)$ plantea serios problemas debido al comportamiento fuertemente oscilatorio del integrando de (2.56). En este sentido merece destacarse [64], que desarrolla un algoritmo eficiente para el cálculo numérico de $E(p_n)$ en el caso de que Δ sea un intervalo acotado.

Conviene aclarar que los valores de las integrales (2.56)–(3.51) dependen de la normalización escogida para los polinomios, lo que obviamente no ocurre para la entropía de las distribuciones de probabilidad dadas por (2.54). Este hecho no supone ningún problema desde el punto de vista práctico, puesto que una vez obtenido un resultado para una normalización es fácilmente traducible a cualquier otra. Para conocer las fórmulas explícitas que relacionan las entropías de posición y momento del oscilador armónico, el átomo de hidrógeno y otros sistemas cuánticos con $E(p_n)$ consúltense [56] y [59], éste último es un artículo de revisión muy completo que cubre todos los avances en el problema de la entropía de los polinomios ortogonales hasta el año 2001.

2.4. Entrelazamiento

El entrelazamiento es uno de los fenómenos más característicos y sorprendentes de la teoría cuántica, que, al igual, que el principio de incertidumbre no tiene análogo clásico. Este fenómeno aparece al considerar sistemas cuánticos compuestos. En esta tesis en general consideraremos el caso de sistemas bipartitos, que es el caso más sencillo, y a la vez suficientemente representativo, para el estudio de este fenómeno. De hecho, muchos de los

resultados de este caso se extienden fácilmente al escenario multipartito, si bien éste tiene sus particularidades. Así mismo, nos restringiremos al caso de sistemas cuánticos de dimensión finita, aunque el entrelazamiento también ha sido exhaustivamente considerado en el caso de dimensión infinita (lo que se conoce como entrelazamiento en variables continuas, véase por ejemplo [65]). Así pues, consideraremos sistemas cuánticos compuestos por dos subsistemas A y B con espacios de Hilbert asociados H_A y H_B tales que $\dim H_A = M$ y $\dim H_B = N$ donde $M, N \in \mathbb{N} \setminus \{1\}$. Además, sin pérdida de generalidad, consideraremos que $M \leq N$. $H = H_A \otimes H_B$ denotará el espacio de Hilbert del sistema conjunto. Para poder dar una primera idea del fenómeno y su relevancia consideremos una definición preliminar del entrelazamiento, sólo aplicable al caso de estados puros.

Definición 2.1. *Sea un estado bipartito puro caracterizado por $|\psi\rangle \in H$. Éste se dice separable si es un estado producto, i.e. si existen $|\phi\rangle_A \in H_A$ y $|\varphi\rangle_B \in H_B$ tales que*

$$|\psi\rangle = |\phi\rangle_A \otimes |\varphi\rangle_B. \quad (2.58)$$

Si el estado no admite esta factorización se dice entonces entrelazado.

La existencia del entrelazamiento es una consecuencia directa del principio de superposición y del Postulado 2. En una primera aproximación, podemos entender los estados entrelazados como aquellos en los que no tiene sentido considerar individualmente los subsistemas y que precisan ser descritos de forma conjunta. Mientras que los estados separables responden a situaciones sencillas en las que el estado de cada subsistema es fruto de una preparación local (i.e. independiente) y los resultados de medidas son estadísticamente independientes, los estados entrelazados son típicamente consecuencia de estados que interaccionaron conjuntamente en el pasado, de forma que, aunque ahora A y B puedan estar espacialmente separados, su descripción queda íntima e irrevocablemente unida (entrelazada).

Sea, por ejemplo, el estado de dos qubits (recuérdese (2.5))

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \quad (2.59)$$

Es sencillo demostrar que es imposible escribir este estado en la forma (2.58) y, por tanto, es entrelazado. Como explicamos en la sección 2.1.2, el estado de los subsistemas A y B (que subrayamos pueden estar separados por una distancia arbitrariamente grande), para sus respectivos observadores viene dado por la traza parcial, $\rho_A = \text{Tr}_B(|\psi\rangle\langle\psi|) = I_2/2$ y $\rho_B = \text{Tr}_A(|\psi\rangle\langle\psi|) = I_2/2$. Así pues, sorprendentemente, mientras que el estado total es puro y los observadores de A y B tienen máxima información acerca del estado del sistema conjunto, individualmente el estado de cada subsistema es mezcla y además proporcional a la identidad, es decir, se trata del estado maximalmente mezclado, en el que el resultado de cualquier medida está siempre sujeto a una máxima incertidumbre, de manera que cada observador tiene una máxima ignorancia acerca del estado de su correspondiente subsistema. De esta forma se sustenta una interpretación

holística de la teoría cuántica en la que la información acerca del todo no es igual a la suma de la información acerca de las partes.

El hecho de que el estado de los subsistemas sea mezcla caracteriza el entrelazamiento de los estados puros. O sea, se puede demostrar que ψ , es separable si y sólo si su número de Schmidt es 1 y, por tanto, si y sólo si sus reducciones son también puras (basta comprobar una reducción debido a que las reducciones son isoespectrales). Además, como más adelante consideraremos en rigor, intuitivamente se deduce que a mayor grado de mezcla de los subsistemas mayor grado de entrelazamiento del estado. El estado que acabamos de considerar forma parte de los conocidos como estados de Bell,

$$\begin{aligned} |\phi^\pm\rangle &= \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \\ |\psi^\pm\rangle &= \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle), \end{aligned} \tag{2.60}$$

que forman una base del espacio de Hilbert de dos qubits ($H \simeq \mathbb{C}^2 \otimes \mathbb{C}^2$). Todos los estados de Bell están máximamente entrelazados.

De acuerdo con la intuición clásica, el entrelazamiento es posiblemente la consecuencia más extraña y chocante de la Mecánica Cuántica. Schrödinger, de forma visionaria, fue el primero en reconocer la particularidad de este fenómeno y el que acuñó el término por el que hoy es conocido. En su serie de artículos de 1936 [66] ya escribía “I would not call that *one* but rather *the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought” y “Another way of expressing the peculiar situation is: the best possible knowledge of a whole does not necessarily include the best possible knowledge of all its parts, even though they may be entirely separate and therefore virtually capable of being best possibly known”. No sorprende que el entrelazamiento haya sido parte activa en las discusiones conceptuales sobre los fundamentos e interpretación de la teoría cuántica. Por ejemplo, el entrelazamiento (aunque todavía no había sido bautizado como tal) está en el núcleo del razonamiento que llevó a Einstein, Podolsky y Rosen (EPR) en 1935 [67] a argumentar que la Mecánica Cuántica no era una teoría completa, especulando con la posibilidad de una teoría más amplia realista⁴ cuyas variables, ocultas hasta ahora por alguna razón desconocida, darían una explicación de la estadística predicha por la teoría cuántica⁵. En la década de los 60 este debate, que muchos consideraban meras especulaciones filosóficas, quedó magistralmente cerrado por el Teorema de Bell [68]. En él se demuestra que la estadística predicha por cualquier teoría de variables ocultas local (es decir, en la que no sea posible transmitir información a velocidad mayor que la de la luz como exige la Teoría de la Relatividad) debe satisfacer una serie de desigualdades (hoy conocidas como desigualdades de Bell) que no se cumplen en Mecánica

⁴O sea, en el que las propiedades de los objetos tienen una entidad absoluta no dependiente de la interacción con un observador como en la teoría cuántica.

⁵Es decir, la Mecánica Cuántica sería a esta teoría lo que la Mecánica Estadística es a la Mecánica Clásica.

Cuántica. Remarcablemente, son precisamente los estados entrelazados los que permiten que se den violaciones de las desigualdades de Bell y que quede refutada la compatibilidad del realismo local y la teoría cuántica. En los años 80 se produjo un cambio de mentalidad en el que, dejando aparte las implicaciones conceptuales, se comenzó a pensar en formas positivas de explotar la rareza de la Mecánica Cuántica. Por ejemplo, se empezó a dar forma a la idea de un ordenador cuántico (i.e. un ordenador construido a una escala en la que su funcionamiento sea regido por las leyes de la teoría cuántica) ya que los trabajos de Deutsch y Jozsa [69, 70] mostraron que de esta forma se podrían resolver algunos problemas (académicos) de manera más eficiente que con un ordenador clásico. La revolución de la computación cuántica llegó en los 90 cuando Shor [71] desarrolló un algoritmo capaz de factorizar números primos en una escala de tiempo polinomial con el tamaño del número (mucho más rápido que en el caso clásico donde la escala es exponencial de forma que factorizar un número suficientemente grande exige una cantidad de tiempo inabarcable). El algoritmo cuántico de Grover [72] para la búsqueda en bases de datos también mostró una aceleración considerable con respecto al caso clásico. Se sabe que el entrelazamiento juega un papel importante en estos algoritmos y que, de hecho, es una condición necesaria (aunque no suficiente) para la aceleración de los algoritmos cuánticos con respecto a los clásicos (véase por ejemplo [73, 74, 75]). A su vez, en estas últimas décadas también se viene desarrollando con gran éxito la Teoría Cuántica de la Información que surgió a partir del desarrollo de los primeros protocolos de criptografía con sistemas cuánticos ([76, 77]), que se demuestran ser completamente seguros (i.e. imposibles de descifrar por terceras partes). Así mismo, también se han desarrollado protocolos de comunicación cuántica en los que la complejidad comunicacional se reduce enormemente con respecto al caso clásico (véase por ejemplo [78]) y en los que es posible la denominada codificación densa [79], que permite enviar una mayor cantidad de información que en el caso clásico. En este sentido, también merece ser destacado [80], que mostró la posibilidad del teletransporte cuántico, es decir, pasar el estado cuántico, que es desconocido, de un sistema a otro separado de él arbitrariamente. Destacablemente, el entrelazamiento es un ingrediente imprescindible para llevar a cabo estos revolucionarios procesos, lo que pone de manifiesto la necesidad de una teoría del entrelazamiento que explique su caracterización, cuantificación y manipulación.

2.4.1. El problema de la separabilidad. Antecedentes

De la sección anterior se deduce que la caracterización del entrelazamiento, poder responder con absoluta certeza a la pregunta de si un estado está entrelazado o no, es de interés fundamental para la teoría cuántica. Por un lado, ahondaría en una mejor descripción y comprensión teórica de este fenómeno clave en los fundamentos de la Mecánica Cuántica. Por otro lado, y desde un punto de vista más práctico, permitiría clarificar (por ejemplo en un contexto experimental) qué estados son útiles para un determinado protocolo en Teoría Cuántica de la Información o para computación cuántica. Para ello, en primer

lugar, necesitamos extender la Definición 2.1 al contexto más general de estados mezcla, lo cual fue desarrollado por primera vez por Werner en [81].

Definición 2.2. *Sea un estado bipartito arbitrario $\rho \in \mathcal{S}(H)$. Éste se dice separable si existen $\rho_i^A \in \mathcal{S}(H_A)$ y $\rho_i^B \in \mathcal{S}(H_B)$ tales que*

$$\rho = \sum_i p_i \rho_i^A \otimes \rho_i^B, \quad (2.61)$$

donde $0 \leq p_i \leq 1$ y $\sum_i p_i = 1$. Además, debido a la descomposición espectral de las matrices densidad, sin pérdida de generalidad podemos suponer que todas las ρ_i^A y ρ_i^B corresponden a estados puros. Si el estado ρ no admite esta descomposición se dice entonces entrelazado.

Efectivamente, (2.61) es la extensión natural de (2.58): los estados mezcla separables son mezclas estadísticas de estados producto. De forma análoga, los estados separables en este caso también responden a preparaciones locales en las que A y B van preparando independientemente $\rho_i^A \otimes \rho_i^B$ con probabilidad p_i . La principal diferencia en este caso reside en que, mientras que en el caso puro los estados separables no están correlacionados (los subsistemas siguen estadísticas independientes) siendo el entrelazamiento la única fuente de correlación, los estados mezcla separables sí están correlacionados por causa de la mezcla estadística. Estas correlaciones se denominan clásicas a diferencia de las correlaciones puramente cuánticas debidas al entrelazamiento. Volveremos sobre este tema en el capítulo 6 de esta tesis.

El subconjunto de $\mathcal{S}(H)$ formado por los estados separables lo denotaremos $\mathcal{D}(H)$. Nótese que es un conjunto convexo. Además, no es de medida cero en $\mathcal{S}(H)$ [82]. La caracterización de los estados entrelazados, que se reduce entonces a un problema de pertenencia a un conjunto convexo, se denomina problema de la separabilidad. Concretamente, se refiere a encontrar una forma operacional de distinguir los estados separables de los entrelazados. Por operacional, nos referimos a condiciones matemáticas sencillas que ofrezcan una forma práctica de comprobar esta propiedad ya que resulta tremendamente complicado, si no imposible, verificar directamente si un estado admite una descomposición del tipo (2.61). Notablemente, este problema resulta ser altamente complicado que, de hecho, recientemente se ha demostrado ser NP-duro [83]. A pesar de ello, pueden destacarse algunos acercamientos computacionales al problema en los que se han desarrollado algoritmos no triviales para su resolución (véase [84] y referencias allí citadas). Nuestro enfoque se mantendrá en la caracterización operacional, que, además de poder detectar el entrelazamiento o la separabilidad de un estado, da información acerca de la estructura matemática de los estados entrelazados y mejora la comprensión del fenómeno. No obstante la complejidad del problema, existe una considerable cantidad de literatura con criterios dirigidos a este fin. Revisamos a continuación los resultados más relevantes. Como, tanto a un nivel práctico como teórico, lo interesante son los estados con entrelazamiento, haremos especial hincapié en los que proporcionan condiciones suficientes para que un estado sea entrelazado (o, lo que es lo mismo, necesarias

para la separabilidad). No obstante, también existen trabajos que deducen interesantes condiciones suficientes para la separabilidad como [82] y [85] (que, por ejemplo, permiten aseverar que en un determinado experimento nunca se va a generar entrelazamiento [86]). Sin embargo, y como ilustración de la dificultad del problema, dado un estado que se sabe separable no es conocido como descomponerlo de acuerdo a (2.61) salvo en el caso $M = N = 2$ [87, 88]. Así mismo, también nos centraremos exclusivamente en criterios de aplicabilidad general aunque existen condiciones que permiten caracterizar el entrelazamiento en casos particulares como el de matrices densidad de bajo rango [89]. El lector también es remitido a los artículos de revisión [90], [91] y al más reciente [92].

El criterio PPT o de Peres-Horodecki

En [93] Peres introdujo el primer criterio operacional que ofrece una forma sistemática para la detección de estados entrelazados.

Definición 2.3. *La operación de trasposición parcial con respecto al subsistema A , T_A , se define como la operación lineal en $\mathcal{L}(H_A \otimes H_B)$ tal que $(X \otimes Y)^{T_A} = X^T \otimes Y$ para $X \in \mathcal{L}(H_A)$ e $Y \in \mathcal{L}(H_B)$. Es decir, es la operación de trasposición actuando únicamente en H_A . Así, sea $\rho \in \mathcal{S}(H_A \otimes H_B)$, si consideramos los elementos matriciales de ρ , $\rho_{ij,kl}$, al escribir el operador en, por ejemplo, la base canónica producto de $\mathcal{L}(H_A \otimes H_B)$, i.e.*

$$\rho = \sum_{i,k=0}^{M-1} \sum_{j,l=0}^{N-1} \rho_{ij,kl} |ij\rangle\langle kl|,$$

los elementos matriciales de ρ^{T_A} en esta base serán los correspondientes a la trasposición de la parte correspondiente a A , o sea, $\rho_{kj,il}$.

Teorema 2.2. *Si $\rho \in \mathcal{D}(H)$ entonces $\rho^{T_A} \in \mathcal{S}(H)$, lo que particularmente implica que es un operador semidefinido positivo. O sea, si ρ^{T_A} no es semidefinido positivo entonces ρ es entrelazado.*

Nótese que aunque la forma concreta de la trasposición parcial depende de la base escogida esto es irrelevante de cara a la aplicación del criterio porque los autovalores son independientes de la base. Igualmente, resulta irrelevante trasponer con respecto a A o B porque ρ^{T_A} y ρ^{T_B} tienen los mismos autovalores ya que $T = T_A \circ T_B$ y, por tanto, $\rho^{T_A} = (\rho^{T_B})^T$ (recuérdese que la trasposición no altera los autovalores). La demostración de este teorema es sumamente sencilla ya que si ρ es separable entonces admite la descomposición (2.61) y así $\rho^{T_A} = \sum_i p_i (\rho_i^A)^T \otimes \rho_i^B$, que es también un estado (separable además) puesto que si $\rho_i^A \in \mathcal{S}(H_A)$ entonces $(\rho_i^A)^T \in \mathcal{S}(H_A)$ también.

Este criterio se puede enunciar en una forma alternativa basada en la descomposición en valores singulares y la norma matricial en traza $\|\cdot\|_{tr}$ (véase el apéndice A) que utilizaremos con frecuencia más adelante. Puesto que los valores singulares de una matriz hermítica son el valor absoluto de sus autovalores, se tiene que $\|\rho\|_{tr} = 1$. Así, el Teorema 2.2 se puede enunciar diciendo que si ρ es

separable entonces $\|\rho^{TA}\|_{tr} = 1$ mientras que si $\|\rho^{TA}\|_{tr} > 1$ entonces el estado es entrelazado.

A pesar de la simplicidad en la demostración de este criterio, el Teorema 2.2 es profundo y muy relevante. En primer lugar, por su interpretación física que tiene que ver con el comportamiento de los estados bajo inversión temporal [90]. En segundo lugar, y más importante, este criterio es sumamente potente. De hecho, caracteriza el entrelazamiento cuando $M = 2$ y $N = 2, 3$ como mostró la familia Horodecki en [94]. Es decir, para estados en estas dimensiones la positividad⁶ bajo trasposición parcial implica la separabilidad. La demostración de este hecho es más complicada que la implicación en el otro sentido y utiliza como ingrediente fundamental el teorema de Hahn-Banach como veremos más adelante. Notablemente, para dimensiones mayores esto deja de ser cierto y la positividad bajo trasposición parcial es únicamente una condición necesaria para la separabilidad [94]. Es decir, si $MN > 6$ existen estados entrelazados con una traspuesta parcial positiva, que denominaremos estados entrelazados PPT. De ahí que este criterio se conozca como PPT o de Peres-Horodecki. Lógicamente, los esfuerzos posteriores a este resultado en el marco del problema de la separabilidad se han dirigido principalmente a encontrar criterios con una fuerte capacidad para detectar estados entrelazados PPT y así complementar al criterio de Peres-Horodecki. No obstante, por completitud, enunciaremos un par de criterios interesantes aunque estrictamente más débiles que el criterio PPT⁷. En este sentido se tiene el criterio de reducción [95, 96] (relevante en la destilación del entrelazamiento, tema que queda fuera de los objetivos de esta tesis):

Teorema 2.3. *Si ρ es separable entonces $\rho_A \otimes I_B - \rho$ y $I_A \otimes \rho_B - \rho$ son semidefinidos positivos.*

Como ya hemos visto vagamente en la introducción del entrelazamiento, este fenómeno está directamente relacionado con el concepto de información. De hecho, los conceptos y medidas introducidos en el capítulo 1 juegan un importante papel también en el estudio del entrelazamiento. Por ejemplo, este es el caso del criterio de la mayorización [97]:

Teorema 2.4. *Si ρ es separable entonces $\rho \prec \rho_A$ y $\rho \prec \rho_B$.*

Se entiende que una matriz mayoriza a otra si esta relación se da entre sus autovalores (nótese que los autovalores de una matriz densidad forman una distribución de probabilidad y que si las matrices tienen distinta dimensión se supone que se añaden ceros a la más pequeña). Este resultado explica con más generalidad la idea introducida anteriormente para estados puros de que un menor grado de incertidumbre a nivel global que a nivel local implica la presencia de entrelazamiento. Del criterio de mayorización se deduce que si una

⁶Este es el término habitual en la literatura aunque sería más correcto no negatividad.

⁷En lo que sigue diremos que un criterio es más débil que otro si no es capaz de detectar estados entrelazados que no sean detectados por el segundo y estrictamente más débil si además el segundo detecta estados entrelazados no reconocidos como tales por el primero. Análogamente utilizaremos los términos “más fuerte” y “estrictamente más fuerte”.

medida entrópica de las reducciones es mayor que la del estado total entonces se tiene entrelazamiento, por ejemplo, si $S(\rho_A), S(\rho_B) \geq S(\rho)$ (fenómeno irreproducible clásicamente ya que $H(X), H(Y) \leq H(X, Y)$). Véase [98] y referencias allí citadas para estudios exhaustivos de entropías generalizadas de los operadores densidad y criterios de entrelazamiento (muchos de estos resultados son anteriores a la deducción del Teorema 2.4, como, por ejemplo, [99]). Merece la pena también mencionarse que la mayorización juega un papel fundamental en otras facetas de la teoría del entrelazamiento como es el caso de las transformaciones de entrelazamiento (¿bajo qué condiciones es posible transformar localmente un estado entrelazado en otro?) [100, 101].

El criterio CCNR

El criterio CCNR fue introducido independientemente por Rudolph (“computable cross norm”) [102] y Chen y Wu (“realignment”) [103] con acercamientos completamente diferentes al problema. De hecho, este criterio se puede expresar en diversas formas equivalentes. Aquí desarrollaremos dos, las cuales utilizaremos más adelante en esta tesis. Una de ellas, particularmente interesante por su analogía con el criterio PPT, fue destacada en [104].

Definición 2.4. *La operación de realineamiento R es una operación lineal que lleva operadores de $\mathcal{L}(H_A \otimes H_B)$ a operadores que llevan operadores de $\mathcal{L}(H_B)$ a operadores de $\mathcal{L}(H_A)$ de forma que $R(X \otimes Y) = |X\rangle\langle Y^*|$ para $X \in \mathcal{L}(H_A)$ e $Y \in \mathcal{L}(H_B)$, donde $\langle \cdot | \cdot \rangle$ es el producto escalar de Hilbert-Schmidt en \mathcal{L} , i.e. $\langle X | Y \rangle = \text{Tr}(X^\dagger Y)$. Equivalentemente, podemos entender el realineamiento como una operación que lleva operadores de $\mathcal{L}(H_A \otimes H_B)$ a operadores de H_B en H_A tal que si*

$$\rho = \sum_{i,k=0}^{M-1} \sum_{j,l=0}^{N-1} \rho_{ij,kl} |ij\rangle\langle kl|,$$

entonces

$$R(\rho) = \sum_{i,j=0}^{M-1} \sum_{k,l=0}^{N-1} \rho_{ik,jl} |ij\rangle\langle kl|.$$

Así, desde un punto de vista matricial las operaciones de trasposición parcial y realineamiento se reducen a reordenamientos particulares de las entradas de ρ .

Teorema 2.5. *Si $\rho \in \mathcal{D}(H)$ entonces $\|R(\rho)\|_{tr} \leq 1$. O sea, si $\|R(\rho)\|_{tr} > 1$ entonces el estado es entrelazado.*

Al igual que el criterio PPT, el Teorema 2.5 caracteriza el entrelazamiento en estados puros. Por contra, a diferencia del criterio PPT, el criterio CCNR no es una condición necesaria para el entrelazamiento en ninguna dimensión. Sin embargo, los criterios son independientes, y el Teorema 2.5 reconoce mucho estados PPT entrelazados. De esta forma, este criterio complementa adecuadamente al de Peres-Horodecki y nuestra capacidad para identificar

estados entrelazados se incrementa considerablemente con el uso conjunto de estos dos resultados.

Testigos de entrelazamiento y aplicaciones no completamente positivas

En esta subsección presentamos dos caracterizaciones no operacionales del entrelazamiento (aunque no por ello ni mucho menos inútiles). Puesto que $\mathcal{D}(H)$ es un conjunto convexo y compacto, intuitivamente, parece claro que dado cualquier elemento de $\mathcal{S}(H) \setminus \mathcal{D}(H)$ existe un hiperplano que separa a este elemento de $\mathcal{D}(H)$. Esta idea se desarrolla en rigor utilizando el teorema de Hahn-Banach geométrico dando lugar a una interesante caracterización del conjunto de los estados entrelazados [94].

Definición 2.5. *Un testigo de entrelazamiento es un operador $W \in \mathcal{B}(H_A \otimes H_B)$ (y, por tanto, un observable) tal que sin ser semidefinido positivo satisface $\text{Tr}(W|\psi\rangle\langle\psi| \otimes |\phi\rangle\langle\phi|) \geq 0 \forall |\psi\rangle \in H_A \text{ y } \forall |\phi\rangle \in H_B$.*

Teorema 2.6. *Un estado ρ es separable si y sólo si $\text{Tr}(W\rho) \geq 0$ para cualquier testigo de entrelazamiento W .*

Por consiguiente, dado un estado no separable existirá al menos un testigo de entrelazamiento para el cual $\text{Tr}(W\rho) < 0$, de forma que quede al descubierto el entrelazamiento del estado. De cara a la aplicación práctica de este resultado el principal problema radica en que los testigos de entrelazamiento deben ser escogidos astutamente en cada caso. De hecho, no es accesible construir el conjunto de todos los testigos $W \in \mathcal{B}(H)$, así que para tener una versión operacional de este criterio la única solución posible es fijar a priori el conjunto de testigos que se va a utilizar ($\{W_i\}$) teniendo de esta manera un condición suficiente para el entrelazamiento ($\text{Tr}(W_i\rho) < 0$ para al menos un i). En la literatura se ha estudiado en profundidad como escoger los mejores testigos (los que detectan más estados entrelazados) o como construir testigos que detecten entrelazamiento PPT (véase, por ejemplo, [105]). Si bien criterios como el PPT o CCNR tienen la ventaja de su sencilla aplicación, los testigos de entrelazamiento pueden ser más convenientes de cara a la implementación experimental cuando se desconoce el estado concreto del sistema cuántico en cuestión. Esto se debe a que mientras que los primeros requieren obviamente un conocimiento absoluto de la matriz densidad que caracteriza al sistema, la aplicación del Teorema 2.6 sólo exige el conocimiento del valor esperado de un observable (recuérdese que $\langle W \rangle = \text{Tr}(W\rho)$) lo cual suele requerir un número de medidas locales menor que la determinación de ρ .

El isomorfismo de Jamiolkowski [106], que conecta operadores de $\mathcal{B}(H)$ y aplicaciones (lineales) actuando en $\mathcal{B}(H)$, permite también caracterizar el entrelazamiento en términos de las aplicaciones positivas pero no completamente positivas [94].

Definición 2.6. *Un endomorfismo actuando en $\mathcal{B}(H)$ se dice positivo si lleva operadores semidefinidos positivos a operadores semidefinidos positivos.*

Definición 2.7. *Un endomorfismo positivo Λ que actúa en $\mathcal{B}(H_A)$ se dice completamente positivo si su extensión trivial para que actúe en $\mathcal{B}(H_A \otimes H_B)$ (o sea, $\Lambda \otimes I_B$, donde I_B es la aplicación identidad actuando en $\mathcal{B}(H_B)$) es positiva.*

Teorema 2.7. *Un estado $\rho \in \mathcal{S}(H_A \otimes H_B)$ es separable si y sólo si $\Lambda \otimes I_B(\rho)$ es un operador semidefinido positivo para cualquier aplicación positiva (pero no necesariamente completamente positiva) Λ en $\mathcal{B}(H_A)$.*

En conclusión, dado un estado entrelazado siempre existe una aplicación positiva pero no completamente positiva que mapea el operador densidad del sistema a un operador que no es semidefinido positivo. Así, el problema de la separabilidad es equivalente a la inconmesurable tarea de clasificación de todas las aplicaciones positivas pero no completamente positivas, problema que ya había sido estudiado en la literatura matemática con anterioridad al “boom” del entrelazamiento (véase por ejemplo [107, 108]). Sin embargo, esta caracterización permite entender en profundidad criterios anteriormente citados como el PPT, ya que la trasposición es una aplicación positiva pero no completamente positiva. Es más, resulta que cuando $H_A \simeq \mathbb{C}^2$ y $H_B \simeq \mathbb{C}^2, \mathbb{C}^3$ todas las aplicaciones de este tipo se pueden descomponer en términos de la trasposición [108, 109], lo que permitió a los Horodecki demostrar que el criterio PPT da también una condición necesaria para el entrelazamiento en este caso como ya mencionamos.

2.4.2. Separabilidad y relaciones de incertidumbre

Remarcablemente, existe una conexión entre los dos objetos de estudio en esta tesis, el principio de incertidumbre y el entrelazamiento, y las relaciones de incertidumbre constituyen una herramienta útil en el marco del problema de la separabilidad. La idea que permite establecer esta relación consiste en considerar conjuntos de observables que posean autoestados entrelazados o, en general, que no tengan autoestados simultáneos separables. De esta forma, mientras que la incertidumbre conjunta en el resultado de una medida de los operadores puede ser arbitrariamente pequeña si los estados son entrelazados, existirá un cota inferior a esta incertidumbre para los estados separables. La relación entre el entrelazamiento y el principio de incertidumbre (o más concretamente, cómo obtener condiciones de separabilidad a partir de relaciones de incertidumbre) fue inicialmente explorada en el caso de sistemas cuánticos de variables continuas (véase [110] por ejemplo) mediante la relación de Kennard (2.22). Esta idea se desarrolló por primera vez en el caso de sistemas cuánticos de dimensión finita en [111] donde se demuestra lo siguiente:

Teorema 2.8. *Sean $\{A_i\}$ y $\{B_i\}$ observables de $\mathcal{B}(H_A)$ y $\mathcal{B}(H_B)$ respectivamente, que satisfagan relaciones de incertidumbre*

$$\sum_i \Delta_\rho^2(A_i) \geq C_A, \quad \sum_i \Delta_\rho^2(B_i) \geq C_B, \quad (2.62)$$

entonces, todos los estados separables $\rho \in \mathcal{D}(H_A \otimes H_B)$ cumplen

$$\sum_i \Delta_\rho^2(A_i \otimes I_B + I_A \otimes B_i) \geq C_A + C_B. \quad (2.63)$$

Dicho de otra forma, mientras que no hay restricción a la incertidumbre en la medida de los observables en estados entrelazados, los estados separables heredan las relaciones de incertidumbre satisfechas a nivel local (i.e., por los subsistemas). De ahí que este criterio se conozca como de las relaciones de incertidumbre locales (LURs). Esta es otra manifestación de la ya comentada idea de que los estados entrelazados son los únicos que pueden tener un menor grado de incertidumbre a nivel global que local. Además de ofrecer una nueva ventana para el estudio del problema de la separabilidad, este criterio (como los que veremos a continuación) tiene el atractivo de formularse en términos de valores esperados de observables, por lo que, al igual que en el caso de los testigos de entrelazamiento, puede ser más adecuado para la implementación experimental que los que requieren un conocimiento total del estado del sistema, siempre y cuando el número de medidas exigidas por el criterio sea menor a las $\dim H - 1$ necesarias para determinar ρ . Sin embargo, como también sucede con los testigos, el Teorema 2.8 adolece de una falta de operacionalidad puesto que no queda claro cual es la elección adecuada en cada caso como conjunto de observables. Es más, nótese que la medida de incertidumbre viene dada por la desviación estándar, lo cual lleva a la relación de incertidumbre de Heisenberg-Robertson (2.25). Sin embargo, para poder aplicar el criterio también a estados mezcla, éste se formula a partir de relaciones de incertidumbre para la suma de desviaciones, no el producto como en (2.25). Esto implica que no existe una receta para obtener directamente C_A y C_B en (2.62) dados $\{A_i\}$ y $\{B_i\}$, lo que significa que cada caso debe ser trabajado desde cero una vez escogidos los conjuntos de observables. En [111] se demuestra que

$$\Delta_\rho^2(S_x) + \Delta_\rho^2(S_y) + \Delta_\rho^2(S_z) \geq \frac{N-1}{2}, \quad (2.64)$$

donde S_i denota el observable de spin N -dimensional a lo largo del eje i [14]. Así, se tiene una condición necesaria para la separabilidad de sistemas cuánticos bipartitos $N \times N$:

$$\sum_{i=x,y,z} \Delta_\rho^2(S_i^A \otimes I_B + I_A \otimes S_i^B) \geq N-1. \quad (2.65)$$

Por ejemplo, en el caso de dos qubits esto se reduce a

$$\sum_{i=x,y,z} \Delta_\rho^2(\sigma_i^A \otimes I_B + I_A \otimes \sigma_i^B) \geq 4, \quad (2.66)$$

donde $\{\sigma_x, \sigma_y, \sigma_z\}$ son las famosas matrices de Pauli,

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.67)$$

directamente relacionadas con el observable de spin en sistemas cuánticos de dos niveles. Más relaciones de este tipo se obtuvieron en [112, 25]. Destacablemente, en estos artículos se demuestra que el criterio de las LURs es capaz de detectar entrelazamiento PPT. Una elección particularmente interesante de observables se desarrolló en [113]: la de observables locales ortogonales (LOOs) [114] que denotaremos $\{G_i^A\}$ y $\{G_i^B\}$. Los LOOs son bases ortonormales de $\mathcal{B}(H_A)$ y $\mathcal{B}(H_B)$. En [113] se demuestra que para cualquier conjunto de LOOs se tiene que

$$\sum_{i=1}^{M^2} \Delta_\rho^2(G_i^A) \geq M - 1, \quad \sum_{i=1}^{N^2} \Delta_\rho^2(G_i^B) \geq N - 1, \quad (2.68)$$

lo que implica que si $\rho \in \mathcal{D}(H_A \otimes H_B)$ entonces

$$\sum_{i=1}^{N^2} \Delta_\rho^2(G_i^A \otimes I_B + I_A \otimes G_i^B) \geq M + N - 2 \quad (2.69)$$

para cualquier conjunto de LOOs $\{G_i^A\}$ y $\{G_i^B\}$ de $\mathcal{B}(H_A)$ y $\mathcal{B}(H_B)$ respectivamente. Si $M < N$ se entiende que $G_i^A = 0$ para $M^2 + 1 \leq i \leq N^2$. La principal importancia de la condición de las LURs formulada en términos de LOOs reside en que el criterio de separabilidad dado por (2.69) es no sólo capaz de detectar estados PPT entrelazados sino en que es estrictamente más fuerte que el criterio CCNR [113]. Aún así, la aplicación del criterio requiere de cierta dosis de intuición para elegir el conjunto de LOOs más apropiado en cada caso.

Una pregunta natural que surge en este contexto es la posibilidad de construir también criterios de separabilidad a partir de relaciones de incertidumbre entrópicas. Como hemos visto, éstas dan una formulación más satisfactoria del principio de incertidumbre y tienen una estructura matemática más trabajada y manejable, por lo que en caso de poder establecerse un análogo del Teorema 2.8 sería inmediato obtener condiciones de separabilidad del tipo (2.63) para conjuntos de pares de observables a partir de la EUR de Maassen-Uffink (2.39). Esto sería además posible para conjuntos de más observables debido a que existen generalizaciones no triviales de esta EUR para conjuntos de más de dos observables en el caso de observables complementarios [115, 116, 117]. En [118] se investigó por primera vez la posibilidad de criterios de separabilidad a partir de relaciones de incertidumbre entrópicas básicamente mediante exploraciones numéricas considerando diversos observables significativos. Esta relación se estableció en rigor en [119].

Lema 2.1. *Sea $\rho = \rho_A \otimes \rho_B$ un estado producto de $\mathcal{S}(H_A \otimes H_B)$, y sean A y B observables de $\mathcal{B}(H_A)$ y $\mathcal{B}(H_B)$ con autovalores no nulos. Entonces, se tiene que*

$$\mathcal{P}(A \otimes B, \rho) \prec \mathcal{P}(A, \rho_A), \quad \mathcal{P}(A \otimes B, \rho) \prec \mathcal{P}(B, \rho_B), \quad (2.70)$$

donde $\mathcal{P}(X, \rho)$ es la distribución de probabilidad asociada a la medida del observable X en el estado ρ .

De este lema se sigue entonces un análogo del Teorema 2.8 para cualquier medida entrópica, por ejemplo, la entropía de Shannon.

Teorema 2.9. Sean $\{A_i\}$ y $\{B_i\}$ conjuntos de observables con autovalores no nulos de $\mathcal{B}(H_A)$ y $\mathcal{B}(H_B)$ respectivamente, para uno de los cuales se tenga al menos una relación de incertidumbre entrópica

$$\sum_i H_\rho(A_i) \geq C, \quad (2.71)$$

entonces, todos los estados separables $\rho \in \mathcal{D}(H_A \otimes H_B)$ cumplen

$$\sum_i H_\rho(A_i \otimes B_i) \geq C. \quad (2.72)$$

Así, como en el caso de la suma de desviaciones estándar, los estados separables heredan las EURs de sus subsistemas. A partir de (2.39) y sus extensiones para conjuntos de más de dos observables, se obtiene ahora directamente la C adecuada en cada caso. Por ejemplo, puesto que para estados en dos dimensiones se tiene que [115]

$$\sum_{i=x,y,z} H_\rho(\sigma_i) \geq 2 \log 2, \quad (2.73)$$

entonces, todos los estados separables de dos qubits deben satisfacer

$$\sum_{i=x,y,z} H_\rho(\sigma_i^A \otimes \sigma_i^B) \geq 2 \log 2, \quad (2.74)$$

condición análoga a (2.66). Desgraciadamente, las exploraciones llevadas a cabo en [119] muestran que las condiciones de separabilidad provenientes de relaciones de incertidumbre entrópicas son más débiles que sus hermanas basadas en sumas de desviaciones estándar, incluso cuando consideramos entropías generalizadas en vez de únicamente la de Shannon. Así, sorprendentemente, aunque las EURs dan una expresión matemática más rigurosa del principio de incertidumbre, las relaciones basadas en la desviación estándar parecen más fuertes para la detección del entrelazamiento.

2.4.3. Medidas de entrelazamiento

Una teoría completa del entrelazamiento no se limita sólo a su caracterización. Desde el punto de vista de la Teoría de la Información Cuántica, el entrelazamiento es una fuente que permite llevar a cabo procesos inaccesibles clásicamente. Por ello, precisa desarrollar una teoría que explique satisfactoriamente como manipular este valioso recurso. Esto lleva, entre otras cosas, a elaborar formas rigurosas de cuantificar el grado de entrelazamiento presente en los estados cuánticos. Por un lado, esto ofrece una clasificación físicamente motivada del entrelazamiento, y, por otro, permite discernir la capacidad de los distintos estados de cara a llevar a cabo un determinado protocolo en Información Cuántica, así como establecer una calidad para los estados entrelazados generados en experimentos. Esta tarea pasa por un

desarrollo similar al del Capítulo 1, en el que se estudió como construir medidas de la incertidumbre, que tiene como objetivo ahora el diseño de medidas de entrelazamiento (de hecho, veremos que ambas medidas tienen muchos puntos en común).

Los caminos para la obtención de este tipo de medidas son, como veremos, múltiples. Sin embargo, existe un paradigma unificador que fija la estructura de una medida de entrelazamiento razonable: el de las operaciones cuánticas locales y comunicación clásica (LOCC). Como ya hemos discutido, el entrelazamiento es un fenómeno de naturaleza no local. Es decir, su existencia se debe a una interacción a nivel global en un conjunto de subsistemas que nunca puede surgir operando localmente (i.e., actuando independientemente sobre los subsistemas). Por ello, si los observadores correspondientes a cada subsistema sólo realizan medidas de forma local, aunque se les permita comunicar sus resultados y coordinar en base a ellos las medidas utilizando un canal de transmisión clásico (por ejemplo, un teléfono), nunca se generará entrelazamiento. Este tipo de operaciones es lo que se conoce como LOCC, y son las más generales posibles incapaces de crear entrelazamiento. Por ello, un estado obtenido a partir de otro mediante un protocolo LOCC determinista no podrá contener más entrelazamiento que el estado original. Si el protocolo LOCC es estocástico (SLOCC), es decir, se pueden obtener diversos estados, cada uno con una determinada probabilidad, entonces se debe satisfacer que el entrelazamiento en media de los estados finales no sea mayor al original. Las medidas de entrelazamiento deben satisfacer por tanto esta propiedad. Consecuentemente, es necesario caracterizar matemáticamente el conjunto de las operaciones LOCC. Esta tarea resulta ser extremadamente difícil, por lo que en la práctica se consideran un conjunto de operaciones más amplio que contiene al de las LOCC. Este conjunto se denomina de operaciones separables y está formado por las medidas generalizadas cuyos operadores de Kraus factorizan, o sea, cuando el conjunto de todos los operadores de Kraus que constituyen el POVM $\{F_k \in \mathcal{L}(H_A \otimes H_B)\}$ ($\sum_k F_k^\dagger F_k = I$) tienen una descomposición producto tal que $F_k = F_k^A \otimes F_k^B$ con $F_k^A \in \mathcal{L}(H_A)$ y $F_k^B \in \mathcal{L}(H_B)$ (lo que implica $\sum_k (F_k^A)^\dagger F_k^A = I_A$ y $\sum_k (F_k^B)^\dagger F_k^B = I_B$). Los $\{F_k^A\}$ y $\{F_k^B\}$ se denominan a veces filtros locales. No sorprende que en este contexto no se limite uno a considerar medidas proyectivas ya que se busca la mayor generalidad posible y nada impide que los observadores añadan sistemas auxiliares para realizar las medidas y luego los descarten llevando a cabo así medidas generalizadas. El hecho de que los operadores de Kraus factoricen responde a la exigencia de que las medidas sean locales.

En conclusión, este es entonces el conjunto de axiomas que debe satisfacer una medida de entrelazamiento E [120, 121]:

- (1) $E(\rho)$ es una aplicación que lleva operadores densidad $\rho \in \mathcal{S}(H_A \otimes H_B)$ a números reales no negativos.
- (2) $E(\rho) = 0$ si, y sólo si, ρ es separable.

(3) E no aumenta en media bajo operaciones LOCC,

$$E(\rho) \geq \sum_k p_k E \left(\frac{F_k^A \otimes F_k^B \rho (F_k^A)^\dagger \otimes (F_k^B)^\dagger}{\text{Tr}(F_k^A \otimes F_k^B \rho (F_k^A)^\dagger \otimes (F_k^B)^\dagger)} \right), \quad (2.75)$$

donde $F_k^A \in \mathcal{L}(H_A)$, $F_k^B \in \mathcal{L}(H_B)$, $\sum_k (F_k^A)^\dagger F_k^A \otimes (F_k^B)^\dagger F_k^B = I_A \otimes I_B$ y $p_k = \text{Tr}(F_k^A \otimes F_k^B \rho (F_k^A)^\dagger \otimes (F_k^B)^\dagger)$ es la probabilidad de obtener el resultado k -ésimo en la medida generalizada representada por $\{F_k^A \otimes F_k^B\}$.

Nótese que si consideramos un POVM constituido por un único operador de Kraus, que debe ser entonces unitario, $U = U_A \otimes U_B$, el protocolo es entonces invertible y, por tanto,

$$E(\rho) = E(U_A \otimes U_B \rho U_A^\dagger \otimes U_B^\dagger) \quad (2.76)$$

para matrices unitarias $U_A \in \mathcal{L}(H_A)$ y $U_B \in \mathcal{L}(H_B)$. Dicho de otra forma, la cantidad de entrelazamiento es invariante bajo transformaciones unitarias locales. Resultado no sorprendente puesto que estas transformaciones equivalen a un cambio de base local, i.e., ρ y $U_A \otimes U_B \rho U_A^\dagger \otimes U_B^\dagger$ representan el mismo estado físico.

Muchos autores también exigen a E que sea una medida convexa, o sea,

$$\sum_k p_k E(\rho_k) \geq E\left(\sum_k p_k \rho_k\right). \quad (2.77)$$

No obstante, esta condición no es estrictamente necesaria [122], aunque sí matemáticamente muy conveniente y todas las medidas que consideremos aquí satisfarán esta propiedad.

Resulta que no existe una única medida que satisfaga los axiomas (1)-(3). Esto se debe a que, como ya brevemente mencionamos, el hecho de que un estado se pueda transformar en otro mediante LOCC depende de que se dé una relación de mayorización [101]. Por lo tanto, al igual que sucedía con las medidas de incertidumbre en el Capítulo 1, no existe un orden total en el conjunto de los estados entrelazados de acuerdo al paradigma LOCC y, de esta manera, existen diferentes medidas de entrelazamiento que cumplen (1)-(3) y cada cual induce un orden distinto. Las distintas medidas que se conocen se construyen a partir de diferentes motivaciones para luego verificar que satisfacen las condiciones (1)-(3). Por ejemplo, se tienen medidas construidas a partir de claras interpretaciones físicas, otras de carácter geométrico (como distancia al conjunto de los estados separables) y otras son puramente de carácter axiomático. En la etapa presente no existe ninguna medida preferible a otra y todas aportan información de una u otra forma. Véase [123] para una introducción muy completa y actual a las medidas de entrelazamiento, la cual considera una gran variedad de medidas. Aquí, sólo mostraremos algunos ejemplos de medidas, centrándonos particularmente en las que tendremos en cuenta más adelante.

Ejemplos de medidas

Una posibilidad a la hora de obtener una medida del entrelazamiento es considerar algún tipo de distancia en $\mathcal{S}(H_A \otimes H_B)$. Cabría esperar, entonces, que cuanto más alejado esté un estado del conjunto de los estados separables $\mathcal{D}(H_A \otimes H_B)$, mayor será su grado de entrelazamiento. Así se verifica, efectivamente, cuando uno considera la entropía relativa cuántica:

$$S(\rho||\sigma) \equiv \text{Tr}(\rho \log \rho - \rho \log \sigma) \quad (2.78)$$

para $\rho, \sigma \in \mathcal{S}(H_A \otimes H_B)$, aunque debe subrayarse que $S(\rho||\sigma)$ no constituye rigurosamente una distancia en el sentido matemático, ya que ni siquiera es una medida simétrica. Sin embargo, en un sentido laxo en el contexto de la Teoría de Información esta cantidad está considerada como una distancia entre distribuciones de probabilidad [5]. Utilizando la entropía relativa cuántica se define la siguiente medida del entrelazamiento, denominada entropía relativa de entrelazamiento [120, 121],

$$E_R(\rho) = \min_{\sigma \in \mathcal{D}(H)} S(\rho||\sigma). \quad (2.79)$$

Otra posibilidad consiste en definir una medida de entrelazamiento en el caso más fácil de entender de los estados puros, para luego extenderla de una manera apropiada a los estados mezcla. Como ya hemos analizado, un estado puro está más entrelazado cuanto más mezcla sean sus reducciones (mayor desorden a nivel local). Por consiguiente, una medida de la pureza de sus reducciones es un buen candidato para una medida de entrelazamiento. Por ejemplo, puede utilizarse la entropía de Von Neumann S dada por (2.4) lo que da lugar a la medida conocida como entrelazamiento de formación E_F [124],

$$E_F(|\psi\rangle) = S(\rho_A) = -\text{Tr}(\rho_A \log \rho_A), \quad (2.80)$$

que además tiene una interesante interpretación física [124, 123]. Como hemos visto, los estados puros también se caracterizan por $\text{Tr}(\rho^2) = 1$. Este punto de vista da lugar a la medida conocida como concurrencia C [125],

$$C(|\psi\rangle) = \sqrt{2(1 - \text{Tr}\rho_A^2)}. \quad (2.81)$$

A veces resulta conveniente extraer la raíz de la anterior fórmula lo que lleva a considerar la medida conocida como tangle τ ,

$$\tau(|\psi\rangle) = C^2(|\psi\rangle) = 2(1 - \text{Tr}\rho_A^2). \quad (2.82)$$

En definitiva, mientras en el entrelazamiento de formación se usa la entropía de Shannon para medir el desorden a nivel local, en la concurrencia y el tangle se utiliza la entropía de Tsallis de parámetro 2.

En [126] se estudia de una forma sistemática cómo crear medidas de entrelazamiento rigurosas considerando exclusivamente sus propiedades en

estados puros. Éstas son verificadas por las anteriores medidas y la forma adecuada de extenderlas a estados mezcla es mediante el “techo convexo”, que es el mínimo valor en media de la medida sobre todas las posibles mezclas estadísticas de estados puros que dan lugar al estado. Así, las anteriores medidas para estados generales se definen como

$$E_F(\rho) = \min_{\{p_i, |\psi_i\rangle\}} \left\{ \sum_i p_i E_F(\psi_i) : \rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \right\}, \quad (2.83)$$

$$C(\rho) = \min_{\{p_i, |\psi_i\rangle\}} \left\{ \sum_i p_i C(\psi_i) : \rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \right\}, \quad (2.84)$$

$$\tau(\rho) = \min_{\{p_i, |\psi_i\rangle\}} \left\{ \sum_i p_i C^2(\psi_i) : \rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \right\}. \quad (2.85)$$

Como muestran los ejemplos presentados, en general la evaluación explícita de una medida de entrelazamiento se reduce a un problema de optimización tremendamente complicado de resolver ya no sólo analíticamente sino incluso numéricamente. Esto no es sorprendente dado que poder calcular una de estas medidas para cualquier estado equivaldría a resolver el problema de la separabilidad debido al axioma (2), y, como ya dijimos, este es un problema NP-duro. A pesar de esto, debe señalarse que existe algún algoritmo útil [127] para su evaluación numérica. Por otro lado, estas medidas únicamente se han podido evaluar explícitamente en forma analítica en el caso de conjuntos de estados con un elevado grado de simetría (véase por ejemplo [128, 129, 130, 131]). Una muy destacable excepción aparece en el caso de las íntimamente relacionadas E_F , C y τ para las cuales se conoce una expresión general cerrada en el caso de sistemas cuánticos 2×2 [87],

$$C(\rho) = \max(0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4), \quad (2.86)$$

donde los $\{\lambda_i\}$ son la raíz cuadrada de los autovalores en orden decreciente de la matriz $\rho\hat{\rho}$ con

$$\hat{\rho} = \sigma_y \otimes \sigma_y \rho^* \sigma_y \otimes \sigma_y, \quad (2.87)$$

donde la conjugación $*$ se toma en la base de los autoestados de σ_z . A su vez,

$$E_F(\rho) = h \left(\frac{1 + \sqrt{1 - C^2(\rho)}}{2} \right), \quad (2.88)$$

donde h es la entropía binaria, $h(p) = H(\{p, 1 - p\})$. Así mismo se puede comprobar (ver [132]) que en el caso de dos qubits $\tau(\rho) = C^2(\rho)$ (aunque en el caso general simplemente $\tau(\rho) \geq C^2(\rho)$). Algunos autores prefieren el tangle a la concurrencia porque, para éste también se conoce una fórmula explícita para matrices densidad de rango 2 [132] y, además, contrariamente a la concurrencia, su comportamiento es análogo al entrelazamiento de formación para ciertos estados en dimensiones mayores [130].

Por último, merece la pena mencionarse que con el fin de obtener una medida fácil de calcular, se permite relajar el axioma (2) a

(2') $E(\rho) = 0$ si ρ es separable.

En este caso, poder evaluar una medida no equivale a resolver el problema de la separabilidad ya que se permite a los estados entrelazados tener medida de entrelazamiento nula, y es posible entonces construir medidas fácilmente computables. Tal es el caso de la negatividad [133],

$$N(\rho) = \frac{\|\rho^{TA}\|_{tr} - 1}{2}, \quad (2.89)$$

directamente relacionada con el criterio PPT, que si bien se puede evaluar sencillamente para cualquier estado es incapaz de cuantificar el entrelazamiento PPT.

Parte II

Incertidumbre

Capítulo 3

Entropía de polinomios ortogonales

Los polinomios de Jacobi de grado n , $P_n^{(\alpha, \beta)}$, que dependen de dos parámetros α, β , se definen como

$$P_n^{(\alpha, \beta)}(x) = \frac{(\alpha + 1)_n}{n!} {}_2F_1\left(\begin{matrix} -n, n + \alpha + \beta + 1 \\ \alpha + 1 \end{matrix}; \frac{1-x}{2}\right). \quad (3.1)$$

donde $(a)_n = \Gamma(a+n)/\Gamma(a)$ denota el símbolo de Pochhammer y ${}_2F_1$ la función hipergeométrica,

$${}_2F_1\left(\begin{matrix} a, b \\ c \end{matrix}; x\right) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n n!} x^n. \quad (3.2)$$

Cuando $\alpha, \beta > -1$ estos polinomios son ortogonales en $[-1, 1]$ con respecto a la función peso $(1-x)^\alpha(1+x)^\beta$

$$\begin{aligned} & \int_{-1}^1 P_n^{(\alpha, \beta)}(x) P_m^{(\alpha, \beta)}(x) (1-x)^\alpha (1+x)^\beta dx \\ &= \frac{2^{\alpha+\beta+1} \Gamma(n+\alpha+1) \Gamma(n+\beta+1)}{(2n+\alpha+\beta+1) n! \Gamma(n+\alpha+\beta+1)} \delta_{nm}. \end{aligned} \quad (3.3)$$

Un importante caso particular de estos polinomios son los ya referidos de Gegenbauer o ultrasféricos $C_n^{(\lambda)}$,

$$C_n^{(\lambda)}(x) = \frac{(2\lambda)_n}{(\lambda + \frac{1}{2})_n} P_n^{(\lambda - \frac{1}{2}, \lambda - \frac{1}{2})}(x), \quad (3.4)$$

que para $\lambda > -1/2$ satisfacen, por tanto, la relación de ortogonalidad (2.52) mencionada en la introducción. En este capítulo intentaremos desarrollar fórmulas analíticas para la entropía de información (2.56) de los polinomios

de Jacobi,

$$E(P_n^{(\alpha,\beta)}) = - \int_{-1}^1 (P_n^{(\alpha,\beta)}(x))^2 \log (P_n^{(\alpha,\beta)}(x))^2 (1-x)^\alpha (1+x)^\beta dx, \quad (3.5)$$

haciendo particular énfasis en el caso particular de los Gegenbauer,

$$E(C_n^{(\lambda)}) = - \int_{-1}^1 (C_n^{(\lambda)}(x))^2 \log (C_n^{(\lambda)}(x))^2 (1-x^2)^{\lambda-\frac{1}{2}} dx. \quad (3.6)$$

Como hemos explicado en la introducción, estas integrales controlan la parte angular de las entropías de información en posición y momento de cualquier sistema cuántico D -dimensional en un potencial central, como por ejemplo el oscilador armónico isotrópico o el átomo de hidrógeno (potencial de Coulomb). Esto se debe a la conexión directa que hay entre los armónicos hiperesféricos, que dan la correspondiente función de onda angular, y los polinomios de Gegenbauer de parámetro entero y semientero (véase (2.51) para el caso tridimensional y, por ejemplo, [49, 134] para el caso general). Por ello, evaluar $E(C_n^{(\lambda)})$ en estos casos es especialmente relevante. No obstante, también son de interés otros casos puesto que los polinomios de Gegenbauer también aparecen en la componente radial de la función de onda en el espacio de momentos del átomo de hidrógeno D -dimensional y controlan la función de onda de algunos potenciales destacados como el de Poschl-Teller [49]. Así mismo, los polinomios de Jacobi fuera del caso particular de Gegenbauer también constituyen la función de onda de sistemas asociados a potenciales relevantes, tales como el de Woods-Saxon [135] o el de Hulthén [136].

Como se mencionó en la introducción, evaluar explícitamente las integrales (2.56) correspondientes a la entropía de una familia de polinomios ortogonales es un problema altamente complicado. De hecho, esto sólo se ha conseguido para unos pocos casos particulares de los polinomios de Gegenbauer. Puesto que la entropía de esta familia ha sido la única que ha dado muestras de ser tratable analíticamente resulta razonable centrar en ella por el momento los esfuerzos matemáticos dirigidos a obtener fórmulas exactas para las entropías de polinomios. Por esta razón, aparte del interés práctico recién comentado, nos limitamos aquí a considerar esta familia, aunque, como veremos, posteriormente extenderemos nuestras ideas a la clase más amplia de Jacobi. Antes de desarrollar nuestros resultados, exponemos brevemente los antecedentes en esta materia.

En lugar de considerar la normalización estándar de los polinomios de Gegenbauer (2.52), en este contexto a veces resulta más conveniente trabajar con los polinomios

$$\tilde{C}_n^{(\lambda)}(x) = \left(\frac{(n+\lambda)n!}{\lambda(2\lambda)_n} \right)^{\frac{1}{2}} C_n^{(\lambda)}(x), \quad (3.7)$$

que son ortonormales en $[-1, 1]$ con respecto a la densidad de probabilidad

$$\hat{w}_\lambda(x) = \frac{\Gamma(\lambda+1)}{\sqrt{\pi} \Gamma(\lambda+\frac{1}{2})} (1-x^2)^{\lambda-\frac{1}{2}}. \quad (3.8)$$

Las correspondientes entropías,

$$E(\widehat{C}_n^{(\lambda)}) = - \int_{-1}^1 [\widehat{C}_n^{(\lambda)}(x)]^2 \log [\widehat{C}_n^{(\lambda)}(x)]^2 \widehat{w}_\lambda(x) dx, \quad (3.9)$$

se relacionan con $E(C_n^{(\lambda)})$ mediante la fórmula

$$E(\widehat{C}_n^{(\lambda)}) = \log \left(\frac{\lambda (2\lambda)_n}{(n+\lambda) n!} \right) + \frac{\Gamma(\lambda)(n+\lambda) n!}{\sqrt{\pi} \Gamma(\lambda + \frac{1}{2})(2\lambda)_n} E(C_n^{(\lambda)}), \quad (3.10)$$

lo cual se sigue de las definiciones previas sin más que tomar en cuenta la relación de ortogonalidad (2.52).

Los casos particulares más simples dentro de los polinomios de Gegenbauer son los polinomios de Chebyshev de primera y segunda especie,

$$T_n(x) = \lim_{\lambda \rightarrow 0} \frac{n!}{(2\lambda)_n} C_n^{(\lambda)}(x), \quad U_n(x) = C_n^{(1)}(x). \quad (3.11)$$

Utilizando propiedades particulares de estos polinomios, las entropías de información de ambas familias se pueden expresar en forma cerrada [137, 138],

$$E(\widehat{T}_n) = \begin{cases} 0 & \text{si } n = 0 \\ \log 2 - 1 & \text{si } n \neq 0 \end{cases}, \quad (3.12)$$

$$E(\widehat{U}_n) = -\frac{n}{n+1}. \quad (3.13)$$

En el caso $\lambda = 2$, se demostró por primera vez en [139] que

$$E(\widehat{C}_n^{(2)}) = -\log \left(\frac{3(n+1)}{n+3} \right) - \frac{n(n^2+2n-1)}{(n+1)(n+2)(n+3)} - \frac{2}{\sqrt{(n+1)^3(n+3)^3}} \frac{T_{n+2}'''(\xi)}{T_{n+2}''(\xi)}, \quad (3.14)$$

donde

$$\xi = \frac{n+2}{\sqrt{(n+1)(n+3)}}, \quad (3.15)$$

resultado que más tarde se simplificó a [140]

$$E(\widehat{C}_n^{(2)}) = -\log \left(\frac{3(n+1)}{n+3} \right) - \frac{n^3 - 5n^2 - 29n - 27}{(n+1)(n+2)(n+3)} - \frac{1}{n+2} \left(\frac{n+3}{n+1} \right)^{n+2}. \quad (3.16)$$

En este trabajo [140], se obtuvo así mismo una generalización de (3.14) a valores enteros arbitrarios del parámetro, $\lambda = l \in \mathbb{N}$:

$$E(\widehat{C}_n^{(l)}) = -s_{nl} - r_{nl} \sum_{j=1}^{2l-2} (1 - \xi_j^2) \frac{H(\xi_j)}{P'(\xi_j)} \frac{\widehat{C}_{n-1}^{(l+1)}(\xi_j)}{\widehat{C}_n^{(l)}(\xi_j)}, \quad (3.17)$$

donde s_{nl} y r_{nl} son constantes que dependen exclusivamente de n y l , los polinomios auxiliares P y H se definen a partir de la secuencia $\{P_k\}$ ($\deg P_k = k$), que se genera a su vez mediante la relación de recurrencia

$$P_{k+1}(x) = (2l - 2k - 3)xP_k(x) - (n + k + 1)(n + 2l - k - 1)(1 - x^2)P_{k-1}(x) \quad (3.18)$$

con los valores iniciales $P_{-1}(x) = 0$, $P_0(x) = 1$, a través de las fórmulas

$$P(x) = P_{2l-2}(x), \quad H(x) = \sum_{s=0}^{2l-2} (-1)^s P_{s-1}(x) P_{2l-s-3}(x), \quad (3.19)$$

y ξ_j ($j = 1, 2, \dots, 2l - 2$) denotan los ceros de P . La expresión explícita del polinomio P se desarrolló posteriormente [141],

$$P(x) = \frac{(-1)^{l-1}(n+2l-1)!}{(n+l)n!} \sum_{\mu=0}^{l-1} \frac{(1-l)_\mu (l)_\mu (1/2)_\mu}{(1-n-l)_\mu (1+n+l)_\mu \mu!} (1-x^2)^{l-1-\mu}. \quad (3.20)$$

Desgraciadamente, (3.17) no es fácil de utilizar en la práctica. Lo que es más, no es una fórmula completamente analítica salvo para valores pequeños de l ya que, como se sigue de (3.20), los ceros ξ_j de P deben ser determinados numéricamente cuando $l \geq 6$. Algo similar sucede con los resultados de [142], que obtienen una expresión para $E(C_n^{(\lambda)})$ a partir de su potencial logarítmico. Esta estrategia, válida para calcular la entropía de cualquier familia (de hecho, es la utilizada en [138] para obtener (3.12)-(3.13)), consiste en observar que si $\{x_j^{(n)}\}$ son los ceros de $p_n(x)$, que satisface la relación de ortogonalidad (2.53), entonces

$$p_n(x) = \gamma_n \prod_{j=1}^n (x - x_j^{(n)}), \quad (3.21)$$

y así reescribir $E(p_n)$, dada por (2.56), como

$$E(p_n) = -2 \log \gamma_n + 2 \sum_{j=1}^n V(x_j^{(n)}, p_n), \quad (3.22)$$

donde

$$V(x, p_n) = - \int_{\Delta} \log |x - t| \omega(t) dt \quad (3.23)$$

es el potencial logarítmico asociado a los polinomios $\{p_n\}$. Así, en [142] se obtiene una fórmula para $V(x, C_n^{(\lambda)})$, lo que redundaría en una expresión para $E(C_n^{(\lambda)})$. Sin embargo, ésta no es analítica salvo para valores pequeños de n , puesto que se expresa en términos de los ceros de $C_n^{(\lambda)}(x)$ ¹.

Estos resultados comprenden todos los avances en entropías de polinomios en lo que a fórmulas exactas se refiere. Además, es interesante señalar la observación

¹Merece la pena señalarse que también se conocen expresiones análogas para la entropía de los polinomios de Hermite [143] y Laguerre [144] a partir de su potencial logarítmico.

de Sánchez-Lara y Sánchez-Ruiz [145] de que, aparte de las estrategias llevadas a cabo en [137] y [138], $E(T_n)$ y $E(U_n)$ son fácilmente calculables a partir de las bien conocidas representaciones trigonométricas de estos polinomios,

$$T_n(\cos \theta) = \cos n\theta, \quad U_n(\cos \theta) = \frac{\sin(n+1)\theta}{\sin \theta}, \quad (3.24)$$

con $x = \cos \theta$, ya que de esta forma las integrales entrópicas correspondientes se reducen a integrales bien conocidas de funciones trigonométricas. Motivados por esta observación, nuestro objetivo en el presente capítulo es obtener expresiones para $E(C_n^{(\lambda)})$ a partir de representaciones trigonométricas de los polinomios de Gegenbauer en general. Por simplicidad, comenzamos ateniéndonos a esta familia, aunque, como ya mencionamos, también consideraremos esta idea en el marco más general de los polinomios de Jacobi.

3.1. Entropía de polinomios de Gegenbauer de parámetro entero

3.1.1. Representaciones trigonométricas para los polinomios de Gegenbauer

Aunque quizá no tan conocidas como en el caso de los polinomios de Chebyshev, en la literatura especializada se encuentran diversas representaciones trigonométricas para los polinomios de Gegenbauer. La más extendida, y que, por tanto, denominaremos representación estándar, es (ver por ejemplo [146, p. 302])

$$C_n^{(\lambda)}(\cos \theta) = \sum_{m=0}^n d_{m,n}^{(\lambda)} e^{i(n-2m)\theta} = \sum_{m=0}^n d_{m,n}^{(\lambda)} \cos(n-2m)\theta, \quad (3.25)$$

donde

$$d_{m,n}^{(\lambda)} = \frac{(\lambda)_m (\lambda)_{n-m}}{m!(n-m)!}. \quad (3.26)$$

Así mismo, se conoce otra representación debida a Szegő [147, 148],

$$C_n^{(\lambda)}(\cos \theta) = \frac{c_n^{(\lambda)}}{(\sin \theta)^{2\lambda-1}} \sum_{\nu=0}^{\infty} \alpha_{\nu,n}^{(\lambda)} \sin(n+2\nu+1)\theta, \quad \lambda > 0, \lambda \notin \mathbb{N}, \quad (3.27)$$

donde

$$c_n^{(\lambda)} = \frac{2^{2-2\lambda} \Gamma(n+2\lambda)}{\Gamma(\lambda) \Gamma(n+\lambda+1)}, \quad \alpha_{\nu,n}^{(\lambda)} = \frac{(1-\lambda)_\nu (n+1)_\nu}{\nu! (n+\lambda+1)_\nu}. \quad (3.28)$$

A primera vista, esta representación puede parecer menos útil que la primera al contener infinitos términos. Además, en la literatura previa se da por cierta con la restricción $\lambda \notin \mathbb{N}$. Sin embargo, no es difícil probar que la validez de (3.27) se extiende al caso en el que λ es un entero positivo.

Proposición 3.1. *La representación de Szegö (3.27) es válida cuando $\lambda \in \mathbb{N}$. En este caso se tiene que*

$$C_n^{(\lambda)}(\cos \theta) = \frac{c_n^{(\lambda)}}{(\sin \theta)^{2\lambda-1}} \sum_{\nu=0}^{\lambda-1} \alpha_{\nu,n}^{(\lambda)} \sin(n+2\nu+1)\theta. \quad (3.29)$$

Demostración. Nótese que si $\lambda \in \mathbb{N}$ entonces $\alpha_{\nu,n}^{(\lambda)} = 0$ cuando $\nu \geq \lambda$, de forma que (3.27) se reduce a (3.29). Probaremos su validez por inducción en λ . Cuando $\lambda = 1$, (3.29) es obviamente cierta ya que se reduce a la bien conocida representación trigonométrica de los polinomios de Chebyshev de segunda especie (segunda ecuación en (3.24)). Ahora, asumamos que el resultado es cierto para $\lambda = m - 1$ ($m \in \mathbb{N}$). Consideremos la siguiente relación de recurrencia para los polinomios de Gegenbauer [148, Eq. (4.7.27)],

$$nC_n^{(\lambda)}(x) = (2\lambda + n - 1)x C_{n-1}^{(\lambda)}(x) - 2\lambda(1 - x^2)C_{n-2}^{(\lambda+1)}(x), \quad (3.30)$$

que en forma trigonométrica ($x = \cos \theta$) se reescribe como

$$C_n^{(\lambda)}(\cos \theta) = \frac{1}{2(\lambda - 1)\sin^2 \theta} \times \left[(2\lambda + n - 1) \cos \theta C_{n+1}^{(\lambda-1)}(\cos \theta) - (n + 2) C_{n+2}^{(\lambda-1)}(\cos \theta) \right]. \quad (3.31)$$

Utilizando esta fórmula para $\lambda = m$ y sustituyendo (3.29) en el miembro derecho de ésta se llega a

$$\begin{aligned} C_n^{(m)}(\cos \theta) &= \frac{2c_n^{(m)}}{(\sin \theta)^{2m-1}} \left[\cos \theta \sum_{\nu=0}^{m-2} \alpha_{\nu,n+1}^{(m-1)} \sin(n+2\nu+2)\theta \right. \\ &\quad \left. - \frac{n+2}{n+m+1} \sum_{\nu=0}^{m-2} \alpha_{\nu,n+2}^{(m-1)} \sin(n+2\nu+3)\theta \right] \\ &= \frac{c_n^{(m)}}{(\sin \theta)^{2m-1}} \left[\sum_{\nu=0}^{m-2} \alpha_{\nu,n+1}^{(m-1)} \sin(n+2\nu+3)\theta + \sum_{\nu=0}^{m-2} \alpha_{\nu,n+1}^{(m-1)} \sin(n+2\nu+1)\theta \right. \\ &\quad \left. - \frac{2(n+2)}{n+m+1} \sum_{\nu=0}^{m-2} \alpha_{\nu,n+2}^{(m-1)} \sin(n+2\nu+3)\theta \right] \\ &= \frac{c_n^{(m)}}{(\sin \theta)^{2m-1}} \sum_{\nu=0}^{m-1} \left(\alpha_{\nu-1,n+1}^{(m-1)} + \alpha_{\nu,n+1}^{(m-1)} - \frac{2(n+2)}{n+m+1} \alpha_{\nu-1,n+2}^{(m-1)} \right) \sin(n+2\nu+1)\theta, \end{aligned} \quad (3.32)$$

donde en el último paso hemos usado que $\alpha_{m-1,n+1}^{(m-1)} = 0$ y $\alpha_{\nu,n+1}^{(m-1)} = 0$ cuando $\nu < 0$. Una cuenta sin complicaciones muestra que

$$\alpha_{\nu-1,n+1}^{(m-1)} + \alpha_{\nu,n+1}^{(m-1)} - \frac{2(n+2)}{n+m+1} \alpha_{\nu-1,n+2}^{(m-1)} = \alpha_{\nu,n}^{(m)}, \quad (3.33)$$

probándose el resultado anunciado. \square

De esta forma la representación de Szegö no sólo es válida para $\lambda \in \mathbb{N}$ sino que además la suma se trunca conteniendo solamente λ sumandos. Así, llegamos a la conclusión de que dentro del problema de la evaluación de $E(C_n^{(\lambda)})$, si nos atenemos al caso $\lambda \in \mathbb{N}$ (con λ no demasiado grande) esta representación contribuirá a simplificar considerablemente nuestro problema independientemente del grado del polinomio (contrariamente a la representación estándar cuyo número de términos depende directamente del grado). Nótese además que la representación de Szegö ya incluye la función peso (tal y como se escribe en (3.29) está en el denominador del miembro derecho), lo que redundará en una mayor simplificación de la integral entrópica. Consecuentemente, en lo que sigue asumimos que $\lambda \in \mathbb{N}$ a menos que se especifique lo contrario.

3.1.2. Evaluación de la integral entrópica

Con el cambio de variable $x = \cos \theta$, la integral (3.6) se convierte en

$$E(C_n^{(\lambda)}) = - \int_0^\pi (C_n^{(\lambda)}(\cos \theta))^2 \log (C_n^{(\lambda)}(\cos \theta))^2 \sin^{2\lambda} \theta d\theta . \quad (3.34)$$

Utilizando la representación de Szegö (3.29) para uno de los dos polinomios de Gegenbauer en $(C_n^{(\lambda)}(\cos \theta))^2$, (3.34) se reescribe como

$$E(C_n^{(\lambda)}) = -\frac{1}{2} c_n^{(\lambda)} \sum_{\nu=0}^{\lambda-1} \alpha_{\nu,n}^{(\lambda)} (J_{\nu,n}^{(\lambda)} - J_{\nu+1,n}^{(\lambda)}), \quad (3.35)$$

donde

$$J_{\nu,n}^{(\lambda)} := \int_0^\pi C_n^{(\lambda)}(\cos \theta) \cos(n+2\nu)\theta \log (C_n^{(\lambda)}(\cos \theta))^2 d\theta . \quad (3.36)$$

Ahora, utilizando la representación estándar (3.25) se tiene que

$$C_n^{(\lambda)}(\cos \theta) \cos(n+2\nu)\theta = \frac{1}{2} \sum_{m=0}^n d_{m,n}^{(\lambda)} \cos 2(m+\nu)\theta + \frac{1}{2} \sum_{m=0}^n d_{m,n}^{(\lambda)} \cos 2(n-m+\nu)\theta . \quad (3.37)$$

Teniendo en cuenta la propiedad de simetría $d_{m,n}^{(\lambda)} = d_{n-m,n}^{(\lambda)}$, que se deduce a partir de la expresión explícita de los coeficientes $d_{m,n}^{(\lambda)}$ (3.26), la ecuación anterior se simplifica a

$$C_n^{(\lambda)}(\cos \theta) \cos(n+2\nu)\theta = \sum_{m=0}^n d_{m,n}^{(\lambda)} \cos 2(m+\nu)\theta , \quad (3.38)$$

de manera que

$$J_{\nu,n}^{(\lambda)} = \sum_{m=0}^n d_{m,n}^{(\lambda)} \int_0^\pi \cos 2(m+\nu)\theta \log (C_n^{(\lambda)}(\cos \theta))^2 d\theta . \quad (3.39)$$

Definiendo las integrales

$$I_{m,n}^{(\lambda)} := \int_0^\pi \cos(2m\theta) \log(C_n^{(\lambda)}(\cos\theta))^2 d\theta \quad (3.40)$$

y usando (3.35) y (3.39), encontramos que $E(C_n^{(\lambda)})$ viene dado por

$$E(C_n^{(\lambda)}) = -\frac{1}{2} c_n^{(\lambda)} \sum_{\nu=0}^{\lambda-1} \alpha_{\nu,n}^{(\lambda)} \sum_{m=0}^n d_{m,n}^{(\lambda)} \left(I_{\nu+m,n}^{(\lambda)} - I_{\nu+1+m,n}^{(\lambda)} \right). \quad (3.41)$$

Se puede obtener una expresión alternativa para $E(C_n^{(\lambda)})$ que resultará más conveniente más adelante considerando que

$$\begin{aligned} E(C_n^{(\lambda)}) &= -\frac{1}{2} c_n^{(\lambda)} \sum_{\nu=0}^{\lambda-1} \alpha_{\nu,n}^{(\lambda)} \left\{ \sum_{m=0}^n d_{m,n}^{(\lambda)} I_{\nu+m,n}^{(\lambda)} - \sum_{m=1}^{n+1} d_{m-1,n}^{(\lambda)} I_{\nu+m,n}^{(\lambda)} \right\} \\ &= -\frac{1}{2} c_n^{(\lambda)} \sum_{\nu=0}^{\lambda-1} \alpha_{\nu,n}^{(\lambda)} \left\{ \sum_{m=1}^n \left(d_{m,n}^{(\lambda)} - d_{m-1,n}^{(\lambda)} \right) I_{\nu+m,n}^{(\lambda)} + d_{0,n}^{(\lambda)} I_{\nu,n}^{(\lambda)} - d_{n,n}^{(\lambda)} I_{\nu+n+1,n}^{(\lambda)} \right\}. \end{aligned} \quad (3.42)$$

De acuerdo con (3.26), $d_{m,n}^{(\lambda)} = 0$ cuando $-\lambda < m < 0$ o $n < m < n + \lambda$. Restrinjámonos inicialmente al supuesto de que $\lambda \neq 1$, de forma que $d_{-1,n}^{(\lambda)} = 0$ y $d_{n+1,n}^{(\lambda)} = 0$. Esto permite escribir la fórmula anterior en la forma más compacta

$$\begin{aligned} E(C_n^{(\lambda)}) &= -\frac{1}{2} c_n^{(\lambda)} \sum_{\nu=0}^{\lambda-1} \sum_{m=0}^{n+1} \alpha_{\nu,n}^{(\lambda)} \left(d_{m,n}^{(\lambda)} - d_{m-1,n}^{(\lambda)} \right) I_{\nu+m,n}^{(\lambda)} \\ &= -\frac{1}{2} c_n^{(\lambda)} \sum_{\nu=0}^{\lambda-1} \sum_{m=\nu}^{n+1+\nu} \alpha_{\nu,n}^{(\lambda)} \left(d_{m-\nu,n}^{(\lambda)} - d_{m-\nu-1,n}^{(\lambda)} \right) I_{m,n}^{(\lambda)}. \end{aligned} \quad (3.43)$$

Beneficiándonos de nuevo del hecho de que $d_{m,n}^{(\lambda)} = 0$ cuando $-\lambda < m < 0$ así como cuando $n < m < n + \lambda$, es posible extender los límites inferior y superior de sumación a 1 y $n + \lambda - 1$, respectivamente, siempre y cuando tratemos separadamente los términos correspondientes a $m = 0$ y $m = n + \lambda$. Así, se encuentra que

$$E(C_n^{(\lambda)}) = -\frac{1}{2} c_n^{(\lambda)} \left(\alpha_{0,n}^{(\lambda)} d_{0,n}^{(\lambda)} I_{0,n}^{(\lambda)} - \alpha_{\lambda-1,n}^{(\lambda)} d_{n,n}^{(\lambda)} I_{n+\lambda,n}^{(\lambda)} + \sum_{m=1}^{n+\lambda-1} \beta_{m,n}^{(\lambda)} I_{m,n}^{(\lambda)} \right), \quad (3.44)$$

donde

$$\beta_{m,n}^{(\lambda)} = \sum_{\nu=0}^{\lambda-1} \alpha_{\nu,n}^{(\lambda)} \left(d_{m-\nu,n}^{(\lambda)} - d_{m-\nu-1,n}^{(\lambda)} \right). \quad (3.45)$$

Es directo comprobar que la validez de (3.44) se extiende fuera del supuesto original de $\lambda \neq 1$, puesto que cuando $\lambda = 1$ el miembro derecho de esta ecuación coincide con el de (3.42).

En definitiva, hemos logrado expresar $E(C_n^{(\lambda)})$ como una suma de integrales trigonométricas $I_{m,n}^{(\lambda)}$. Si queremos beneficiarnos de esta fórmula para calcular las entropías de los polinomios de Gegenbauer el siguiente paso es, lógicamente, evaluar estas integrales, al menos cuando $0 \leq m \leq n + \lambda$. Como veremos a continuación, esto es posible utilizando técnicas de integración en el plano complejo.

Teorema 3.1. Para $\lambda \in \mathbb{N}$,

$$I_{0,n}^{(\lambda)} = 2\pi \log \left(\frac{(\lambda)_n}{n!} \right) \quad (3.46)$$

y, cuando $m \geq 1$,

$$I_{m,n}^{(\lambda)} = \frac{(2\lambda - 1)\pi}{m} + \frac{\pi}{(2m)!} \frac{d^{2m}}{dz^{2m}} \left(\log \sum_{\nu=0}^{\lambda-1} \alpha_{\nu,n}^{(\lambda)} (z^{2n+2\lambda+2\nu} - z^{2\lambda-2\nu-2}) \right) \Big|_{z=0}. \quad (3.47)$$

Demostración. Teniendo en cuenta que $C_n^{(\lambda)}(-x) = (-1)^n C_n^{(\lambda)}(x)$, (3.40) se puede escribir como

$$\begin{aligned} I_{m,n}^{(\lambda)} &= \frac{1}{2} \int_0^{2\pi} \cos(2m\theta) \log |C_n^{(\lambda)}(\cos \theta)|^2 d\theta \\ &= \int_0^{2\pi} \cos(2m\theta) \log |C_n^{(\lambda)}(\cos \theta)| d\theta \\ &= \int_0^{2\pi} \cos(2m\theta) \log |e^{in\theta} C_n^{(\lambda)}(\cos \theta)| d\theta, \end{aligned} \quad (3.48)$$

donde en el último paso hemos introducido el factor $e^{in\theta}$ para simplificar desarrollos posteriores. Utilizando la representación de Szegő (3.29) para el polinomio de Gegenbauer $C_n^{(\lambda)}(\cos \theta)$, la anterior ecuación se convierte en

$$I_{m,n}^{(\lambda)} = \int_0^{2\pi} \cos(2m\theta) \log \left| \frac{c_n^{(\lambda)} e^{in\theta}}{\sin^{2\lambda-1} \theta} \sum_{\nu=0}^{\lambda-1} \alpha_{\nu,n}^{(\lambda)} \sin(n+2\nu+1)\theta \right| d\theta. \quad (3.49)$$

En lo que sigue calcularemos la integral

$$\mathcal{I}_{m,n}^{(\lambda)} = \int_0^{2\pi} \cos(2m\theta) \log \left(\frac{c_n^{(\lambda)} e^{in\theta}}{\sin^{2\lambda-1} \theta} \sum_{\nu=0}^{\lambda-1} \alpha_{\nu,n}^{(\lambda)} \sin(n+2\nu+1)\theta \right) d\theta, \quad (3.50)$$

cuya parte real es $I_{m,n}^{(\lambda)}$. Introduciendo el cambio de variable $z = \exp(i\theta)$ se llega a

$$\mathcal{I}_{m,n}^{(\lambda)} = \frac{1}{2i} \oint_{|z|=1} \frac{z^{4m} + 1}{z^{2m+1}} \log q(z) dz, \quad (3.51)$$

donde

$$q(z) = z^n C_n^{(\lambda)} \left(\frac{z + z^{-1}}{2} \right) = c_n^{(\lambda)} 2^{2\lambda-2} (-1)^\lambda \frac{\sum_{\nu=0}^{\lambda-1} \alpha_{\nu,n}^{(\lambda)} (z^{2n+2\lambda+2\nu} - z^{2\lambda-2\nu-2})}{(1-z^2)^{2\lambda-1}}. \quad (3.52)$$

Las singularidades del integrando son $z = 0$, que es un polo de orden $2m + 1$, y todos los ceros de $q(z)$, que constituyen puntos de ramificación. Si $\{x_{n,j}\}_{j=1}^n$ denotan los ceros de $C_n^{(\lambda)}(x)$, que, como es bien sabido (ver por ejemplo [49]), son reales y simples y se encuentran en $(-1, 1)$, entonces los ceros $\{z_{n,j}\}_{j=1}^{2n}$ de la función $q(z)$ son

$$z_{n,j+\frac{n}{2} \mp \frac{n}{2}} = \exp(i \arccos x_{n,j}) = x_{n,j} \pm i \sqrt{1 - x_{n,j}^2}, \quad j = 1, 2, \dots, n. \quad (3.53)$$

Esto significa que todos los $\{z_{n,j}\}_{j=1}^{2n}$ se encuentran en la circunferencia unidad, lo que también puede verse a partir del hecho de que el cambio de variable llevado a cabo $z = \exp(i \arccos x)$ lleva $(-1, 1)$ a la circunferencia unidad. En conclusión, el integrando de (3.51) tiene $2n$ puntos de ramificación situados en el contorno de integración. Para evitar esta dificultad consideramos la misma integral a lo largo de un contorno levemente modificado que denotamos Γ (véase la figura 3.1), y que también es cerrado. Nótese que las ramas de los logaritmos se pueden escoger desde los puntos rama al exterior del disco unidad, de forma que Γ no corta a ninguna de ellas. Dado que la única singularidad en el interior de Γ es $z = 0$ se tiene ahora que

$$\oint_{\Gamma} \frac{z^{4m} + 1}{z^{2m+1}} \log q(z) dz = 2\pi i \operatorname{Res} \left(\frac{z^{4m} + 1}{z^{2m+1}} \log q(z), z = 0 \right). \quad (3.54)$$

La integral a lo largo de Γ se puede descomponer como

$$\oint_{\Gamma} \frac{z^{4m} + 1}{z^{2m+1}} \log q(z) dz = \sum_{j=1}^{2n} \left(\int_{\epsilon_j} \frac{z^{4m} + 1}{z^{2m+1}} \log q(z) dz + \int_{\gamma_j} \frac{z^{4m} + 1}{z^{2m+1}} \log q(z) dz \right), \quad (3.55)$$

donde ϵ_j denota el arco de circunferencia de radio ε con centro en $z_{n,j}$ utilizado para evitar este punto rama y γ_j denota el arco de la circunferencia unidad que conecta ϵ_j y ϵ_{j+1} (γ_{2n} conecta ϵ_{2n} y ϵ_1). Mediante la parametrización $z = z_{n,j} + \varepsilon e^{i\theta}$ se encuentra que

$$\int_{\epsilon_j} \frac{z^{4m} + 1}{z^{2m+1}} \log q(z) dz = \int_{\epsilon_j} \frac{(z_{n,j} + \varepsilon e^{i\theta})^{4m} + 1}{(z_{n,j} + \varepsilon e^{i\theta})^{2m+1}} \log q(z_{n,j} + \varepsilon e^{i\theta}) i \varepsilon e^{i\theta} d\theta \xrightarrow{\varepsilon \rightarrow 0} 0, \quad (3.56)$$

donde hemos utilizado que $x \log x \rightarrow 0$ cuando $x \rightarrow 0$. Así, tomando el límite $\varepsilon \rightarrow 0$ en (3.55) se concluye que

$$\oint_{|z|=1} \frac{z^{4m} + 1}{z^{2m+1}} \log q(z) dz = \oint_{\Gamma} \frac{z^{4m} + 1}{z^{2m+1}} \log q(z) dz. \quad (3.57)$$

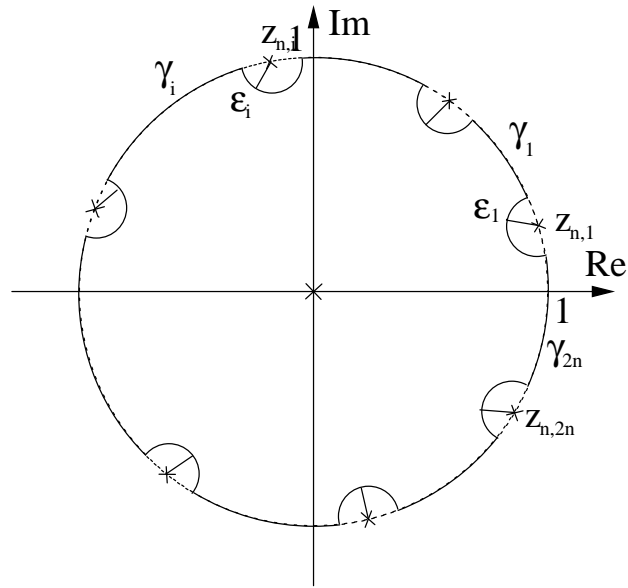


Figura 3.1: Contorno de integración Γ utilizado para evitar los puntos de ramificación situados en la circunferencia unidad. Nótese que las correspondientes ramas se escogerían hacia el exterior del círculo unidad.

Recordando que el residuo de la función meromorfa $h(z)$ en un polo z_0 de orden $2m + 1$ viene dado por

$$\text{Res}(h(z), z = z_0) = \frac{1}{(2m)!} \frac{d^{2m}}{dz^{2m}} \left((z - z_0)^{2m+1} h(z) \right) \Big|_{z=z_0}, \quad (3.58)$$

el uso de (3.54) y (3.57) en (3.51) lleva a

$$\mathcal{I}_{m,n}^{(\lambda)} = \frac{\pi}{(2m)!} \frac{d^{2m}}{dz^{2m}} \left[(z^{4m} + 1) \log q(z) \right] \Big|_{z=0}. \quad (3.59)$$

En el caso $m = 0$, la ecuación anterior se reduce a

$$\mathcal{I}_{0,n}^{(\lambda)} = 2\pi \log q(0) = 2\pi \log \left(c_n^{(\lambda)} 2^{2\lambda-2} (-1)^{\lambda+1} \alpha_{\lambda-1,n}^{(\lambda)} \right), \quad (3.60)$$

por lo que

$$\mathcal{I}_{0,n}^{(\lambda)} = 2\pi \log \left| c_n^{(\lambda)} 2^{2\lambda-2} \alpha_{\lambda-1,n}^{(\lambda)} \right| = 2\pi \log \left(\frac{(\lambda)_n}{n!} \right), \quad (3.61)$$

lo que prueba la primera parte del teorema². Por otro lado, si $m \geq 1$ se sigue de (3.59) que $\mathcal{I}_{m,n}^{(\lambda)} \in \mathbb{R}$, de forma que $I_{m,n}^{(\lambda)} = \mathcal{I}_{m,n}^{(\lambda)}$. Más aún, en este caso el

²Esta parte también se puede probar utilizando el teorema del valor medio para funciones armónicas (ver [137, Sec. VI]).

factor $(z^{4m} + 1)$ del miembro derecho de (3.59) se puede omitir, ya que su valor en $z = 0$ es la unidad mientras que todas sus derivadas en este punto se anulan. Se encuentra así que

$$I_{m,n}^{(\lambda)} = \frac{\pi}{(2m)!} \frac{d^{2m}}{dz^{2m}} \left[\log \left(\frac{\sum_{\nu=0}^{\lambda-1} \alpha_{\nu,n}^{(\lambda)} (z^{2n+2\lambda+2\nu} - z^{2\lambda-2\nu-2})}{(1-z^2)^{2\lambda-1}} \right) \right] \Big|_{z=0}, \quad (3.62)$$

y (3.47) se obtiene finalmente si observamos que

$$\frac{d^{2m}}{dz^{2m}} (\log(1-z^2)) \Big|_{z=0} = \frac{d^{2m}}{dz^{2m}} \left(- \sum_{k=1}^{\infty} \frac{z^{2k}}{k} \right) \Big|_{z=0} = - \frac{(2m)!}{m} \quad (3.63)$$

cuando $m \neq 0$. □

El hecho de que la representación de Szegö (3.29) contiene un número finito de sumandos juega un papel esencial en la demostración del Teorema 3.1. Aunque estamos principalmente interesados en evaluar las integrales $I_{m,n}^{(\lambda)}$ cuando $\lambda \in \mathbb{N}$, debe apuntarse que estas integrales se pueden calcular de forma similar para todos los posibles valores de λ utilizando la representación trigonométrica estándar (3.25) en lugar de la de Szegö.

Teorema 3.2. Para $\lambda \in \mathbb{R}$, $\lambda > -\frac{1}{2}$,

$$I_{0,n}^{(\lambda)} = 2\pi \log \left(\frac{(\lambda)_n}{n!} \right) \quad (3.64)$$

y, cuando $m \geq 1$,

$$I_{m,n}^{(\lambda)} = \frac{\pi}{(2m)!} \frac{d^{2m}}{dz^{2m}} \left(\log \sum_{j=0}^n d_{j,n}^{(\lambda)} z^{2n-2j} \right) \Big|_{z=0}. \quad (3.65)$$

Demostración. Procedemos como en la demostración del Teorema 3.1, pero utilizando la forma compleja de la representación estándar (3.25) para el polinomio de Gegenbauer $C_n^{(\lambda)}(\cos \theta)$ en (3.48). De esta forma se llega a

$$I_{m,n}^{(\lambda)} = \frac{\pi}{(2m)!} \frac{d^{2m}}{dz^{2m}} \left[(z^{4m} + 1) \log \left(\sum_{j=0}^n d_{j,n}^{(\lambda)} z^{2n-2j} \right) \right] \Big|_{z=0}, \quad (3.66)$$

de lo que, análogamente, se siguen (3.64) y (3.65). □

Una vez evaluadas las integrales, en nuestro objetivo de llevar a cabo la suma en (3.44) el siguiente paso es obtener fórmulas cerradas para las derivadas que expresan $I_{m,n}^{(\lambda)}$ cuando $1 \leq m \leq n + \lambda$. A pesar de su mayor generalidad, el Teorema 3.2 resulta menos útil que el Teorema 3.1 porque (3.47) expresa las integrales en términos de logaritmos de un polinomio con 2λ términos,

mientras que en (3.65) tenemos el logaritmo de un polinomio de $n + 1$ términos. Como veremos, la dificultad para obtener fórmulas cerradas para tales funciones aumenta con el número de sumandos del polinomio. Por lo tanto, si buscamos una expresión para $I_{m,n}^{(\lambda)}$ para un valor fijo de λ y cualquier $n \in \mathbb{N}$, el Teorema 3.1 será de más ayuda, particularmente para valores pequeños de λ .

En el caso $\lambda = 1$, se sigue claramente de (3.47) que, si $1 \leq m \leq n + 1$, entonces

$$\begin{aligned} I_{m,n}^{(1)} &= \frac{\pi}{m} + \frac{\pi}{(2m)!} \frac{d^{2m}}{dz^{2m}} (\log(1 - z^{2n+2})) \Big|_{z=0} \\ &= \frac{\pi}{m} + \frac{\pi}{(2m)!} \frac{d^{2m}}{dz^{2m}} \left(- \sum_{k=1}^{\infty} \frac{z^{(2n+2)k}}{k} \right) \Big|_{z=0} \\ &= \pi \left(\frac{1}{m} - \delta_{m,n+1} \right). \end{aligned} \quad (3.67)$$

Cuando $\lambda \geq 2$, el polinomio contenido en el logaritmo tiene más términos por lo que el truco antes utilizado no funciona en este caso. Sin embargo, es posible obtener fórmulas cerradas para las derivadas de (3.47) por medio de la fórmula de Faà di Bruno para las derivadas de una composición de funciones, que establece que (véase por ejemplo [149])

$$\frac{d^m f(g(z))}{dz^m} = m! \sum_{k=0}^m f^{(k)}(g(z)) \sum_{k_1, k_2, \dots, k_m} \prod_{j=1}^m \frac{[g^{(j)}(z)]^{k_j}}{(j!)^{k_j} k_j!}, \quad (3.68)$$

donde la suma interna se extiende sobre todas las particiones que satisfagan

$$k_1 + k_2 + \dots + k_m = k, \quad k_1 + 2k_2 + \dots + mk_m = m. \quad (3.69)$$

Esta fórmula nos permite derivar expresiones explícitas para $I_{m,n}^{(\lambda)}$ con $\lambda \geq 2$, como mostramos en las siguientes dos proposiciones.

Proposición 3.2. *En el caso $\lambda = 2$, cuando $1 \leq m \leq n + 2$*

$$I_{m,n}^{(2)} = \frac{\pi}{m} \left[3 - \left(\frac{n+3}{n+1} \right)^m \right] + \pi \frac{n+3}{n+1} \delta_{m,n+2}. \quad (3.70)$$

Demostración. En este caso, la aplicación de la fórmula Faà di Bruno (3.68) a las derivadas de (3.47) da³

$$\begin{aligned} & \frac{d^{2m}}{dz^{2m}} \left(\log \sum_{\nu=0}^1 \alpha_{\nu,n}^{(2)} (z^{2n+4+2\nu} - z^{2-2\nu}) \right) \Big|_{z=0} = (2m)! \sum_{k=1}^{2m} \frac{d^k}{dz^k} (\log z) \Big|_{z=-\alpha_{1,n}^{(2)}} \\ & \times \sum_{k_1, k_2, \dots, k_{2m}} \prod_{j=1}^{2m} \frac{\left[\frac{d^j}{dz^j} \left(\sum_{\nu=0}^1 \alpha_{\nu,n}^{(2)} (z^{2n+4+2\nu} - z^{2-2\nu}) \right) \Big|_{z=0} \right]^{k_j}}{(j!)^{k_j} k_j!}. \end{aligned} \quad (3.71)$$

³Nótese que, de acuerdo a (3.69), $k = 0$ se corresponde con $m = 0$, por lo que podemos empezar la suma en $k = 1$.

Por otro lado, para $k \geq 1$,

$$\frac{d^k}{dz^k}(\log z) = \frac{(-1)^{k+1}(k-1)!}{z^k}, \quad (3.72)$$

de forma que

$$\frac{d^k}{dz^k}(\log z) \Big|_{z=-\alpha_{1,n}^{(2)}} = -\frac{(k-1)!}{(\alpha_{1,n}^{(2)})^k}. \quad (3.73)$$

Por otro lado, todas las derivadas del polinomio en (3.71) se anulan en $z = 0$ excepto cuando $j = 2$ y $j = 2n + 4$, por lo que debemos fijar $k_j = 0$ si $j \neq 2$ y $j \neq 2n + 4$. Las condiciones (3.69) se reducen entonces a

$$k_2 + k_{2n+4} = k, \quad 2k_2 + (2n+4)k_{2n+4} = 2m. \quad (3.74)$$

Como k_2 y k_{2n+4} son enteros no negativos, estas ecuaciones sólo admiten la solución $k_{2n+4} = 0, k_2 = k = m$ cuando $m \leq n + 1$, mientras que en el caso $m = n + 2$ se debe añadir a la anterior la solución $k_{2n+4} = k = 1, k_2 = 0$. Consecuentemente, (3.71) se simplifica a

$$\begin{aligned} & \frac{d^{2m}}{dz^{2m}} \left(\log \sum_{\nu=0}^1 \alpha_{\nu,n}^{(2)} (z^{2n+4+2\nu} - z^{2-2\nu}) \right) \Big|_{z=0} \\ &= -(2m)! \left[\frac{1}{m} \left(-\frac{\alpha_{0,n}^{(2)}}{\alpha_{1,n}^{(2)}} \right)^m + \frac{\alpha_{0,n}^{(2)}}{\alpha_{1,n}^{(2)}} \delta_{m,n+2} \right], \end{aligned} \quad (3.75)$$

y se obtiene el resultado final usando la segunda ecuación en (3.28). \square

Proposición 3.3. *Para cualquier $\lambda \in \mathbb{N}$, $\lambda \geq 3$, cuando $1 \leq m \leq n + \lambda$*

$$\begin{aligned} I_{m,n}^{(\lambda)} &= \frac{(2\lambda-1)\pi}{m} - \pi \frac{\alpha_{0,n}^{(\lambda)}}{\alpha_{\lambda-1,n}^{(\lambda)}} \delta_{m,n+\lambda} \\ &\quad - \pi \left(\frac{\alpha_{\lambda-3,n}^{(\lambda)}}{\alpha_{\lambda-2,n}^{(\lambda)}} \right)^m \sum_{k=1}^m \sum_{k_6=0}^{[m/3]} \sum_{k_8=0}^{[m/4]} \cdots \sum_{k_{2(\lambda-1)}=0}^{[m/(\lambda-1)]} \left(-\frac{(\alpha_{\lambda-2,n}^{(\lambda)})^2}{\alpha_{\lambda-1,n}^{(\lambda)} \alpha_{\lambda-3,n}^{(\lambda)}} \right)^k \\ &\quad \times \frac{(k-1)!}{(2k-m + \sum_{r=1}^{\lambda-3} r k_{2r+4})! (m-k - \sum_{s=1}^{\lambda-3} (s+1) k_{2s+4})!} \\ &\quad \times \prod_{j=3}^{\lambda-1} \frac{1}{(k_{2j})!} \left(-\frac{\alpha_{\lambda-1-j,n}^{(\lambda)} (\alpha_{\lambda-2,n}^{(\lambda)})^{j-2}}{(\alpha_{\lambda-3,n}^{(\lambda)})^{j-1}} \right)^{k_{2j}}, \end{aligned} \quad (3.76)$$

donde en los límites superiores de las sumas sobre $k_6, k_8, \dots, k_{2(\lambda-1)}$, $[\cdot]$ denota la parte entera.

En particular, en el caso $\lambda = 3$, se tiene que cuando $1 \leq m \leq n + 3$

$$I_{m,n}^{(3)} = \frac{\pi}{m} \left[5 - 2\operatorname{Re}(f(n)^m) \right] - \pi \frac{(n+4)(n+5)}{(n+1)(n+2)} \delta_{m,n+3}, \quad (3.77)$$

donde

$$f(n) = \frac{(n+1)(n+5) + i\sqrt{3(n+1)(n+5)}}{(n+1)(n+2)}. \quad (3.78)$$

Demostración. En el caso general ($\lambda \in \mathbb{N}$, $\lambda \geq 3$), la aplicación de la fórmula de Faà di Bruno (3.68) a las derivadas de (3.47) y el uso de (3.72) llevan a

$$\begin{aligned} & \frac{d^{2m}}{dz^{2m}} \left(\log \sum_{\nu=0}^{\lambda-1} \alpha_{\nu,n}^{(\lambda)} (z^{2n+2\lambda+2\nu} - z^{2\lambda-2\nu-2}) \right) \Big|_{z=0} \\ &= -(2m)! \sum_{k=1}^{2m} \frac{(k-1)!}{(\alpha_{\lambda-1,n}^{(\lambda)})^k} \sum_{k_1, k_2, \dots, k_{2m}} \prod_{j=1}^{2m} \frac{\left[\frac{d^j}{dz^j} \left(\sum_{\nu=0}^{\lambda-1} \alpha_{\nu,n}^{(\lambda)} z^{2n+2\lambda+2\nu} \right) \Big|_{z=0} \right]^{k_j}}{(j!)^{k_j} k_j!} \\ & - (2m)! \sum_{k=1}^{2m} \frac{(k-1)!}{(\alpha_{\lambda-1,n}^{(\lambda)})^k} \sum_{k_1, k_2, \dots, k_{2m}} \prod_{j=1}^{2m} \frac{\left[\frac{d^j}{dz^j} \left(- \sum_{\nu=0}^{\lambda-1} \alpha_{\nu,n}^{(\lambda)} z^{2\lambda-2\nu-2} \right) \Big|_{z=0} \right]^{k_j}}{(j!)^{k_j} k_j!}. \end{aligned} \quad (3.79)$$

En el primer término del miembro derecho todas las derivadas se anulan en $z = 0$ excepto cuando $j = 2n + 2\lambda$, por lo que $k_j = 0$ cuando $j \neq 2n + 2\lambda$ de forma que las condiciones (3.69) se simplifican a

$$k_{2n+2\lambda} = k, \quad (2n + 2\lambda)k_{2n+2\lambda} = 2m, \quad (3.80)$$

que sólo admite la solución $k_{2n+2\lambda} = k = 1$ cuando $m = n + \lambda$. En el segundo término las derivadas que no se anulan son las de j par, $2 \leq j \leq 2\lambda - 2$, por lo que las condiciones (3.69) toman la forma

$$\sum_{r=1}^{\lambda-1} k_{2r} = k, \quad \sum_{s=1}^{\lambda-1} s k_{2s} = m. \quad (3.81)$$

Así, (3.79) se reduce a

$$\begin{aligned} & \frac{d^{2m}}{dz^{2m}} \left(\log \sum_{\nu=0}^{\lambda-1} \alpha_{\nu,n}^{(\lambda)} (z^{2n+2\lambda+2\nu} - z^{2\lambda-2\nu-2}) \right) \Big|_{z=0} = -(2n + 2\lambda)! \frac{\alpha_{0,n}^{(\lambda)}}{\alpha_{\lambda-1,n}^{(\lambda)}} \delta_{m,n+\lambda} \\ & - (2m)! \sum_{k=1}^{2m} \frac{(k-1)!}{(\alpha_{\lambda-1,n}^{(\lambda)})^k} \sum_{k_2, k_4, \dots, k_{2(\lambda-1)}} \prod_{j=1}^{2m} \frac{(-\alpha_{\lambda-1-j,n}^{(\lambda)})^{k_{2j}}}{(k_{2j})!}. \end{aligned} \quad (3.82)$$

Finalmente, la expresión anterior se puede simplificar para obtener (3.76) utilizando las condiciones (3.81) para escribir k_2 y k_4 en términos de los restantes índices, i.e.

$$\begin{aligned} k_2 &= k_6 + 2k_8 + \dots + (\lambda - 3)k_{2(\lambda-1)} + 2k - m, \\ k_4 &= -2k_6 - 3k_8 - \dots - (\lambda - 2)k_{2(\lambda-1)} + m - k. \end{aligned} \quad (3.83)$$

Nótese que en (3.76) está garantizado el cumplimiento de las condiciones (3.81) porque para los valores de los índices que no las satisfacen se obtiene el inverso del factorial de un número negativo, lo que se considera igual a cero ya que cuando $x \rightarrow -n$ se tiene que $1/\Gamma(x+1) \rightarrow 0$. Se ha cambiado el límite superior de la suma sobre k de $2m$ a m porque cuando $m+1 \leq k \leq 2m$ las condiciones (3.81) no se cumplen.

En el caso $\lambda = 3$, (3.76) se reduce a

$$I_{m,n}^{(3)} = \frac{5\pi}{m} - \pi \left(\frac{\alpha_{0,n}^{(3)}}{\alpha_{1,n}^{(3)}} \right)^m \sum_{k=1}^m \frac{(k-1)!}{(2k-m)!(m-k)!} \left(-\frac{(\alpha_{1,n}^{(3)})^2}{\alpha_{2,n}^{(3)}\alpha_{0,n}^{(3)}} \right)^k - \pi \frac{\alpha_{0,n}^{(3)}}{\alpha_{2,n}^{(3)}} \delta_{m,n+3}, \quad (3.84)$$

por lo que necesitamos evaluar una suma de la forma

$$\sum_{k=1}^m \frac{(k-1)!}{(2k-m)!(m-k)!} x^k = \sum_{j=0}^{m-1} \frac{(m-j-1)!}{j!(m-2j)!} x^{m-j} = x^m \sum_{j=0}^{m-1} \binom{m-j}{j} \frac{(x^{-1})^j}{m-j}. \quad (3.85)$$

Esto se puede llevar a cabo utilizando que

$$\sum_{j=0}^{m-1} \binom{m-j}{j} \frac{m}{m-j} z^j = \left(\frac{1+\sqrt{1+4z}}{2} \right)^m + \left(\frac{1-\sqrt{1+4z}}{2} \right)^m, \quad (3.86)$$

lo cual se demuestra en [150, Eq. (5.75)] mediante la teoría de las series binomiales generalizadas. En el apéndice B damos una prueba alternativa de esta identidad utilizando propiedades de las funciones hipergeométricas. Ahora, utilizando (3.86) junto con la segunda ecuación en (3.28), se encuentra que

$$I_{m,n}^{(3)} = \frac{5\pi}{m} - \frac{\pi}{m} [f(n)^m + f^*(n)^m] - \pi \frac{(n+4)(n+5)}{(n+1)(n+2)} \delta_{m,n+3}, \quad (3.87)$$

que es equivalente a (3.77). \square

3.1.3. Resultados

Puesto que hemos sido capaces de obtener fórmulas analíticas cerradas para $I_{m,n}^{(\lambda)}$ cuando $\lambda = 1, 2, 3$, las ecuaciones (3.46), (3.67), (3.70) y (3.77) nos conducirán igualmente a fórmulas cerradas para $E(C_n^{(\lambda)})$ en estos casos, siempre y cuando podamos llevar a cabo la suma en (3.44). Como mostramos a continuación esto es efectivamente posible.

Cuando $\lambda = 1$, después de sustituir los correspondientes valores de las constantes $d_{m,n}^{(\lambda)}$, $c_n^{(\lambda)}$ y $\alpha_{\nu,n}^{(\lambda)}$ (véase (3.26) y (3.28)), (3.44) se reduce a

$$E(C_n^{(1)}) = -\frac{1}{2} \left(I_{0,n}^{(1)} - I_{n+1,n}^{(1)} \right), \quad (3.88)$$

lo que mediante (3.46) y (3.67) inmediatamente lleva a

$$E(C_n^{(1)}) = E(U_n) = \frac{\pi}{2} \left(\frac{1}{n+1} - 1 \right). \quad (3.89)$$

Cuando $\lambda = 2$, (3.44) toma la forma

$$E(C_n^{(2)}) = -\frac{1}{8} \left((n+1)(n+3)I_{0,n}^{(2)} - (n+1)^2 I_{n+2,n}^{(2)} - 4 \sum_{m=1}^{n+1} m I_{m,n}^{(2)} \right), \quad (3.90)$$

así que utilizando (3.46) y (3.70) conjuntamente con la bien conocida fórmula para la serie geométrica,

$$\sum_{m=1}^n x^m = \frac{x(1-x^n)}{1-x}, \quad (3.91)$$

se encuentra que

$$E(C_n^{(2)}) = -\frac{\pi}{8} \left(2(n+1)(n+3) \log(n+1) + \frac{n^3 - 5n^2 - 29n - 27}{n+2} + \frac{(n+3)^{n+3}}{(n+2)(n+1)^{n+1}} \right). \quad (3.92)$$

Recordando (3.10), se comprueba que (3.89) y (3.92) son equivalentes a (3.13) y (3.16), respectivamente.

En el caso $\lambda = 3$, (3.44) se reescribe como

$$E(C_n^{(3)}) = -\frac{1}{128} \left[(n+1)(n+2)(n+4)(n+5)I_{0,n}^{(3)} - (n+1)^2(n+2)^2 I_{n+3,n}^{(3)} - 12 \sum_{m=1}^{n+2} m(n^2 + 6n + 7 - 2m^2) I_{m,n}^{(3)} \right]. \quad (3.93)$$

Sustituyendo (3.46) y (3.77) en esta expresión, se encuentra de nuevo la suma geométrica, así como una suma de la forma $\sum_m m^2 x^m$. Utilizando (3.91) y la fórmula [151, Eq. (5.14.9)]

$$\sum_{m=1}^n m^2 x^m = \frac{x(1+x) - x^{n+1} [(n+1)^2 - (2n^2 + 2n - 1)x + n^2 x^2]}{(1-x)^3}, \quad (3.94)$$

después de un cálculo tedioso pero directo se encuentra la siguiente fórmula analítica cerrada para $E(C_n^{(3)})$, que es un resultado nuevo:

$$E(C_n^{(3)}) = -\frac{\pi}{128} \left\{ 2(n+1)(n+2)(n+4)(n+5) \log \left(\frac{(n+1)(n+2)}{2} \right) + \frac{n^5 - 16n^4 - 269n^3 - 1200n^2 - 2102n - 1250}{n+3} + \frac{2(n+5)^2}{(n+2)(n+3)} \operatorname{Re} \left[\left(\frac{(n+1)(n+5) + i\sqrt{3(n+1)(n+5)}}{(n+1)(n+2)} \right)^{n+1} \times \left(2n^2 + 13n + 14 - i(n+1)(n+6) \sqrt{\frac{(n+1)(n+5)}{3}} \right) \right] \right\}. \quad (3.95)$$

Cuando $\lambda \geq 4$ no parece muy probable lograr expresar las derivadas de $I_{n+\lambda,n}^{(\lambda)}$ en forma cerrada, por lo que no se puede esperar una fórmula de este tipo para $E(C_n^{(\lambda)})$. No obstante, el método que hemos desarrollado aquí no está exento de utilidad en este caso. Esto se debe a que combinando las ecuaciones (3.44) y (3.76) se obtiene una expresión para la entropía $E(C_n^{(\lambda)})$ en términos de sumas con un número finito de sumandos. Por brevedad, en (3.44) conviene absorber el término correspondiente a $I_{n+\lambda,n}^{(\lambda)}$ en la suma sobre m fijando

$$\beta_{n+\lambda,n}^{(\lambda)} := -\alpha_{\lambda-1,n}^{(\lambda)} d_{n,n}^{(\lambda)} \quad (3.96)$$

en lugar de usar (3.45) para $m = n + \lambda$, que daría $\beta_{n+\lambda,n}^{(\lambda)} = -\alpha_{\lambda-1,n}^{(\lambda)} d_{n,n}^{(\lambda)} + \alpha_{0,n}^{(\lambda)} d_{n+\lambda,n}^{(\lambda)}$. De esta manera se tiene que

$$\begin{aligned} E(C_n^{(\lambda)}) = & -\frac{\pi}{2} c_n^{(\lambda)} \left\{ 2\alpha_{0,n}^{(\lambda)} d_{0,n}^{(\lambda)} \log \left(\frac{(\lambda)_n}{n!} \right) + \alpha_{0,n}^{(\lambda)} d_{n,n}^{(\lambda)} + (2\lambda - 1) \sum_{m=1}^{n+\lambda} \frac{\beta_{m,n}^{(\lambda)}}{m} \right. \\ & - \sum_{m=1}^{n+\lambda} \sum_{k=1}^m \sum_{k_6=0}^{[m/3]} \sum_{k_8=0}^{[m/4]} \cdots \sum_{k_{2(\lambda-1)}=0}^{[m/(\lambda-1)]} \beta_{m,n}^{(\lambda)} \left(\frac{\alpha_{\lambda-3,n}^{(\lambda)}}{\alpha_{\lambda-2,n}^{(\lambda)}} \right)^m \left(-\frac{(\alpha_{\lambda-2,n}^{(\lambda)})^2}{\alpha_{\lambda-1,n}^{(\lambda)} \alpha_{\lambda-3,n}^{(\lambda)}} \right)^k \\ & \times \frac{(k-1)!}{(2k-m+\sum_{r=1}^{\lambda-3} r k_{2r+4})! (m-k-\sum_{s=1}^{\lambda-3} (s+1) k_{2s+4})!} \\ & \left. \times \prod_{j=3}^{\lambda-1} \frac{1}{(k_{2j})!} \left(-\frac{\alpha_{\lambda-1-j,n}^{(\lambda)} (\alpha_{\lambda-2,n}^{(\lambda)})^{j-2}}{(\alpha_{\lambda-3,n}^{(\lambda)})^{j-1}} \right)^{k_{2j}} \right\}. \quad (3.97) \end{aligned}$$

A diferencia de los resultados previos (3.17) y el derivado de (3.22) de [140] y [142] respectivamente, (3.97) es, a pesar de su forma enrevesada, completamente analítica para todo $\lambda \in \mathbb{N}$, lo que permite utilizarla para obtener expresiones exactas para la entropía utilizando computación simbólica. Como ejemplo, damos en las tablas a continuación (cuadros 3.1 y 3.2) las expresiones obtenidas para $E(C_n^{(4)})$ y $E(C_n^{(5)})$, con $1 \leq n \leq 15$, mediante una implementación en Maple de (3.97). En estas tablas indicamos tanto el valor exacto como el numérico (que se obtiene a partir de éste) para que el lector interesado pueda comparar con los obtenidos mediante el algoritmo numérico diseñado en [64] para el cálculo de las entropías de los polinomios de Jacobi que ya mencionamos en la introducción. Creemos que la fórmula (3.97) puede complementar en la práctica a este acercamiento numérico.

3.2. Entropía de polinomios de Jacobi de parámetros semienteros

Algunos de los resultados de la entropía de los polinomios de Gegenbauer se pueden extender al caso más general de los polinomios de Jacobi. Una posibilidad

Cuadro 3.1: Valores exactos y numéricos de la entropía $E(C_n^{(4)})$ para $1 \leq n \leq 15$.

n	Valor exacto	Valor numérico
1	$-7\pi \log(2) + \frac{119}{240}\pi$	-13,685
2	$-\frac{105}{8}\pi \log(10) + \frac{580771}{300000}\pi$	-88,862
3	$-\frac{75}{2}\pi \log(20) + \frac{95}{21}\pi$	-338,714
4	$-\frac{5775}{64}\pi \log(35) + \frac{4883222845}{632481024}\pi$	-983,613
5	$-\frac{385}{2}\pi \log(56) + \frac{17355685}{1806336}\pi$	-2404,173
6	$-\frac{3003}{8}\pi \log(84) + \frac{6449434961}{1058158080}\pi$	-5206,005
7	$-\frac{1365}{2}\pi \log(120) - \frac{1396715852287}{139218750000}\pi$	-10296,556
8	$-\frac{75075}{64}\pi \log(165) - \frac{24757176334716125}{493018566815808}\pi$	-18974,368
9	$-1925\pi \log(220) - \frac{1200329915}{9135984}\pi$	-33031,075
10	$-\frac{12155}{4}\pi \log(286) - \frac{325291539600149215255}{1172732412725203616}\pi$	-54866,421
11	$-4641\pi \log(364) - \frac{31458443588344487293819}{60436675052957701680}\pi$	-87616,538
12	$-\frac{440895}{64}\pi \log(455) - \frac{25537984326378849719971131}{28270687046875000000000}\pi$	-135295,739
13	$-9975\pi \log(560) - \frac{1779685691911133495}{1202109806542848}\pi$	-202952,031
14	$-\frac{56525}{4}\pi \log(680) - \frac{36234350694889865223938313068785}{15613637127259094259005915136}\pi$	-296836,555
15	$-19635\pi \log(816) - \frac{130243656594168370141034405}{37115886521993021558784}\pi$	-424587,139

Cuadro 3.2: Valores exactos y numéricos de la entropía $E(C_n^{(5)})$ para $1 \leq n \leq 15$.

n	Valor exacto	Valor numérico
1	$-\frac{525}{128} \pi \log(5) + \frac{945}{1024} \pi$	-17,839
2	$-\frac{2475}{128} \pi \log(15) + \frac{27685925}{5225472} \pi$	-147,857
3	$-\frac{17325}{256} \pi \log(35) + \frac{61634724075}{3373232128} \pi$	-698,499
4	$-\frac{25025}{128} \pi \log(70) + \frac{5573831525}{115605504} \pi$	-2457,981
5	$-\frac{63063}{128} \pi \log(126) + \frac{338107973281463}{3173748645888} \pi$	-7150,909
6	$-\frac{143325}{128} \pi \log(210) + \frac{20887195}{101376} \pi$	-18162,369
7	$-\frac{75075}{32} \pi \log(330) + \frac{1408430247274269205}{3944148534526464} \pi$	-41620,201
8	$-\frac{294525}{64} \pi \log(495) + \frac{806559968327725}{1438588584576} \pi$	-87940,792
9	$-\frac{546975}{64} \pi \log(715) + \frac{29915266041851863399425}{37527437207206515712} \pi$	-173958,634
10	$-\frac{969969}{64} \pi \log(1001) + \frac{97664804776687286561309}{96698680084732322688} \pi$	-325775,232
11	$-\frac{6613425}{256} \pi \log(1365) + \frac{230209361727271224010045}{212679240849405517824} \pi$	-582478,486
12	$-\frac{2723175}{64} \pi \log(1820) + \frac{10188450005911283085}{12635587626401792} \pi$	-1000899,539
13	$-\frac{4352425}{64} \pi \log(2380) - \frac{39663465263970548089202600252605}{249818194036145508144094642176} \pi$	-1661590,212
14	$-\frac{6774075}{64} \pi \log(3060) - \frac{717231543067734581588054629334401175}{306761704661739640893688309874688} \pi$	-2676220,464
15	$-\frac{5148297}{32} \pi \log(3876) - \frac{535111116210266542852402527915814650511}{82233794352493419438828330115762176} \pi$	-4196611,889

de este tipo se apuntó en [145] ya que las transformaciones cuadráticas [146]

$$P_{2n}^{(m-1/2, m-1/2)}(x) = \frac{\Gamma(2n+m+1/2)n!}{\Gamma(n+m+1/2)(2n)!} P_n^{(m-1/2, -1/2)}(2x^2-1), \quad (3.98)$$

$$P_{2n+1}^{(m-1/2, m-1/2)}(x) = \frac{\Gamma(2n+m+3/2)n!}{\Gamma(n+m+1/2)(2n+1)!} x P_n^{(m-1/2, 1/2)}(2x^2-1), \quad (3.99)$$

permiten escribir $E(P_n^{(\pm 1/2, m-1/2)})$ en términos de $E(C_n^{(m)})$. En esta sección mostraremos que podemos extender el método utilizado en la sección anterior para calcular la entropía de los polinomios de Gegenbauer de parámetro entero al caso de polinomios de Jacobi de parámetros semienteros, de los que los primeros son un caso particular. A lo largo de nuestros desarrollos será útil tener en cuenta que la entropía en este caso es una función simétrica de los parámetros del polinomio, i.e.

$$E(P_n^{(\alpha, \beta)}) = E(P_n^{(\beta, \alpha)}), \quad (3.100)$$

lo que se deduce de la propiedad $P_n^{(\alpha, \beta)}(-x) = (-1)^n P_n^{(\beta, \alpha)}(x)$. Para abreviar $\omega_{\alpha\beta}$ denotará el peso de ortogonalidad de los polinomios de Jacobi, i.e.

$$\omega_{\alpha\beta}(x) = (1-x)^\alpha(1+x)^\beta. \quad (3.101)$$

3.2.1. Representaciones trigonométricas para los polinomios de Jacobi

Mientras que en los anteriores casos las representaciones trigonométricas de los polinomios estaban disponibles en la literatura especializada en funciones especiales y polinomios ortogonales, el caso de los polinomios de Jacobi no se encuentra desarrollado. Sin embargo, como veremos a continuación es posible obtener representaciones análogas también en este caso. Nuestra principal herramienta es la siguiente proposición ([146], capítulo 7, ejercicio 26), que el lector podrá probar sin demasiada dificultad.

Proposición 3.4. Sean $\{p_n(x)\}$ y $\{q_n(x)\}$ secuencias de polinomios ortogonales en Δ con respecto a las funciones peso $\omega(x)$ y $\Omega(x)$ respectivamente, tales que

$$q_n(x) = \sum_{k=0}^n c_{kn} p_k(x). \quad (3.102)$$

Entonces, se tiene que

$$\omega(x)p_k(x) = \|p_k\|_\omega^2 \sum_{n=k}^{\infty} \frac{c_{kn}}{\|q_n\|_\Omega^2} q_n(x)\Omega(x),$$

donde

$$\|p_k\|_\omega^2 = \int_{\Delta} (p_k(x))^2 \omega(x) dx, \quad \|q_n\|_\Omega^2 = \int_{\Delta} (q_n(x))^2 \Omega(x) dx.$$

Las relaciones de la forma (3.102) se conocen como fórmulas de conexión. Éstas han sido estudiadas concienzudamente en la literatura y se conocen muchas fórmulas explícitas de este tipo para los polinomios ortogonales clásicos (véase por ejemplo [146], capítulo 7.1). Así, utilizando una de estas fórmulas, es posible obtener una representación trigonométrica para $\omega(x)p_n(x)$ con la ayuda de la Proposición 3.4 si se conoce una representación trigonométrica para q_n . Por ejemplo, consideremos la siguiente fórmula de conexión [146]

$$P_n^{(\alpha,\delta)}(x) = \frac{(\alpha+1)_n}{(\alpha+\beta+2)_n} \sum_{k=0}^n \frac{(-1)^{n-k}(\delta-\beta)_{n-k}(\alpha+\beta+1)_k(\alpha+\beta+2k+1)(\alpha+\delta+n+1)_k}{(n-k)!(\alpha+1)_k(\alpha+\beta+1)(\alpha+\beta+n+2)_k} P_k^{(\alpha,\beta)}(x). \quad (3.103)$$

Fijemos $\alpha = \delta = m - 1/2$ para, con la ayuda de (3.4), llegar a

$$C_n^{(m)}(x) = \frac{(2m)_n}{(m+\beta+3/2)_n} \sum_{k=0}^n P_k^{(m-1/2,\beta)}(x) \times \frac{(-1)^{n-k}(m-\beta-1/2)_{n-k}(m+\beta+1/2)_k(m+\beta+2k+1/2)(2m+n)_k}{(n-k)!(m+1/2)_k(m+\beta+1/2)(m+\beta+n+3/2)_k}. \quad (3.104)$$

Ahora, la aplicación de la Proposición 3.4 tras un cálculo tedioso pero directo nos conduce a

$$\omega_{m-1/2,\beta}(x)P_k^{(m-1/2,\beta)}(x) = \frac{\Gamma(m)\Gamma(k+m)\Gamma(k+m+1/2)\Gamma(k+\beta+1)}{2^{1/2-3m-2k-\beta}\pi\Gamma(k+1)\Gamma(m+\beta+2k+3/2)} \sum_{n=0}^{\infty} \frac{(-1)^n(m-1/2-\beta)_n(2m+2k)_n(n+m+k)(n+k)!}{n!(m+\beta+2k+3/2)_n\Gamma(n+k+2m)} C_{n+k}^{(m)}(x)(1-x^2)^{m-1/2}, \quad (3.105)$$

donde hemos usado

$$(a)_{n+k} = (a)_k(a+k)_n, \quad (a+n)_k = \frac{(a+k)_n(a)_k}{(a)_n}, \quad (2a)_{2n} = 2^{2n}(a)_n(a+1/2)_n. \quad (3.106)$$

Puesto que buscamos representaciones que contengan un número finito de sumandos debemos restringir los posibles valores de los parámetros para poder truncar la suma infinita. Esto se logra considerando $\beta = q - 1/2$, $m, q \in \mathbb{N}$ y $q \geq m$ dado que de esta forma el factor $(m-1/2-\beta)_n$ se anula cuando $n > q-m$. Nótese que, de acuerdo a la propiedad de simetría (3.100), la restricción $q \geq m$ es irrelevante de cara a la evaluación de la entropía. Por simplicidad, reescribamos (3.105) como

$$\omega_{m-1/2,q-1/2}(x)P_k^{(m-1/2,q-1/2)}(x) = a_k^{(m,q)} \sum_{n=0}^{q-m} b_{nk}^{(m,q)} C_{n+k}^{(m)}(x)(1-x^2)^{m-1/2}, \quad (3.107)$$

donde

$$a_k^{(m,q)} = \frac{2^{3m+2k+q-1}\Gamma(m)\Gamma(k+m)\Gamma(k+m+1/2)\Gamma(k+q+1/2)}{\pi\Gamma(k+1)\Gamma(m+q+2k+1)} \quad (3.108)$$

y

$$b_{nk}^{(m,q)} = \frac{(-1)^n(m-q)_n(2m+2k)_n(n+m+k)(n+k)!}{n!(m+q+2k+1)_n\Gamma(n+k+2m)}. \quad (3.109)$$

Ahora, si añadimos la restricción $m \neq 0^4$ con el objetivo de utilizar la representación (3.29) para $C_{n+k}^{(m)}(x)(1-x^2)^{m-1/2}$, encontramos que, si $x = \cos \theta$,

$$\begin{aligned} \omega_{m-1/2,q-1/2}(x)P_k^{(m-1/2,q-1/2)}(x) &= \\ &= a_k^{(m,q)} \sum_{\nu=0}^{m-1} \sum_{n=2\nu}^{q-m+2\nu} b_{n-2\nu,k}^{(m,q)} c_{n-2\nu+k}^{(m)} \alpha_{\nu,n-2\nu+k}^{(m)} \sin(n+k+1)\theta = \\ &= a_k^{(m,q)} \sum_{n=0}^{q+m-2} \left(\sum_{\nu=0}^{m-1} b_{n-2\nu,k}^{(m,q)} c_{n-2\nu+k}^{(m)} \alpha_{nu,n-2nu+k}^{(m)} \right) \sin(n+k+1)\theta. \end{aligned} \quad (3.110)$$

El intercambio en el orden de sumación se lleva a cabo transformando $\sum_{n=2\nu}^{q-m+2\nu}$ en $\sum_{n=0}^{q+m-2}$. El cambio en el límite superior se justifica porque $b_{nk}^{(m,q)}$ se anula cuando $n > q - m$; sin embargo, $b_{nk}^{(m,q)}$ no se anula cuando $n < 0$, por tanto, debemos imponer $b_{nk}^{(m,q)} = 0$ si $n < 0$ en (3.109) para permitir el cambio en el límite inferior. Así, se obtiene la siguiente representación trigonométrica para polinomios de Jacobi de parámetros semienteros, donde renombramos los coeficientes y reetiquetamos los índices para simplificar

$$\omega_{m-1/2,q-1/2}(\cos \theta)P_n^{(m-1/2,q-1/2)}(\cos \theta) = c_n^{(m,q)} \sum_{k=0}^{m+q-2} \alpha_{kn}^{(m,q)} \sin(n+k+1)\theta, \quad (3.111)$$

donde

$$c_n^{(m,q)} = \frac{2^{3m+2n+q-1}\Gamma(m)\Gamma(n+m)\Gamma(n+m+1/2)\Gamma(n+q+1/2)}{\pi\Gamma(n+1)\Gamma(m+q+2n+1)}, \quad (3.112)$$

$$\alpha_{kn}^{(m,q)} = \frac{(-1)^k 2^{2-2m}}{\Gamma(m)} \sum_{\nu=0}^{m-1} \beta_{\nu kn}^{(m,q)} \quad (3.113)$$

y

$$\beta_{\nu kn}^{(m,q)} = \begin{cases} 0 & (k < 2\nu) \\ \frac{(1-m)_\nu \Gamma(n+k-\nu+1)(m-q)_{k-2\nu} (2m+2n)_{k-2\nu} (n+m+k-2\nu)}{\nu! \Gamma(n+k+m+1-\nu) \Gamma(k-2\nu+1) (2n+m+q+1)_{k-2\nu}} & (k \geq 2\nu) \end{cases} \quad (3.114)$$

⁴Esta restricción no es demasiado importante pues significa que estamos perdiendo el caso de los polinomios de Jacobi para los cuales un parámetro es $-1/2$. Sin embargo, como ya mencionamos, fórmulas para la entropía de estos polinomios se pueden obtener a partir de la transformación cuadrática (3.98) y los resultados previos para los polinomios de Gegenbauer.

La representación (3.111) es la generalización de la representación de Szegő (3.29) para polinomios de Gegenbauer. Así, mientras que la Proposición 3.4 es útil para obtener representaciones trigonométricas de polinomios conjuntamente con su peso de ortogonalidad, siguiendo nuestros pasos anteriores, sería conveniente hallar un análogo en este caso de la representación estándar (3.25), lo cual puede hacerse trabajando únicamente sobre la fórmula de conexión. Por ejemplo, se tiene que [146]

$$P_n^{(\gamma, \delta)}(x) = \sum_{k=0}^n c_{nk} P_k^{(\alpha, \beta)}(x) \quad (3.115)$$

con

$$c_{nk} = \frac{(n + \gamma + \delta + 1)_k (k + \gamma + 1)_{n-k} (2k + \alpha + \beta + 1) \Gamma(k + \alpha + \beta + 1)}{(n - k)! \Gamma(2k + \alpha + \beta + 2)} \\ \times {}_3F_2 \left(\begin{matrix} -n + k, n + k + \gamma + \delta + 1, k + \alpha + 1 \\ k + \gamma + 1, 2k + \alpha + \beta + 2 \end{matrix} ; 1 \right). \quad (3.116)$$

Fijemos $\gamma = m - 1/2$, $\delta = q - 1/2$ y $\alpha = \beta = 1/2$ y usemos las segundas ecuaciones de (3.11) y (3.24) de forma que

$$P_n^{(m-1/2, q-1/2)}(\cos \theta) = \sum_{k=0}^n d_{kn}^{(m, q)} \frac{\sin(k+1)\theta}{\sin \theta}, \quad (3.117)$$

donde

$$d_{kn}^{(m, q)} = \frac{(n + m + q)_k (k + m + 1/2)_{n-k}}{(n - k)! k! 2^{2k}} \\ \times {}_3F_2 \left(\begin{matrix} -n + k, n + k + m + q, k + 3/2 \\ k + m + 1/2, 2k + 3 \end{matrix} ; 1 \right). \quad (3.118)$$

Se pueden hallar otras representaciones trigonométricas para los polinomios de Jacobi siguiendo estas ideas, pero para nuestro objetivo bastarán (3.111) y (3.117).

3.2.2. Evaluación de la integral entrópica. Resultados

Puesto que las representaciones (3.111) y (3.117) son análogas a las del caso Gegenbauer (3.29) y (3.25), los pasos en este caso vienen a ser una repetición de los de la sección anterior. Por ello, y para no aburrir al lector, intentaremos desarrollar este apartado de la forma más escueta posible.

Llevando a cabo el cambio de variable $x = \cos \theta$ en la integral entrópica (3.5) y sustituyendo $\omega_{m-1/2, q-1/2}(\cos \theta) P_n^{(m-1/2, q-1/2)}(\cos \theta)$ y $P_n^{(m-1/2, q-1/2)}(\cos \theta)$ por sus respectivas representaciones trigonométricas

(3.111) y (3.117) encontramos que

$$\begin{aligned}
E(P_n^{(m-1/2, q-1/2)}) &= -c_n^{(m, q)} \sum_{s=0}^n \sum_{k=0}^{m+q-2} \alpha_{k, n}^{(m, q)} d_{s, n}^{(m, q)} \\
&\times \int_0^\pi \sin(n+k+1)\theta \sin(s+1)\theta \log(P_n^{(m-1/2, q-1/2)}(\cos \theta))^2 d\theta = \\
&= -\frac{c_n^{(m, q)}}{2} \sum_{s=0}^n \sum_{k=0}^{m+q-2} \alpha_{k, n}^{(m, q)} d_{s, n}^{(m, q)} \\
&\times \int_0^\pi (\cos(n+k-s)\theta - \cos(n+k+s+2)\theta) \log(P_n^{(m-1/2, q-1/2)}(\cos \theta))^2 d\theta
\end{aligned} \tag{3.119}$$

Definiendo

$$I_{r, n}^{(m, q)} := \int_0^\pi \cos r\theta \log(P_n^{(m-1/2, q-1/2)}(\cos \theta))^2 d\theta, \tag{3.120}$$

$E(P_n^{(m-1/2, q-1/2)})$ se reescribe como

$$E(P_n^{(m-1/2, q-1/2)}) = -\frac{c_n^{(m, q)}}{2} \sum_{s=0}^n \sum_{k=0}^{m+q-2} \alpha_{k, n}^{(m, q)} d_{s, n}^{(m, q)} \left(I_{n+k-s, n}^{(m, q)} - I_{n+k+s+2, n}^{(m, q)} \right). \tag{3.121}$$

Ahora, tenemos que evaluar como anteriormente las integrales $I_{r, n}^{(m, q)}$ con $0 \leq r \leq 2n + m + q$, lo que hacemos a continuación en un análogo del Teorema 3.1.

Teorema 3.3. Para $m, q \in \mathbb{N}$ ($q \geq m$),

$$I_{0, n}^{(m, q)} = 2\pi \log \left| 2^{m+q-2} c_n^{(m, q)} \alpha_{m+q-2, n}^{(m, q)} \right| \tag{3.122}$$

y, cuando $r \geq 1$,

$$\begin{aligned}
I_{r, n}^{(m, q)} &= \frac{\pi}{r} [2m - 1 + (-1)^r (2q - 1)] \\
&+ \frac{\pi}{r!} \frac{d^r}{dz^r} \left(\log \sum_{k=0}^{m+q-2} \alpha_{k, n}^{(m, q)} (z^{2n+m+q+k} - z^{m+q-k-2}) \right) \Big|_{z=0}.
\end{aligned} \tag{3.123}$$

Demostración. Siguiendo los pasos de la demostración del Teorema 3.1 utilizamos la representación (3.111) y el cambio de variable $z = \exp(i\theta)$, para hallar

$$\mathcal{I}_{m, n}^{(\lambda)} = \frac{1}{2i} \oint_{|z|=1} \frac{z^{2r} + 1}{z^{r+1}} \log \phi(z) dz, \tag{3.124}$$

donde

$$\phi(z) = \frac{c_n^{(m, q)} (-1)^{m+1} 2^{m+q-2}}{(z-1)^{2m-1} (z+1)^{2q-1}} \sum_{k=0}^{m+q-2} \alpha_{k, n}^{(m, q)} (z^{2n+m+q+k} - z^{m+q-k-2}) \tag{3.125}$$

y como anteriormente $I_{m,n}^{(\lambda)} = \text{Re}\mathcal{I}_{m,n}^{(\lambda)}$. Deformando el contorno de integración como en el caso previo se encuentra que

$$\mathcal{I}_{r,n}^{(m,q)} = \frac{\pi}{r!} \frac{d^r}{dz^r} [(z^{2r} + 1) \log \phi(z)] \Big|_{z=0}, \quad (3.126)$$

de donde se deducen (3.122) y (3.123). \square

Una vez transformadas las integrales en derivadas, como en la sección previa, utilizamos la fórmula de Faà di Bruno (3.68) con $f(z) = \log z$ y $g(z) = \sum_{k=0}^{m+q-2} \alpha_{k,n}^{(m,q)} (z^{2n+m+q+k} - z^{m+q-k-2})$, para hallar que (3.47) se reescribe como

$$\begin{aligned} I_{r,n}^{(m,q)} &= \frac{\pi}{r} [2m - 1 + (-1)^r (2q - 1)] - \pi \frac{\alpha_{0,n}^{(m,q)}}{\alpha_{m+q-2,n}^{(m,q)}} \delta_{r,2n+m+q-} \\ &\quad - \pi \sum_{k=1}^r \sum_{k_1=0}^r \sum_{k_2=0}^{\lfloor r/2 \rfloor} \cdots \sum_{k_{m+q-2}=0}^{\lfloor r/(m+q-2) \rfloor} \frac{(k-1)!}{(\alpha_{m+q-2,n}^{(m,q)})^k} \prod_{j=1}^{m+q-2} \frac{(-\alpha_{m+q-2-j,n}^{(m,q)})^{k_j}}{k_j!} \end{aligned} \quad (3.127)$$

donde $1 \leq r \leq 2n + m + q$ y

$$\sum_{i=1}^{m+q-2} k_i = k, \quad \sum_{i=1}^{m+q-2} ik_i = r. \quad (3.128)$$

Como vimos anteriormente, cuando el número de sumas es pequeño éstas pueden expresarse en forma cerrada. Así, si $m + q = 2$ no hay sumas y (3.127) es ya una expresión cerrada. No obstante, esto implica que $m = 1$ y $q = 1$, o sea, $P_n^{(1/2,1/2)}$, que se corresponde con el ya trabajado caso del polinomio de Chebyshev de segunda especie (polinomio de Gegenbauer de parámetro 1).

Cuando $m + q = 3$, sólo tenemos que sumar sobre k y k_1 y, además, (3.128) se reduce a $k_1 = k = r$, por lo que los índices están fijos y en realidad no hay suma. Nótese que $m + q = 3$ implica que $m = 1$ y $q = 2$, por lo que en este caso podríamos desarrollar una fórmula cerrada para $E(P_n^{(1/2,3/2)})$, la cual, sin embargo, también puede obtenerse utilizando la transformación cuadrática (3.99) a partir de la expresión de $E(C_n^{(2)})$. De todas formas, para el lector interesado mostramos brevemente a continuación cómo se obtiene esta fórmula siguiendo nuestro método. En este caso las expresiones (3.113) y (3.118) de los coeficientes de las representaciones trigonométricas se reducen sensiblemente a

$$\alpha_{0,n}^{(1,2)} = 1, \quad \alpha_{1,n}^{(1,2)} = \frac{n+1}{n+2}, \quad d_{k,n}^{(1,2)} = \frac{(3/2)_n}{(3)_n} (-1)^{n-k} (k+1), \quad (3.129)$$

donde, en el último de ellos, hemos utilizado que la ${}_3F_2$ de (3.118) se reduce en este caso a una ${}_2F_1$, la cual se puede evaluar mediante la fórmula de Chu-Vandermonde [146], i.e.

$${}_2F_1 \left(\begin{matrix} -n, a \\ c \end{matrix}; 1 \right) = \frac{(c-a)_n}{(c)_n} \quad (n \in \mathbb{N}). \quad (3.130)$$

Utilizando las simetrías de estos coeficientes la expresión de la entropía (3.121) se puede escribir como

$$E(P_n^{(1/2,3/2)}) = -\frac{c_n^{(1,2)}}{2} \left[\alpha_{0,n}^{(1,2)} d_{n,n}^{(1,2)} I_{0,n}^{(1,2)} + \sum_{s=0}^{n+1} \left(\sum_{k=0}^1 \alpha_{k,n}^{(1,2)} d_{n+k-s,n}^{(1,2)} \right) I_{s,n}^{(1,2)} \right. \\ \left. + \sum_{s=n+2}^{2n+2} \left(\sum_{k=0}^1 \alpha_{k,n}^{(1,2)} d_{s-k-n-2,n}^{(1,2)} \right) I_{s,n}^{(1,2)} + \alpha_{1,n}^{(1,2)} d_{n,n}^{(1,2)} I_{2n+3,n}^{(1,2)} \right]. \quad (3.131)$$

Un sencillo cálculo muestra que

$$\sum_{k=0}^1 \alpha_{k,n}^{(1,2)} d_{n+k-s,n}^{(1,2)} = \frac{(3/2)_n}{(n+2)(3)_n} (-1)^{s+1} s, \quad (3.132)$$

$$\sum_{k=0}^1 \alpha_{k,n}^{(1,2)} d_{s-k-n-2,n}^{(1,2)} = \frac{(3/2)_n}{(n+2)(3)_n} (-1)^s s. \quad (3.133)$$

Por otro lado, (3.122) y (3.127) se transforman en este caso en

$$I_{0,n}^{(1,2)} = 2\pi \log \left(\frac{(2n+2)!}{2^{2n} n! (n+2)!} \right), \quad (3.134)$$

$$I_{r,n}^{(1,2)} = \frac{\pi}{r} + \left(\left(\frac{n+2}{n+1} \right)^r - 3 \right) \frac{(-1)^{r+1} \pi}{r} - \frac{n+2}{n+1} \pi \delta_{r,2n+3}. \quad (3.135)$$

Así, la ecuación (3.131) sólo contiene sumas triviales y la suma geométrica. Después de llevar a cabo los correspondientes cálculos, la expresión de la entropía se reduce a la siguiente fórmula cerrada:

$$E(P_n^{(1/2,3/2)}) = -4 \left(\frac{\Gamma(n+3/2)}{\Gamma(n+3)} \right)^2 \left[2(n+1)(n+2) \log \left(\frac{(2n+2)!}{2^{2n} n! (n+2)!} \right) \right. \\ \left. + n^2 - 4n - 6 + \frac{2(n+1)^{2n+4} + (n+2)^{2n+4}}{(2n+3)(n+1)^{2n+2}} \right]. \quad (3.136)$$

La otra posibilidad para eliminar las sumas en (3.127) es cuando $m+q=4$, cuyo desarrollo es análogo al caso Gegenbauer cuando $\lambda=3$. En este caso (3.128) toman la forma

$$k_1 + k_2 = k, \quad k_1 + 2k_2 = r, \quad (3.137)$$

así que, como anteriormente, k_1 y k_2 se pueden escribir en términos de k y r , i.e.

$$k_1 = 2k - r, \quad k_2 = r - k, \quad (3.138)$$

de forma que sólo resta la ya conocida suma en k (3.85), para la que ya desarrollamos una expresión compacta. De todas formas, las únicas posibilidades para que se cumpla $m+q=4$ son $m=q=2$, que lleva al ya resuelto caso

$E(C_n^{(2)})$, o $m = 1$ y $q = 3$, que conduce a una fórmula para $E(P_n^{(1/2, 5/2)})$, que como ya hemos comentado se puede obtener a partir de $E(C_n^{(3)})$ mediante (3.99).

Cuando $m + q \geq 5$, aunque podemos escribir k_1 y k_2 en términos de los otros índices, restan las sumas cuyos índices van de k_3 a k_{m+q-2} , por lo que, al igual que en la sección previa, parece extremadamente difícil poder extraer en estos casos una fórmula cerrada. Sin embargo, como en el caso Gegenbauer, se pueden combinar (3.127) y (3.121) para alcanzar no sólo una expresión analítica para $E(P_n^{(m-1/2, q-1/2)})$, sino también una forma de obtener valores exactos de esta entropía mediante computación simbólica. Como ejemplo, adjuntamos algunas tablas con valores de la entropía de polinomios de Jacobi de parámetro semientero obtenidas mediante una implementación en Mathematica de estas fórmulas. Como antes, para este objetivo resulta más conveniente reescribir (3.128) como

$$k_1 = 2k - r + \sum_{i=3}^{m+q-2} (i-2)k_i, \quad k_2 = r - k - \sum_{i=3}^{m+q-2} (i-1)k_i, \quad (3.139)$$

para que así (3.127) se reduzca a

$$\begin{aligned} I_{r,n}^{(m,q)} &= \frac{\pi}{r} [2m - 1 + (-1)^r (2q - 1)] - \pi \frac{\alpha_{0,n}^{(m,q)}}{\alpha_{m+q-2,n}^{(m,q)}} \delta_{r, 2n+m+q-} \\ &- \pi \left(\frac{\alpha_{m+q-4,n}^{(m,q)}}{\alpha_{m+q-3,n}^{(m,q)}} \right)^r \sum_{k=1}^r \sum_{k_3=0}^{[r/3]} \cdots \sum_{k_{m+q-2}=0}^{[r/(m+q-2)]} \left(- \frac{(\alpha_{m+q-3,n}^{(m,q)})^2}{\alpha_{m+q-2,n}^{(m,q)} \alpha_{m+q-4,n}^{(m,q)}} \right)^k \\ &\quad (k-1)! \\ &\frac{1}{(2k - r + \sum_{i=3}^{m+q-2} (i-2)k_i)! (r - k - \sum_{i=3}^{m+q-2} (i-1)k_i)!} \\ &\prod_{j=3}^{m+q-2} \frac{1}{k_j!} \left(- \frac{(\alpha_{m+q-3,n}^{(m,q)})^{j-2} \alpha_{m+q-2-j,n}^{(m,q)}}{(\alpha_{m+q-4,n}^{(m,q)})^{j-1}} \right)^{k_j}, \end{aligned} \quad (3.140)$$

de forma que no es necesario verificar que se cumpla (3.128), ya que cuando esto no es así alguno de los factores se va a cero al contener el inverso del factorial de un entero negativo.

Cuadro 3.3: Valores exactos y numéricos de $E(P_n^{(3/2,7/2)})$ para $1 \leq n \leq 10$.

n	Valor exacto	Valor numérico
1	$-\frac{315}{256}\pi \log\left(\frac{7}{4}\right) + \frac{148736383}{1686616064}\pi$	-1,88622
2	$-\frac{693}{512}\pi \log\left(\frac{9}{4}\right) + \frac{744400307509}{12694994583552}\pi$	-3,26402
3	$-\frac{45045}{32768}\pi \log\left(\frac{165}{64}\right) - \frac{7274689276125159}{233727320564531200000}\pi$	-4,09011
4	$-\frac{353925}{262144}\pi \log\left(\frac{715}{256}\right) - \frac{63544801910250084545}{1111556487195273986048}\pi$	-4,53608
5	$-\frac{21900879}{16777216}\pi \log\left(\frac{3003}{1024}\right) - \frac{1660939262673177112483}{15860430960749476577280}\pi$	-4,74126
6	$-\frac{21015995}{16777216}\pi \log\left(\frac{1547}{512}\right) - \frac{151855542354118615163298712800647789}{1069642897404385025419783916016893952}\pi$	-4,79748
7	$-\frac{643089447}{536870912}\pi \log\left(\frac{12597}{4096}\right) - \frac{14062744653941337056251308982673453093}{82593247826975385621808987037478420480}\pi$	-4,7626
8	$-\frac{9826226865}{8589934592}\pi \log\left(\frac{101745}{32768}\right) - \frac{28653598782907703541709864373467607345994709}{149756359509560222426082384732487680000000000}\pi$	-4,67287
9	$-\frac{300245820875}{274877906944}\pi \log\left(\frac{408595}{131072}\right) - \frac{2708390177209545292588995899131488171758813743384091505}{13100551760363121197593191676122897702121495170828468224}\pi$	-4,55105
10	$-\frac{573700476195}{549755813888}\pi \log\left(\frac{408595}{131072}\right) - \frac{556279293455061580836442569848150085551}{255460628234240000000000000000000000000000000000000}\pi$	-4,41159

Cuadro 3.4: Valores exactos y numéricos de $E(P_n^{(5/2,7/2)})$ para $1 \leq n \leq 10$.

n	Valor exacto	Valor numérico
1	$-\frac{35}{32}\pi \log(2) - \frac{16518965}{150994944}\pi$	-2,03804
2	$-\frac{2835}{2048}\pi \log\left(\frac{45}{16}\right) - \frac{105127673499}{880000000000}\pi$	-4,12171
3	$-\frac{12705}{8192}\pi \log\left(\frac{55}{16}\right) - \frac{6788254517}{87241523200}\pi$	-5,77161
4	$-\frac{429429}{262144}\pi \log\left(\frac{1001}{256}\right) - \frac{19362380100474697}{1066752655253569536}\pi$	-6,96046
5	$-\frac{1756755}{1048576}\pi \log\left(\frac{273}{64}\right) - \frac{1188564371237376741349365}{27536118961612901241061376}\pi$	-7,77054
6	$-\frac{56411355}{33554432}\pi \log\left(\frac{4641}{1024}\right) - \frac{32964039221579014211}{331959686846037884928}\pi$	-8,29361
7	$-\frac{223785705}{134217728}\pi \log\left(\frac{4845}{1024}\right) - \frac{2895831374712090560339802358357}{1956894474240000000000000000000000}\pi$	-8,60608
8	$-\frac{14098499415}{8589934592}\pi \log\left(\frac{159885}{32768}\right) - \frac{2849041404043108236793746411450582301}{15077455064403935654635919313469440000}\pi$	-8,76629
9	$-\frac{55245231041}{34359738368}\pi \log\left(\frac{81719}{16384}\right) - \frac{2290775928638516365095486570914881}{10280092002827097961783909731532800}\pi$	-8,81725
10	$-\frac{6905653880125}{4398046511104}\pi \log\left(\frac{5311735}{1048576}\right) - \frac{29053954710689000576207663897279449591843077625}{115989528892619941362060536788936621082862944256}\pi$	-8,79035

Capítulo 4

Cotas más fuertes para una EUR en dimensión finita

Consideremos un sistema cuántico con un espacio de Hilbert asociado de dimensión finita N . Como ya hemos discutido en la introducción, las relaciones de incertidumbre entrópicas dan lugar a una formulación matemática completamente rigurosa del principio de incertidumbre y, además, constituyen herramientas útiles en la Teoría de la Información Cuántica. Sin embargo, a diferencia del caso infinito-dimensional dado por la relación de incertidumbre de Białynicki-Birula-Mycielski (2.34), en este caso la EUR más fuerte conocida hasta la fecha, que es la de Maassen-Uffink (2.39), no es una desigualdad óptima, en el sentido de que la cota para la suma de entropías de un par arbitrario de observables dada por (2.39) no es alcanzada por ningún estado cuántico salvo en el caso particular de observables complementarios. Dicho de otra forma, la cota de la desigualdad de Maassen-Uffink es susceptible de mejora cuando $c \neq 1/\sqrt{N}$. Sería muy conveniente por tanto, desde un punto de vista tanto teórico como práctico, desarrollar una EUR que constituyera una desigualdad óptima, o, al menos, establecer una cota más restrictiva. Como se dijo en la introducción, el primer y más ambicioso objetivo sólo se ha logrado cuando $N = 2$ [46] y parece sumamente complicado poder extender este resultado a dimensión arbitraria. Como muestra de la dificultad de este problema matemático considérese que parte de la cota óptima del caso bidimensional no se ha podido expresar mediante funciones elementales sino que se requiere una función definida implícitamente. Por todo esto, sería ya un logro considerable poder establecer una cota más fuerte que la de Maassen-Uffink, objetivo en el que nos centramos en este capítulo. Una posibilidad para llevar a cabo este fin es considerar otras desigualdades asociadas al principio de incertidumbre que sean lo suficientemente fuertes como para ser más restrictivas que (2.39) en alguna situación. Este es el caso de la relación de incertidumbre de Landau-Pollak (2.42) que, como ya dijimos, no es ni más fuerte ni más débil que la relación de incertidumbre de Maassen-Uffink, puesto que existen distribuciones

de probabilidad permitidas por la primera y prohibidas por la segunda y viceversa, si bien no se han caracterizado hasta la fecha las condiciones que dan lugar a la mayor fortaleza de una relación sobre la otra. La relación de Landau-Pollak satisface, además, muchos de los requerimientos formales de Deutsch para una relación de incertidumbre (ver sección 2.2.2). En concreto, su cota no depende del estado y es función exclusivamente del solapamiento c del par de observables en cuestión, al igual que en las EURs de Deutsch y Maassen-Uffink. Sin embargo, la relación de incertidumbre de Landau-Pollak no descansa sobre la medida de incertidumbre deseada, la entropía de Shannon, y además, como ya aclaramos, a excepción del caso bidimensional no está clara su validez en el caso general ya que, dada la falta de convexidad del funcional de incertidumbre $\arccos \sqrt{M_\infty}$, no es obvia la extensión de la desigualdad del caso de estados puros a estados mezcla, lo que obliga en este último caso a utilizar la desigualdad más débil (2.43). Nuestro objetivo concreto en este capítulo es entonces obtener una EUR a partir de la relación de incertidumbre de Landau-Pollak, lo cual, además de constituir una formulación completamente rigurosa del principio de incertidumbre de acuerdo a la discusión de la sección 2.2.2 y de ser válida tanto para estados puros como mezcla, ofrece la posibilidad de mejorar, al menos parcialmente, la cota para una EUR de Maassen y Uffink. De hecho, estos autores ya muestran [10, 8] que la relación de incertidumbre de Landau-Pollak permite deducir la EUR de Deutsch (2.37) (y es, por tanto, más fuerte). En lo que sigue, demostraremos que la desigualdad de Landau-Pollak en realidad permite obtener una EUR aún más fuerte, más restrictiva incluso que la de Maassen-Uffink para pares de observables tales que

$$c(A, B) \geq \frac{1}{\sqrt{2}} \simeq 0,707. \quad (4.1)$$

Como consecuencia adicional, nuestra discusión clarificará las condiciones bajo las cuales la relación de incertidumbre de Landau-Pollak es más fuerte que la de Maassen-Uffink y viceversa.

Sean entonces A y B observables cualesquiera (no degenerados, como es habitual imponer) asociados a nuestro sistema cuántico de N niveles. El problema matemático que nos planteamos se trata entonces de minimizar la suma de entropías $H(A) + H(B)$ sujeta a la restricción dada por la desigualdad expresada en (2.42). En un principio, la cota obtenida de esta forma será válida sólo para estados puros ya que así sucede con (2.42). Sin embargo, la concavidad de H permite extender directamente la validez de esta cota al caso general de estados mezcla. Para llevar a cabo el problema de minimización dividimos nuestro proceso en dos partes. Dado que la optimización de $H(A) + H(B)$ sólo se restringe por la forma de las probabilidades máximas de A y B , $M_\infty(A)$ y $M_\infty(B)$, está claro que $\min[H(A) + H(B)]$ se atenderá para distribuciones de probabilidad que dadas $M_\infty(A)$ y $M_\infty(B)$ minimizan respectivamente $H(A)$ y $H(B)$. Así pues, en primer lugar minimizaremos la entropía de una distribución de probabilidad para un valor fijo de su probabilidad máxima M_∞ , para luego, en segunda instancia, minimizar la suma de dos entropías mínimas en el sentido anterior sujeta a la restricción de Landau-Pollak (2.42).

De esta forma, primeramente, buscamos los valores mínimos de la función de N variables

$$H(\{p_i\}) = \sum_{i=1}^N p_i \log p_i \quad (4.2)$$

con las restricciones

$$\sum_{i=1}^N p_i = 1, \quad (4.3)$$

$$M_\infty = \underset{i}{\text{máx}} p_i = P, \quad (4.4)$$

Nótese que la máxima probabilidad se puede repetir M veces, con $1 \leq M \leq N$, de manera que la restricción (4.4) es en realidad un conjunto de M ecuaciones que, teniendo en cuenta la irrelevancia del etiquetado de las probabilidades para este problema, se pueden expresar como $p_i = P$ para $i = 1, \dots, M$. Es directo percatarse de que los mínimos valores de H se alcanzan para distribuciones de probabilidad de la forma

$$\{p_i\}_{\min} = \left\{ \underbrace{P, \dots, P}_{M \text{ veces}}, 1 - MP, \underbrace{0, \dots, 0}_{N-M-1 \text{ veces}} \right\}. \quad (4.5)$$

Esto se debe a que en nuestro problema los M primeros sucesos tienen fijada su probabilidad a P , restando por distribuir entonces la cantidad de probabilidad $1 - MP$ entre los $N - M$ sucesos restantes. Ahora bien, esto se puede hacer concentrando toda esta probabilidad restante en un solo suceso como en (4.5), o repartiéndola entre varios sucesos. Sin embargo, distribuciones de probabilidad de este último tipo se obtienen a partir de un proceso de transferencia de la primera. Así, $\{p_i\}_{\min}$ mayoriza a todas las demás posibilidades y, teniendo en cuenta la Schur-concavidad de la entropía, esta distribución de probabilidad es entonces la de menor entropía. Alternativamente, esto se puede comprobar resolviendo directamente el problema de optimización usando multiplicadores de Lagrange, lo que se puede encontrar en [152]. En conclusión, la solución a nuestro problema de minimización es

$$H_{\min} = -MP \log P - (1 - MP) \log(1 - MP), \quad (4.6)$$

para cualesquiera valores de M y P tales que

$$M \leq \frac{1}{P} < M + 1 \quad (M \in \mathbb{N}), \quad (4.7)$$

o sea, $M = \lceil 1/P \rceil$.

Ahora, el resultado previo nos permite reducir el problema original al más sencillo de minimizar la función de dos variables

$$\begin{aligned} \mathcal{H}(P_A, P_B) &\equiv H_{\min}(A) + H_{\min}(B) \\ &= \sum_{i=A,B} [-M_i P_i \log P_i - (1 - M_i P_i) \log(1 - M_i P_i)] \end{aligned} \quad (4.8)$$

con las restricciones

$$M_i \leq \frac{1}{P_i} < M_i + 1 \quad (M_i \in \mathbb{N}, i = A, B) \quad (4.9)$$

$$\arccos \sqrt{P_A} + \arccos \sqrt{P_B} \geq \arccos c, \quad (4.10)$$

donde la primera viene de (4.7) y la segunda de la desigualdad de Landau-Pollak (2.42).

Por conveniencia, resolveremos el problema hallando los máximos de $-\mathcal{H}$, lo cual se puede llevar a cabo aplicando la teoría de (Karush)-Kuhn-Tucker para optimización sujeta a restricciones dadas por desigualdades (ver cualquier libro de texto estándar sobre optimización, por ejemplo [153]). Consideremos el problema de optimización de una función de n variables x_i ,

$$\text{Maximizar } f(\mathbf{x}) \quad \text{sujeto a } g_i(\mathbf{x}) \leq c_i \quad (i = 1, \dots, m). \quad (4.11)$$

Introduciendo m multiplicadores de Lagrange $\{\lambda_i\}$ se obtiene el lagrangiano

$$\mathcal{L} = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i [c_i - g_i(\mathbf{x})]. \quad (4.12)$$

Las condiciones de Kuhn-Tucker, necesarias (pero no suficientes) para que un punto sea solución del problema (4.11), son entonces

$$\frac{\partial \mathcal{L}}{\partial x_i} = 0, \quad i = 1, \dots, n \quad (4.13)$$

$$\lambda_i [c_i - g_i(\mathbf{x})] = 0, \quad i = 1, \dots, m \quad (4.14)$$

$$g_i(\mathbf{x}) \leq c_i, \quad i = 1, \dots, m \quad (4.15)$$

$$\lambda_i \geq 0, \quad i = 1, \dots, m. \quad (4.16)$$

Para resolver nuestro problema excluimos en un principio la posibilidad de que $P_i = 1/M_i$ para al menos una i , la cual trataremos separadamente. El lagrangiano en nuestro caso es entonces

$$\begin{aligned} \mathcal{L} = \sum_{i=A,B} \left[M_i P_i \log P_i + (1 - M_i P_i) \log(1 - M_i P_i) + \mu_i \left(\frac{1}{M_i} - P_i \right) \right. \\ \left. + \nu_i \left(P_i - \frac{1}{M_i + 1} \right) \right] + \lambda (\arccos \sqrt{P_A} + \arccos \sqrt{P_B} - \arccos c), \end{aligned} \quad (4.17)$$

donde $\lambda, \mu_i, \nu_i \geq 0$ son los multiplicadores de Lagrange. Las condiciones de Kuhn-Tucker toman entonces la forma ($i = A, B$)

$$M_i \log \frac{P_i}{1 - M_i P_i} - \frac{\lambda}{2\sqrt{P_i(1 - P_i)}} - \mu_i + \nu_i = 0, \quad (4.18)$$

$$\mu_i \left(\frac{1}{M_i} - P_i \right) = 0, \quad (4.19)$$

$$\nu_i \left(P_i - \frac{1}{M_i + 1} \right) = 0, \quad (4.20)$$

$$\lambda (\arccos \sqrt{P_A} + \arccos \sqrt{P_B} - \arccos c) = 0, \quad (4.21)$$

debiéndose satisfacer, además, (4.9) y (4.10). Recuérdese que las condiciones de Kuhn-Tucker son necesarias para que un punto sea un máximo pero no suficientes, por lo que una vez encontradas todas las soluciones de (4.9)-(4.10) y (4.18)-(4.21) se debe comprobar cuál de ellas se corresponde con el verdadero máximo. Como nos restringimos de momento al caso en el que $P_i \neq 1/M_i$ para $i = A, B$, se debe fijar $\mu_i = \nu_i = 0$ para $i = A, B$ para que (4.19) y (4.20) sean compatibles con (4.9). Si $\lambda = 0$ también, entonces la condición (4.18) se reduciría a $P_i = 1/(1 + M_i)$, lo que llevaría a contradicción con (4.9). En consecuencia, $\lambda \neq 0$, por lo que (4.21) se reduce a

$$\arccos \sqrt{P_A} + \arccos \sqrt{P_B} = \arccos c. \quad (4.22)$$

Esto significa que, como era razonable esperar, la distribución de probabilidad óptima para la EUR satura la relación de incertidumbre de Landau-Pollak. Utilizando la identidad trigonométrica

$$\arccos x + \arccos y = \arccos \left(xy - \sqrt{(1-x^2)(1-y^2)} \right), \quad (4.23)$$

la ecuación (4.22) implica que

$$\begin{aligned} c &= \sqrt{P_A P_B} - \sqrt{(1-P_A)(1-P_B)} \\ &\leq \frac{1 - \sqrt{(M_A - 1)(M_B - 1)}}{\sqrt{M_A M_B}}, \end{aligned} \quad (4.24)$$

donde la desigualdad se sigue de (4.9). Puesto que $c \in (0, 1]$, se deduce de (4.24) que

$$\min(M_A, M_B) = 1, \quad c \leq \frac{1}{\sqrt{\max(M_A, M_B)}}. \quad (4.25)$$

Por otro lado, (4.18) lleva a

$$\begin{aligned} \lambda &= 2M_A \sqrt{P_A(1-P_A)} \log \frac{P_A}{1 - M_A P_A} \\ &= 2M_B \sqrt{P_B(1-P_B)} \log \frac{P_B}{1 - M_B P_B}. \end{aligned} \quad (4.26)$$

La ecuación (4.26) tiene diversas soluciones, dando cada una de ellas un potencial mínimo para \mathcal{H} . Una posibilidad es $P_A = P_B$ ($M_A = M_B$), entonces (4.22) significa que

$$P_A = P_B = \frac{1+c}{2}, \quad (4.27)$$

mientras que la primera ecuación en (4.25) impone que $M_A = M_B = 1$. Por tanto, esta solución da el siguiente candidato para mínimo de \mathcal{H} ,

$$\mathcal{F}(c) \equiv -(1+c) \log \frac{1+c}{2} - (1-c) \log \frac{1-c}{2}. \quad (4.28)$$

Si asumimos ahora que $P_A \neq P_B$, se tiene entonces que $M_A = 1$, $M_B = M \in \mathbb{N}$ debido a (4.25). De acuerdo a esto, los posibles mínimos de \mathcal{H} vendrán en este caso de las soluciones de la ecuación

$$\sqrt{P_A(1-P_A)} \log \frac{P_A}{1-P_A} = M \sqrt{P_B(1-P_B)} \log \frac{P_B}{1-MP_B} \quad (4.29)$$

para $M = 1, 2, \dots$, siendo cada una de ellas válida solamente en el rango $c \leq 1/\sqrt{M}$ como consecuencia de la segunda ecuación en (4.25). Desafortunadamente, la ecuación (4.29) con $P_A \neq P_B$ no puede resolverse analíticamente y su solución sólo puede obtenerse numéricamente. Para ello, resulta más conveniente, recordando (4.22), definir

$$P_A \equiv \cos^2 \alpha, \quad P_B \equiv \cos^2(\theta - \alpha), \quad c \equiv \cos \theta, \quad (4.30)$$

y, así, (4.29) se reescribe como

$$\begin{aligned} & \sin 2\alpha \log \left(\frac{1 + \cos 2\alpha}{1 - \cos 2\alpha} \right) + M \sin 2(\alpha - \theta) \\ & \times \log \left(\frac{1 + \cos 2(\alpha - \theta)}{2(1 - M \cos^2(\alpha - \theta))} \right) = 0, \end{aligned} \quad (4.31)$$

donde $\alpha \neq \theta/2, \theta/2 + \pi/4$ para que se cumpla $P_A \neq P_B$. Denotaremos por $\mathcal{H}_M(c)$ el posible mínimo de \mathcal{H} que se obtiene insertando en (4.8) los valores numéricos de $P_i(c)$ correspondientes a través de (4.30) a la solución $\alpha(\theta)$ de (4.31).

Por último, resta considerar la situación cuando se permite que $P_i = 1/M_i$ para $i = A$ y/o $i = B$. Si esta igualdad se tiene únicamente para una i , digamos $i = A$, entonces $\mu_B = \nu_B = 0$ y se obtienen las ecuaciones (4.22)-(4.25) igual que anteriormente. El análogo de (4.18) para $i = A$,

$$M_A(1 - \log M_A) - \frac{\lambda}{2\sqrt{P_A(1-P_A)}} - \mu_A = 0, \quad (4.32)$$

impone ahora que $M_A = 1, 2$ para que se cumpla la no negatividad de los multiplicadores de Lagrange. Si $M_A = 1$, entonces $P_A = 1$ y $P_B = c^2$ por (4.22), solución que lleva al mínimo potencial

$$\mathcal{G}(c) \equiv -c^2[1/c^2] \log c^2 - (1 - c^2[1/c^2]) \log(1 - c^2[1/c^2]), \quad (4.33)$$

válido para todos los posibles valores de c . Por otro lado, la solución correspondiente a $M_A = 2$ (que, como veremos, será irrelevante) sólo tiene sentido para $c \leq 1/\sqrt{2}$. Si $P_i = 1/M_i$ para $i = A$ e $i = B$, un razonamiento similar lleva a que $M_i = 1, 2$ para $i = A, B$, y las cotas resultantes, $\log 2$ y $\log 4$, son claramente no óptimas.

Mediante el estudio previo hemos obtenido todas las soluciones a las condiciones de Kuhn-Tucker para nuestro problema de optimización. Ahora, debemos seleccionar entre todas estas soluciones el auténtico mínimo de \mathcal{H} , lo que constituirá la nueva cota para $H(A) + H(B)$. Para $c \geq 1/\sqrt{2}$ sólo hay tres

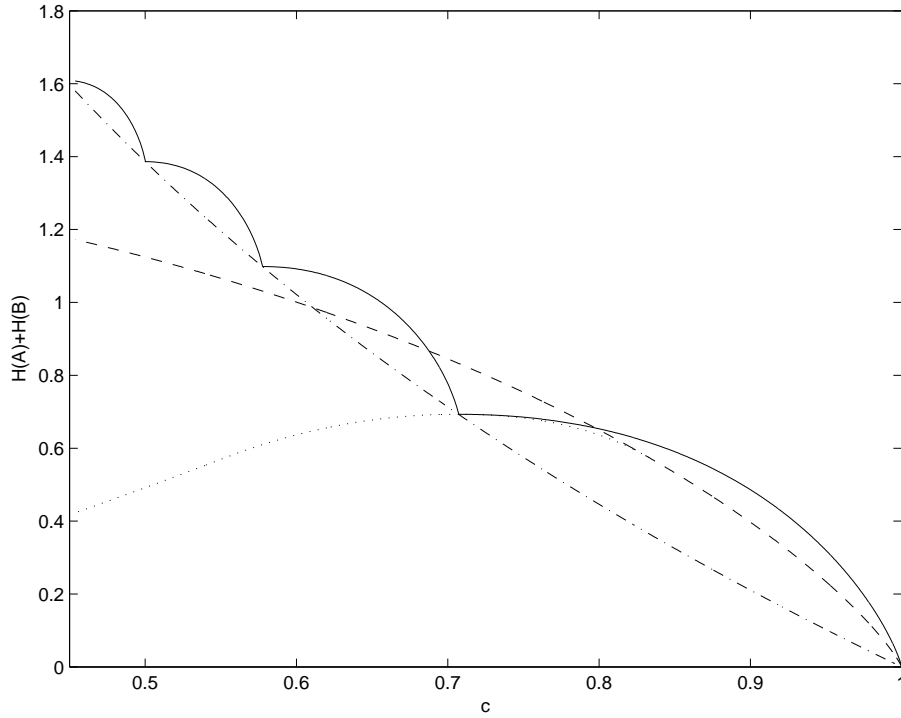


Figura 4.1: Posibles cotas para la suma de entropías de un par de observables con solapamiento c : cota de Maassen-Uffink (línea rayada-punteada), $\mathcal{F}(c)$ (línea rayada), $\mathcal{H}_1(c)$ (línea punteada), y $\mathcal{G}(c)$ (línea continua).

posibilidades, las cotas analíticas $\mathcal{F}(c)$, $\mathcal{G}(c)$ y la cota numérica $\mathcal{H}_1(c)$. Éstas se representan en la figura 4.1 junto con la cota de Maassen-Uffink (2.39). De ahí, se comprueba que el mínimo de \mathcal{H} es $\mathcal{F}(c)$ cuando $c \geq c^* \simeq 0,834$, y $\mathcal{H}_1(c)$ cuando $1/\sqrt{2} \leq c \leq c^*$. Remarcablemente, en ambos casos, se obtiene una cota más restrictiva que la dada por la EUR de Maassen-Uffink. Nótese que $\mathcal{G}(c)$ no se corresponde con el mínimo para ningún valor de c , aunque toma valores muy cercanos a los de la cota numérica $\mathcal{H}_1(c)$ y prácticamente se solapa con ella en un rango considerable.

Por otro lado, cuando $c \leq 1/\sqrt{2}$ el mínimo de \mathcal{H} es inferior a la cota de Maassen-Uffink, como se comprueba al considerar la forma de $\mathcal{H}_1(c)$ para estos valores en la figura 4.1. No obstante, puesto que $\mathcal{G}(c)$ interpola a la cota de Maassen-Uffink en los únicos puntos en los que ésta es óptima, i.e. $c = 1/\sqrt{n}$ con $n \in \mathbb{N}$, uno podría estar tentado a pensar que $\mathcal{G}(c)$ es en realidad la cota inferior óptima de $H(A) + H(B)$ para $c \leq 1/\sqrt{2}$. Sin embargo, es posible encontrar ejemplos con una entropía ligeramente inferior, lo que demuestra la falsedad de esta conjetura.

En resumen, hemos probado el siguiente teorema, que constituye una mejora

de la cota de Maassen-Uffink para EURs para pares de observables que satisfacen la condición de solapamiento grande (4.1).

Teorema 4.1. *Sea un sistema cuántico con espacio de Hilbert asociado $H \simeq \mathbb{C}^N$ y caracterizado por el operador densidad ρ y sean A y B observables no degenerados cualesquiera de este sistema ($A, B \in \mathcal{B}(H)$). Entonces,*

$$H_\rho(A) + H_\rho(B) \geq \begin{cases} -2 \log c & \text{si } 0 < c \leq 1/\sqrt{2}, \\ \mathcal{H}_1(c) & \text{si } 1/\sqrt{2} \leq c \leq c^*, \\ \mathcal{F}(c) & \text{si } c^* \leq c \leq 1, \end{cases} \quad (4.34)$$

donde $c^* \simeq 0,834$, $\mathcal{F}(c)$ viene dado por (4.28), y la cota numérica $\mathcal{H}_1(c)$ se define de acuerdo a (4.31).

Debe resaltarse que para observables de un espacio de Hilbert bidimensional, para los cuales la condición de solapamiento grande (4.1) siempre se cumple, la cota de (4.34) coincide con la cota óptima [46]. Esto significa que en el caso general (dimensiones arbitrarias) la cota de (4.34) es la mejor posible expresada exclusivamente en términos del solapamiento c cuando $c \geq 1/\sqrt{2}$. Como consecuencia adicional, puesto que, debido a (4.22), los estados que alcanzan $\mathcal{F}(c)$ y $\mathcal{H}_1(c)$ también alcanzan la cota de Landau-Pollak, nuestros resultados muestran que la relación de incertidumbre de Landau-Pollak es óptima para sistemas cuánticos bidimensionales. Además, de nuestro desarrollo se sigue también que la relación de incertidumbre de Landau-Pollak es más fuerte que la de Maassen-Uffink para los observables que satisfacen (4.1), y más débil en caso contrario.

Parte III

Entrelazamiento

Capítulo 5

Condiciones de separabilidad a partir de la relación de incertidumbre de Landau-Pollak

Las relaciones de incertidumbre constituyen una herramienta útil en la caracterización del entrelazamiento, como hemos analizado en la sección 2.4.2 en el caso particular de relaciones basadas en la desviación estándar y en la entropía (EURs). Sin embargo, la relación de incertidumbre de Landau-Pollak (2.42) aún no ha sido estudiada en este contexto, y, además, como acabamos de ver en el capítulo 4, esta relación utiliza una medida de la incertidumbre matemáticamente muy sencilla y puede ser efectivamente más restrictiva que la correspondiente EUR (la de Maassen-Uffink). Merece por tanto la pena estudiar en detalle el papel que puede jugar esta relación de incertidumbre en el contexto del problema de la separabilidad, que es nuestro objetivo en este capítulo.

En primer lugar, debe subrayarse que para este propósito, dado que nos vemos obligados a trabajar con estados mezcla y dada la falta de convexidad del funcional $\arccos \sqrt{M_\infty}$, nos debemos restringir a considerar la relación de Landau-Pollak en su formulación débil (2.43). Como veremos más adelante, esto no implica necesariamente una debilitación de los resultados finales y, además, debido a la estrategia utilizada muchos cálculos se harán más simples. En segundo lugar, nótese que los trabajos anteriores explicados en 2.4.2 ya permiten prácticamente sin necesidad de trabajo extra establecer condiciones de separabilidad a partir de la relación de Landau-Pollak. Efectivamente, si consideramos el lema de Gühne y Lewenstein [119] (Lema 2.1 en esta memoria) conjuntamente con el Teorema 1.3 se concluye directamente que si $\rho = \rho_A \otimes \rho_B$ es un estado producto de $\mathcal{S}(H_A \otimes H_B)$, y A y B son observables de $\mathcal{B}(H_A)$ y

$\mathcal{B}(H_B)$ con autovalores no nulos entonces

$$M_\infty(A \otimes B, \rho) \leq M_\infty(A, \rho_A), \quad M_\infty(A \otimes B, \rho) \leq M_\infty(B, \rho_B). \quad (5.1)$$

A partir de esta observación se puede obtener entonces una condición necesaria para la separabilidad que expresamos a continuación.

Corolario 5.1. *Si un estado es separable ($\rho \in \mathcal{D}(H_A \otimes H_B)$), entonces,*

$$M_\infty(A_1 \otimes B_1, \rho) + M_\infty(A_2 \otimes B_2, \rho) \leq 1 + \min\{c(A_1, A_2), c(B_1, B_2)\}, \quad (5.2)$$

donde $A_1, A_2 \in \mathcal{B}(H_A)$ y $B_1, B_2 \in \mathcal{B}(H_B)$ son observables con autovalores no nulos.

Demostración. Puesto que ρ es separable se puede escribir como $\rho = \sum_i p_i \rho_i^A \otimes \rho_i^B$. Así,

$$\begin{aligned} & M_\infty(A_1 \otimes B_1, \rho) + M_\infty(A_2 \otimes B_2, \rho) \\ & \leq \sum_i p_i (M_\infty(A_1 \otimes B_1, \rho_i^A \otimes \rho_i^B) + M_\infty(A_2 \otimes B_2, \rho_i^A \otimes \rho_i^B)) \\ & \leq \sum_i p_i (M_\infty(A_1, \rho_i^A) + M_\infty(A_2, \rho_i^A)) \\ & \leq \sum_i p_i (1 + c(A_1, A_2)) = 1 + c(A_1, A_2), \end{aligned} \quad (5.3)$$

donde en la primera desigualdad utilizamos la convexidad de M_∞ , en la segunda (5.1) y en la tercera la relación de Landau-Pollak débil (2.43). \square

Análogamente a los casos anteriores, observamos que los estados separables heredan la relación de incertidumbre de Landau-Pollak que satisfacen localmente. Sin embargo, este resultado tiene algunas limitaciones. En primer lugar, contrariamente a las EURs, no se conocen generalizaciones no triviales de la relación de incertidumbre de Landau-Pollak a conjuntos de más de dos observables, por lo que el Corolario 5.1 no puede ser fortalecido considerando más de un par de observables por subsistema. Segundo, y más importante, para los observables más relevantes la cota de (5.2) puede ser mejorada. Así, siguiendo una estrategia directa encontramos que, aunque sacrifiquemos la elegancia del procedimiento anterior, podemos cosechar resultados más fuertes.

5.1. Condiciones de separabilidad para sistemas de dos qubits

Los observables más característicos de los sistemas bidimensionales son los que representan las matrices de Pauli (2.67). Como hemos visto, éstos son los utilizados en las condiciones de separabilidad obtenidas a partir de relaciones de incertidumbre basadas en la desviación estándar y en la entropía para sistemas

de dos qubits (2.66) y (2.74) porque, además, al ser complementarios ofrecen las condiciones más restrictivas. Así pues, consideremos los siguientes observables de $\mathcal{B}(\mathbb{C}^2 \otimes \mathbb{C}^2)$:

$$Z = \sigma_z^A \otimes \sigma_z^B, \quad X = \sigma_x^A \otimes \sigma_x^B. \quad (5.4)$$

Como Z y X conmutan, para este par de observables se tiene que $c = 1$, y el miembro derecho de (2.42) se anula y, por tanto, no existe ninguna restricción a los resultados de una medida conjunta de los mismos. La cota inferior trivial 0 de (2.42) se alcanza, por ejemplo, si el estado medido es un elemento de la base de Bell $\{|\phi^\pm\rangle, |\psi^\pm\rangle\}$ (véase (2.60)). Contrariamente, si Z y X actúan en un estado separable, la cota inferior nula es inalcanzable, lo que permite la posibilidad de derivar una condición de separabilidad. La idea para llevar esto a cabo reside en que si $M_\infty(Z, |\phi\rangle_A \langle\phi| \otimes |\varphi\rangle_B \langle\varphi|) + M_\infty(X, |\phi\rangle_A \langle\phi| \otimes |\varphi\rangle_B \langle\varphi|) \leq C$ para cualquier estado producto (i.e., $\forall |\phi\rangle_A \in \mathbb{C}^2, \forall |\varphi\rangle_B \in \mathbb{C}^2$), entonces la cota C será válida para cualquier estado separable debido a la convexidad de M_∞ y a que los estados separables son combinaciones convexas de estados producto. Puesto que un estado producto arbitrario de dos qubits es de la forma

$$|\psi\rangle = |\phi\rangle_A \otimes |\varphi\rangle_B = (\cos \alpha |0\rangle_A + e^{i\delta} \sin \alpha |1\rangle_A) \otimes (\cos \beta |0\rangle_B + e^{i\gamma} \sin \beta |1\rangle_B), \quad (5.5)$$

con $\alpha, \beta \in [0, \pi/2]$ y $\delta, \gamma \in [0, 2\pi)$, son pocos los parámetros libres y es factible determinar la cota C mediante una maximización directa¹.

Proposición 5.1. *Si ρ es un estado separable de dos qubits, entonces*

$$M_\infty(Z, \rho) + M_\infty(X, \rho) \leq \frac{3}{2}. \quad (5.6)$$

Demostración. Tanto Z como X tienen por autovalores $+1$ y -1 , y los correspondientes proyectores en los subespacios propios son

$$\begin{aligned} P_+(Z) &= |00\rangle\langle 00| + |11\rangle\langle 11|, \\ P_-(Z) &= |01\rangle\langle 01| + |10\rangle\langle 10|, \\ P_\pm(X) &= |\phi^\pm\rangle\langle\phi^\pm| + |\psi^\pm\rangle\langle\psi^\pm|. \end{aligned} \quad (5.7)$$

En consecuencia, las probabilidades de hallar el estado puro separable (5.5) en estos subespacios propios son, respectivamente,

$$\begin{aligned} p_+(Z) &= (\cos \alpha \cos \beta)^2 + (\sin \alpha \sin \beta)^2, \\ p_-(Z) &= 1 - (\cos \alpha \cos \beta)^2 - (\sin \alpha \sin \beta)^2, \\ p_\pm(X) &= \frac{1}{2}(1 \pm \cos \delta \cos \gamma \sin 2\alpha \sin 2\beta). \end{aligned} \quad (5.8)$$

¹Nótese que considerar el funcional de incertidumbre de la relación de Landau-Pollak débil (2.43) en vez del de (2.42) hace los cálculos mucho más sencillos y, además, hace inmediata la comparación con el Corolario 5.1.

Dado que $p_{\pm}(Z)$ no dependen ni de δ ni de γ , y que $\sin 2\alpha \sin 2\beta$ es siempre no negativo, el valor máximo de $M_{\infty}(Z) + M_{\infty}(X)$ se corresponde con el máximo de las funciones

$$f_{\pm}(\alpha, \beta) = p_{\pm}(Z) + \frac{1}{2}(1 + \sin 2\alpha \sin 2\beta), \quad (5.9)$$

que se da cuando $\alpha = \pm\beta$, obteniéndose el resultado deseado. \square

Sin embargo, dado que $c(\sigma_z, \sigma_x) = 1/\sqrt{2}$, el Corolario 5.1 simplemente exige a los estados separables

$$M_{\infty}(Z, \rho) + M_{\infty}(X, \rho) \leq 1 + \frac{1}{\sqrt{2}} \approx 1,71 \quad (5.10)$$

y así vemos que, efectivamente, la Proposición 5.1 mejora la cota del anterior resultado y resulta por tanto más fuerte para la detección de estados entrelazados. Pero además, esta estrategia de optimización también permite la introducción de un tercer observable, obviamente

$$Y = \sigma_y^A \otimes \sigma_y^B, \quad (5.11)$$

lo que permite obtener una condición más fuerte aún, completamente análoga a (2.66) y (2.74).

Proposición 5.2. *Si ρ es un estado separable de dos qubits, entonces*

$$M_{\infty}(X, \rho) + M_{\infty}(Y, \rho) + M_{\infty}(Z, \rho) \leq 2. \quad (5.12)$$

Demostración. El observable Y tiene los mismos autovalores que Z y X , con proyectores en los subespacios propios

$$P_{\pm}(Y) = |\phi^{\mp}\rangle\langle\phi^{\mp}| + |\psi^{\pm}\rangle\langle\psi^{\pm}|, \quad (5.13)$$

por lo que las correspondientes probabilidades para el estado puro separable (5.5) son

$$p_{\pm}(Y) = \frac{1}{2}(1 \pm \sin \delta \sin \gamma \sin 2\alpha \sin 2\beta). \quad (5.14)$$

Como $\sin 2\alpha \sin 2\beta$ es no negativo, y el máximo sobre δ y γ de las cuatro funciones de la forma $\pm(\sin \delta \sin \gamma \pm \cos \delta \cos \gamma) = \pm \cos(\delta \mp \gamma)$ es 1, nuestro problema se reduce a hallar el máximo de las funciones

$$g_{\pm}(\alpha, \beta) = p_{\pm}(Z) + 1 + \frac{\sin 2\alpha \sin 2\beta}{2} = f_{\pm}(\alpha, \beta) + \frac{1}{2}, \quad (5.15)$$

donde f_{\pm} viene dada por (5.9), lo que ya resolvimos en la anterior demostración. \square

Teniendo en cuenta que $M_\infty(Y, \rho_{sep}) \geq 1/2$, se aprecia que la Proposición 5.1 puede ser derivada a partir de la Proposición 5.2 por lo cual ésta última es, como cabía esperar, una condición más fuerte para la detección de el entrelazamiento que la anterior. El interés de considerar (5.6), además de para comparar con el Corolario 5.1, estriba en su mayor facilidad de cara a la implementación experimental ya que requiere menos medidas.

Siguiendo [111, 118] también resulta interesante considerar las tres componentes ortogonales del spin total del sistema

$$S_i = \sigma_i^A \otimes \mathbb{I}_B + \mathbb{I}_A \otimes \sigma_i^B \quad (i = x, y, z). \quad (5.16)$$

Proposición 5.3. *Si ρ es un estado separable de dos qubits, entonces*

$$M_\infty(S_x, \rho_{sep}) + M_\infty(S_y, \rho_{sep}) + M_\infty(S_z, \rho_{sep}) \leq 2. \quad (5.17)$$

Demostración. Todos estos observables tienen por autovalores ± 2 (no degenerado) y 0 (de multiplicidad doble), con proyectores en los subespacios propios

$$\begin{aligned} P_\pm(S_x) &= \frac{1}{2}(|\phi^+ \rangle \pm |\psi^+ \rangle)(\langle \phi^+ | \pm \langle \psi^+ |), \\ P_0(S_x) &= |\phi^- \rangle \langle \phi^- | + |\psi^- \rangle \langle \psi^- |, \\ P_\pm(S_y) &= \frac{1}{2}(|\phi^- \rangle \pm |\psi^+ \rangle)(\langle \phi^- | \pm \langle \psi^+ |), \\ P_0(S_y) &= |\phi^+ \rangle \langle \phi^+ | + |\psi^- \rangle \langle \psi^- |, \\ P_+(S_z) &= |00\rangle \langle 00|, \quad P_-(S_z) = |11\rangle \langle 11|, \\ P_0(S_z) &= |01\rangle \langle 01| + |10\rangle \langle 10|, \end{aligned} \quad (5.18)$$

así que las correspondientes probabilidades para el estado producto genérico (5.5) son

$$\begin{aligned} p_\pm(S_x) &= \frac{1}{4}(1 \pm \cos \delta \sin 2\alpha)(1 \pm \cos \gamma \sin 2\beta), \\ p_0(S_x) &= \frac{1}{2}(1 - \cos \delta \cos \gamma \sin 2\alpha \sin 2\beta), \\ p_\pm(S_y) &= \frac{1}{4}(1 \pm \sin \delta \sin 2\alpha)(1 \pm \sin \gamma \sin 2\beta), \\ p_0(S_y) &= \frac{1}{2}(1 - \sin \delta \sin \gamma \sin 2\alpha \sin 2\beta), \\ p_+(S_z) &= (\cos \alpha \cos \beta)^2, \quad p_-(S_z) = (\sin \alpha \sin \beta)^2, \\ p_0(S_z) &= (\cos \alpha \sin \beta)^2 + (\sin \alpha \cos \beta)^2. \end{aligned} \quad (5.19)$$

De aquí se sigue que el valor máximo de $\sum_i M_\infty(S_i)$ para los estados producto es el máximo de la función

$$w(\alpha, \beta) = p(S_z) + 1 + \frac{\sin 2\alpha \sin 2\beta}{2}, \quad (5.20)$$

que el lector fácilmente puede verificar que resulta ser igual a 2. \square

Otra interesante posibilidad es considerar la medida de un observable no degenerado diagonal en la base de Bell,

$$B = \lambda_1|\phi^+\rangle\langle\phi^+| + \lambda_2|\phi^-\rangle\langle\phi^-| + \lambda_3|\psi^+\rangle\langle\psi^+| + \lambda_4|\psi^-\rangle\langle\psi^-|, \quad (5.21)$$

con $\lambda_i \neq \lambda_j$ si $i \neq j$.

Proposición 5.4. *Si ρ es un estado separable de dos qubits, entonces*

$$M_\infty(B) \leq \frac{1}{2}. \quad (5.22)$$

Demostración. La distribución de probabilidad para los resultados de la medida de B en el estado puro separable (5.5) es

$$\begin{aligned} p_{\phi^\pm}(B) &= \frac{1}{2}[(\cos \alpha \cos \beta)^2 + (\sin \alpha \sin \beta)^2 \pm \xi(\alpha, \beta)\zeta_+(\delta, \gamma)], \\ p_{\psi^\pm}(B) &= \frac{1}{2}[(\cos \alpha \sin \beta)^2 + (\sin \alpha \cos \beta)^2 \pm \xi(\alpha, \beta)\zeta_-(\delta, \gamma)], \end{aligned} \quad (5.23)$$

donde $\xi(\alpha, \beta) = \frac{1}{2} \sin 2\alpha \sin 2\beta$ y $\zeta_\pm(\delta, \gamma) = \cos(\delta \pm \gamma)$. La no negatividad de $\xi(\alpha, \beta)$ lleva a que $M_\infty(B)$ es el máximo sobre α y β de las funciones

$$\begin{aligned} h_1(\alpha, \beta) &= \frac{1}{2}[(\cos \alpha \cos \beta)^2 + (\sin \alpha \sin \beta)^2 + \xi(\alpha, \beta)], \\ h_2(\alpha, \beta) &= \frac{1}{2}[(\cos \alpha \sin \beta)^2 + (\sin \alpha \cos \beta)^2 + \xi(\alpha, \beta)], \end{aligned} \quad (5.24)$$

de lo que se sigue el resultado. \square

Debe aclararse que esta condición no es nueva ya que fue obtenido por otros medios en [119]. Como señalan estos autores, es equivalente al conjunto de testigos de entrelazamiento óptimos

$$W_{\phi^\pm} = \frac{1}{2}\mathbb{I} - |\phi^\pm\rangle\langle\phi^\pm|, \quad W_{\psi^\pm} = \frac{1}{2}\mathbb{I} - |\psi^\pm\rangle\langle\psi^\pm|. \quad (5.25)$$

5.1.1. Eficacia de las condiciones de separabilidad

En lo que sigue aplicaremos las condiciones de separabilidad para dos qubits derivadas anteriormente a algunos estados para poner a prueba su capacidad para la identificación de estados entrelazados.

Ejemplo 1: Consideremos los estados de Werner, que en el caso de dos qubits son mezclas de un estado puro maximalmente entrelazado y el estado de máxima ignorancia o completamente mezcla²,

$$\rho_W = \frac{1-p}{4}\mathbb{I}_A \otimes \mathbb{I}_B + p|\psi^-\rangle\langle\psi^-|, \quad (5.26)$$

²Este estado podría modelizar la preparación en el laboratorio de un estado maximalmente entrelazado, el cual se ve contaminado por ruido blanco (es decir, ruido completamente aleatorio)

donde $p \in [0, 1]$. Estos estados son separables si y sólo si $p \leq 1/3$ (lo cual se demostró originalmente en [81] y también se puede comprobar con el criterio PPT [93]). Las probabilidades de encontrar ρ_W en cada subespacio propio al medir los observables anteriores son

$$\begin{aligned}
p_{\pm}(X) &= p_{\pm}(Y) = p_{\pm}(Z) = \frac{1 \mp p}{2}, \\
p_0(S_i) &= \frac{1 + p}{2}, \\
p_{\pm}(S_i) &= p_{\phi^{\pm}}(B) = p_{\psi^+}(B) = \frac{1 - p}{4}, \\
p_{\psi^-}(B) &= \frac{1 + 3p}{4}.
\end{aligned} \tag{5.27}$$

Así, se tiene que

$$\begin{aligned}
\sum_{\tau=X,Z} M_{\infty}(\tau, \rho_W) &= 1 + p, \\
\sum_{\tau=X,Y,Z} M_{\infty}(\tau, \rho_W) &= \frac{3(1+p)}{2}, \\
\sum_{i=x,y,z} M_{\infty}(S_i, \rho_W) &= \frac{3(1+p)}{2}, \\
M_{\infty}(B, \rho_W) &= \frac{1+3p}{4}.
\end{aligned} \tag{5.28}$$

A partir de estos resultados se observa que (5.6) detecta entrelazamiento cuando $p > 1/2$, mientras que (5.12), (5.17), y (5.22) lo reconocen cuando $p > 1/3$. Nótese, entonces, que las tres últimas condiciones de separabilidad son óptimas para la detección del entrelazamiento en este conjunto de estados. La condición (2.66) basada en desviaciones estándar tiene también una capacidad óptima en este caso [111]. Sin embargo, las condiciones basadas en entropías de Shannon tienen un rendimiento sensiblemente peor ya que sólo reconocen con certeza a ρ_W como entrelazado en el mejor de los casos cuando $p > 0,55$ (utilizando los observables $\{S_i\}$, véase [118]). Lo mismo ocurre utilizando entropías de Tsallis que en el mejor de los casos estudiados [119] detectan el entrelazamiento para $p > 1/\sqrt{3}$. Lo que es más, cuando se miden los mismos observables (i.e. cuando se utiliza el mismo montaje experimental), nuestras condiciones siempre mejoran los resultados dados por las condiciones basadas en entropías de Shannon y Tsallis [118, 119], respectivamente: $p > 0,78$ y $p > 1/\sqrt{2}$ al medir X y Z ; $p > 0,65$ y $p > 1/\sqrt{3}$ al medir X , Y , y Z ; $p > 0,55$ al medir S_x , S_y , y S_z ; y $p > 0,74$ al medir B (en los dos últimos casos sólo se disponen de resultados para la entropía

de Shannon).

Ejemplo 2: Consideremos ahora los estados³

$$\rho_0 = p|\psi^-\rangle\langle\psi^-| + (1-p)|00\rangle\langle 00|, \quad (5.29)$$

con $p \in [0, 1]$. El criterio PPT muestra que la separabilidad sólo se da cuando $p = 0$ [93]. Las probabilidades para los anteriores observables son ahora

$$\begin{aligned} p_{\pm}(X) &= p_{\pm}(Y) = \frac{1 \mp p}{2}, \\ p_0(S_x) &= p_0(S_y) = \frac{1+p}{2}, \\ p_{\pm}(S_x) &= p_{\pm}(S_y) = \frac{1-p}{4}, \\ p_+(Z) &= p_+(S_z) = 1-p, \\ p_-(Z) &= p_0(S_z) = p_{\psi^-}(B) = p, \\ p_-(S_z) &= p_{\psi^+}(B) = 0, \\ p_{\phi^{\pm}}(B) &= \frac{1-p}{2}, \end{aligned} \quad (5.30)$$

y, consecuentemente,

$$\begin{aligned} \sum_{\tau=X,Z} M_{\infty}(\tau, \rho_0) &= \max\{p, 1-p\} + \frac{1+p}{2}, \\ \sum_{\tau=X,Y,Z} M_{\infty}(\tau, \rho_0) &= \max\{p, 1-p\} + 1+p, \\ \sum_{i=x,y,z} M_{\infty}(S_i, \rho_0) &= \max\{p, 1-p\} + 1+p, \\ M_{\infty}(B) &= \max\left\{p, \frac{1-p}{2}\right\}. \end{aligned} \quad (5.31)$$

Así pues, se tiene que (5.6) identifica a los estados como entrelazados cuando $p > 2/3$, mientras que basta con tener $p > 1/2$ para que se pueda deducir la presencia de entrelazamiento utilizando las condiciones (5.12), (5.17), y (5.22). Por tanto, para este conjunto de estados las condiciones anteriormente derivadas no son capaces de detectar todos los estados entrelazados. Sin embargo, como en el caso de los estados de Werner, para cualquier esquema de medidas estos resultados son más potentes que sus análogos entrópicos [118]: $p > 0,85$ al medir X y Z ; $p > 0,73$ al medir X , Y , y Z ; $p > 0,55$ al medir S_x , S_y , y S_z ; y $p > 0,78$ al medir B . Así mismo, debe señalarse, que al igual que en el Ejemplo 1, la condición basada en la desviación estándar (2.66) tiene la misma capacidad

³Estos estados también podrían modelizar la preparación experimental del estado ψ^- , que ahora se ve afectado por un ruido “coloreado”.

para la detección del entrelazamiento ($p > 1/2$).

Estos ejemplos ponen de manifiesto situaciones que analizamos a continuación en más detalle. En primer lugar, como se ve en el Ejemplo 2 y como cabía esperar por analogía con las condiciones basadas en la desviación estándar y en las entropías, las condiciones basadas en M_∞ son necesarias para la separabilidad pero no suficientes. En segundo lugar, se observa que las condiciones basadas en la relación de incertidumbre de Landau-Pollak son más fuertes que las entrópicas. Esta tónica se mantiene al considerar más ejemplos, lo que permite conjeturar que las proposiciones 5.1-5.4 son más fuertes que sus análogos entrópicos. Uno podría pensar que esto se debe a que mientras que, por ejemplo, la condición de separabilidad entrópica (2.74) es consecuencia directa del Lema 2.1, en nuestro caso hemos mejorado el resultado que se sigue del Lema 2.1 mediante una estrategia directa de optimización. Sin embargo, resulta que esta estrategia en el caso entrópico no mejora las condiciones sino que reproduce la misma cota [118]. Así pues, la aparente mayor fortaleza de estas condiciones frente a las entrópicas no es accidental debido a la metodología empleada en su derivación sino que responde a la estructura inherente a cada medida ⁴. Además, la estructura matemática de la medida M_∞ es mucho más sencilla que la entrópica por lo que las condiciones basadas en la primera son mucho más fáciles de manejar. De hecho, aplicar las condiciones entrópicas a conjuntos de estados que dependen de parámetros libres es a veces inaccesible analíticamente. Otro hecho que llama la atención en los ejemplos considerados es que las condiciones (5.12), (5.17) y (5.22) arrojan los mismos resultados. No obstante, otros ejemplos muestran que (5.17) es a veces más débil que las otras dos, que como mostramos a continuación sí resultan ser equivalentes.

Proposición 5.5. *La Proposición 5.2 y la Proposición 5.4 son equivalentes.*

Demostración. La condición (5.22) se puede expresar como

$$0 \leq \text{Tr}(|BS_i\rangle\langle BS_i|\rho) \leq \frac{1}{2}, \quad (5.32)$$

donde $|BS_i\rangle$ es cualquier elemento de la base de Bell (2.60). Teniendo en cuenta las identidades [119]

$$\begin{aligned} \text{Tr}(|\phi^\pm\rangle\langle\phi^\pm|\rho) &= \frac{1 \pm \text{Tr}(X\rho) \mp \text{Tr}(Y\rho) + \text{Tr}(Z\rho)}{4}, \\ \text{Tr}(|\psi^\pm\rangle\langle\psi^\pm|\rho) &= \frac{1 \pm \text{Tr}(X\rho) \pm \text{Tr}(Y\rho) - \text{Tr}(Z\rho)}{4}, \end{aligned} \quad (5.33)$$

y observando que

$$\text{Tr}(\tau\rho) = p_+(\tau) - p_-(\tau) \equiv \Delta p(\tau) \quad (\tau = X, Y, Z), \quad (5.34)$$

⁴Nótese, así mismo, que nuestra mejora de una EUR para dos observables del capítulo anterior no conduce a una mejora relevante de las condiciones de separabilidad entrópicas puesto que los observables que dan las cotas más restrictivas son complementarios ($c = 1/\sqrt{2}$).

las desigualdades en (5.32) se pueden escribir como

$$\begin{aligned} -1 &\leq \pm\Delta p(X, \rho) \mp \Delta p(Y, \rho) + \Delta p(Z, \rho) \leq 1, \\ -1 &\leq \pm\Delta p(X, \rho) \pm \Delta p(Y, \rho) - \Delta p(Z, \rho) \leq 1. \end{aligned} \quad (5.35)$$

Esto equivale al conjunto de ocho desigualdades de la forma

$$-1 \leq \pm\Delta p(X, \rho) \pm \Delta p(Y, \rho) \pm \Delta p(Z, \rho) \leq 1, \quad (5.36)$$

o sea,

$$|\Delta p(X, \rho)| + |\Delta p(Y, \rho)| + |\Delta p(Z, \rho)| \leq 1. \quad (5.37)$$

Finalmente, puesto que para $\tau = X, Y, Z$

$$|\Delta p(\tau)| = M_\infty(\tau) - (1 - M_\infty(\tau)) = 2M_\infty(\tau) - 1, \quad (5.38)$$

la ecuación (5.37) se reduce a (5.12). \square

Esto significa que la condición de separabilidad (5.12) (que es entonces la más fuerte con un esquema de medida local) es equivalente al conjunto óptimo de testigos de entrelazamiento dado por (5.25). Esto nos lleva al último punto que queda por comentar: la relación entre esta condición y su análogo basado en la desviación estándar (2.66). En los ejemplos considerados hasta ahora las dos condiciones muestran la misma capacidad para la detección de estados entrelazados. Sin embargo, otros ejemplos muestran que (2.66) es más fuerte. Esto se debe a que al considerar la desviación estándar se tiene tanto un término correspondiente al valor esperado de un observable como un término correspondiente al valor esperado al cuadrado de un observable. Este segundo término se puede interpretar como una corrección de segundo orden al criterio dado por los testigos de entrelazamiento [154], lo que ha dado lugar a la teoría de los testigos de entrelazamiento no lineales [155]. Así pues, mientras que la condición (5.12) es equivalente al mejor conjunto de 4 testigos posible, la condición (2.66) ofrecería su refinamiento no lineal siendo, por consiguiente, más fuerte.

5.2. Condiciones de separabilidad para sistemas de mayor dimensionalidad

Al considerar sistemas bipartitos en dimensiones arbitrarias, el número de parámetros libres de un estado producto arbitrario crece enormemente con la dimensión por lo que el procedimiento de maximización utilizado en el caso de dos qubits se hace intratable analíticamente y muy complicado incluso desde el punto de vista numérico. No obstante, el Corolario 5.1 proporciona una forma de construir condiciones de separabilidad de una forma inmediata a través de una elección adecuada de observables. Para estados con subsistemas bidimensionales la mejor capacidad de detección de entrelazamiento se consigue mediante la medida en cada subsistema de las tres componentes ortogonales del spin, i.e.

las tres matrices de Pauli, que forman así mismo un conjunto maximal de observables complementarios. Sin embargo cuando la dimensión del espacio de Hilbert es estrictamente mayor que dos las componentes ortogonales del spin no son observables complementarios y ambos casos deben tratarse separadamente.

Así, si A_1, A_2 y/o B_1, B_2 son observables complementarios en un espacio de Hilbert D -dimensional, del Corolario 5.1 se sigue que

$$M_\infty(A_1 \otimes B_1, \rho) + M_\infty(A_2 \otimes B_2, \rho) \leq 1 + \frac{1}{\sqrt{D}} \quad (5.39)$$

es una condición necesaria para que ρ sea separable.

Por otro lado, si S_n y $S_{n'}$ denotan observables de spin D -dimensional a lo largo de los ejes n y n' , respectivamente, se tiene que [156]

$$c^2(S_n, S_{n'}) = \binom{D-1}{n^*} \left(\cos^2 \frac{\beta}{2} \right)^{D-1-n^*} \left(\sin^2 \frac{\beta}{2} \right)^{n^*}, \quad n^* = \left[D \sin^2 \frac{\beta}{2} \right], \quad (5.40)$$

donde β es el ángulo que forman n y n' . Por lo tanto, el uso del Corolario 5.1 lleva a

$$\begin{aligned} & M_\infty(S_n^A \otimes S_n^B, \rho) + M_\infty(S_{n'}^A \otimes S_{n'}^B, \rho) \\ & \leq 1 + \sqrt{\binom{D-1}{n^*} \left(\cos \frac{\beta}{2} \right)^{D-1-n^*} \left(\sin \frac{\beta}{2} \right)^{n^*}} \end{aligned} \quad (5.41)$$

como condición necesaria para la separabilidad. Eligiendo ejes ortogonales ($\beta = \pi/2$) la condición anterior se simplifica a

$$M_\infty(S_x^A \otimes S_x^B, \rho) + M_\infty(S_z^A \otimes S_z^B, \rho) \leq 1 + \sqrt{\frac{1}{2^{D-1}} \binom{D-1}{[D/2]}}. \quad (5.42)$$

Debe mencionarse que cuando D es impar los observables de spin tienen un autovalor de multiplicidad simple nulo, por lo que no se satisfacen las hipótesis del Corolario 5.1. Sin embargo, como ya se apunta en [119], el requerimiento de autovalores no nulos es una condición técnica de poca relevancia en la práctica ya que esta condición se puede satisfacer alterando los autovalores, lo cual no varía el proceso de medición ni el valor de c (ni, por tanto, las relaciones de incertidumbre), que sólo dependen de los autovectores.

Debe subrayarse que las condiciones de separabilidad obtenidas en esta sección, contrariamente al caso entrópico, no pueden mejorarse considerando medidas de observables adicionales porque no se conoce ninguna generalización no trivial de la relación de incertidumbre de Landau-Pollak para conjuntos de más de dos observables, lo que restringe de alguna forma su capacidad para la detección de entrelazamiento.

Capítulo 6

Entrelazamiento y medidas de la correlación: Criterios de separabilidad a partir de la representación de Bloch

En el capítulo anterior hemos ampliado el acercamiento al problema de la separabilidad mediante relaciones de incertidumbre considerando la relación de Landau-Pollak. Si bien se han obtenido resultados interesantes, se ha confirmado que el uso del criterio LUR con desviaciones estándar es el más potente. Por un lado, hemos visto que da las condiciones más fuertes en el caso de más asequible análisis que es el de dos qubits. Por otro lado el trabajo desarrollado en [113] (véase (2.68)-(2.69)) proporciona una forma extremadamente fuerte (capaz de detectar entrelazamiento PPT y más fuerte que el criterio CCNR) de aplicar este criterio en dimensiones arbitrarias donde las condiciones basadas en EURs siguen siendo débiles y nuestro acercamiento basado en M_∞ muestra los inconvenientes ya analizados. La única contrapartida del Teorema 2.8 (o más en particular de la condición (2.69)) reside en su falta de operacionalidad dado que su capacidad para la detección del entrelazamiento depende de la elección de observables que se haga en cada caso. Merece, por tanto, la pena estudiar en más detalle este criterio y profundizar en las razones de su fortaleza. Como ya hemos comentado este criterio se puede interpretar como un refinamiento de segundo orden para un determinado testigo de entrelazamiento, lo que ha llevado a la teoría de testigos de entrelazamiento no lineales. Aquí nos acercaremos con un enfoque distinto (siguiendo las ideas de [157]) con el objetivo de desarrollar algún criterio con una capacidad similar pero operacional (i.e., que no dependa de una elección de observables).

Consideremos entonces los observables $\{A_i\}$ y $\{B_i\}$ de $\mathcal{B}(H_A)$ y $\mathcal{B}(H_B)$ respectivamente tales que $\sum_i \Delta_\rho^2(A_i) \geq C_A$ y $\sum_i \Delta_\rho^2(B_i) \geq C_B$. El criterio de

LUR enunciado en el Teorema 2.8 nos dice entonces que los estados separables deben satisfacer la desigualdad (2.63) que reproducimos a continuación para comodidad del lector:

$$\sum_i \Delta_\rho^2(A_i \otimes I_B + I_A \otimes B_i) \geq C_A + C_B.$$

Un sencillo cálculo (utilizando simplemente la definición de desviación estándar (1.20)) nos muestra que el miembro de la izquierda se puede escribir como

$$\sum_i \Delta_\rho^2(A_i \otimes I_B + I_A \otimes B_i) = \sum_i \Delta_{\rho_A}^2(A_i) + \sum_i \Delta_{\rho_B}^2(B_i) + 2 \sum_i \kappa_\rho(A_i, B_i), \quad (6.1)$$

donde $\kappa_\rho(A_i, B_i)$ es la covarianza,

$$\kappa_\rho(A_i, B_i) = \langle A_i \otimes B_i \rangle_\rho - \langle A_i \rangle_{\rho_A} \langle B_i \rangle_{\rho_B}. \quad (6.2)$$

De (6.1) se observa entonces que para que se viole la condición de LUR (2.63) es preciso que alguno de los términos de covarianza sea negativo. Pero, es más, como $\kappa_\rho(-A_i, B_i) = -\kappa_\rho(A_i, B_i)$ y $\Delta_\rho^2(-A_i) = \Delta_\rho^2(A_i)$, tomando adecuadamente los signos siempre existe alguna elección para la cual todos los término de covarianza son negativos y, así, (6.1) se puede escribir como

$$\sum_i \Delta_\rho^2(A_i \otimes I_B + I_A \otimes B_i) = \sum_i \Delta_{\rho_A}^2(A_i) + \sum_i \Delta_{\rho_B}^2(B_i) - 2 \sum_i |\kappa_\rho(A_i, B_i)|, \quad (6.3)$$

como inicialmente se demostró en [157]. De aquí se sigue, por tanto, que la existencia de estados entrelazados que violan la condición (2.63) se debe a su capacidad para tener términos de covarianza "grandes" en valor absoluto, lo que por enésima vez nos lleva al entrelazamiento como manifestación de una diferencia de comportamiento a nivel global y local. Nótese que para los estados producto $\kappa_\rho(A_i, B_i) = 0 \forall i$, es decir, $\langle A_i \otimes B_i \rangle_\rho = \langle A_i \rangle_{\rho_A} \langle B_i \rangle_{\rho_B}$, puesto que que no hay correlación entre los subsistemas y siguen estadísticas independientes. En los estados separables mezcla alguno de los términos de covarianza pueden ser considerable, pero la suma de todos no demasiado grande ya que se tiene entonces la acotación

$$\sum_i |\kappa_\rho(A_i, B_i)| \leq \frac{1}{2} \left(\sum_i \Delta_{\rho_A}^2(A_i) - C_A + \sum_i \Delta_{\rho_B}^2(B_i) - C_B \right). \quad (6.4)$$

Sólo en los estados entrelazados son permitidos términos de covarianza arbitrariamente grandes fruto de la gran diferencia que puede haber entre la estadística a nivel global y a nivel local. En otras palabras, sólo los estados entrelazados pueden exhibir un alto nivel de correlación para varios observables de la forma $A \otimes B$ a la vez. Así, nuestro objetivo en este capítulo será hallar condiciones de separabilidad a través de alguna medida de la correlación que sea función de $\kappa_\rho(A_i, B_i)$ (o bien de $\langle A_i \otimes B_i \rangle_\rho$, $\langle A_i \rangle_{\rho_A}$ y $\langle B_i \rangle_{\rho_B}$ separadamente) para alguna elección adecuada de los $\{A_i\}$ y los $\{B_i\}$. Como veremos, la representación de Bloch, que introducimos a continuación, ofrece un marco muy apropiado para este estudio.

6.1. La representación de Bloch

Consideremos un sistema cuántico monocomponente de N niveles (es decir, con espacio de Hilbert asociado $H \simeq \mathbb{C}^N$), que vendrá, por tanto, caracterizado por matrices densidad $\rho \in \mathcal{S}(H)$. Puesto que este conjunto no tiene estructura de espacio vectorial, resulta complicado hallar una representación sencilla para las matrices que lo componen. Sin embargo, $\mathcal{S}(H) \subset \mathcal{B}(H)$, que sí es un espacio vectorial. De hecho, los operadores hermíticos que actúan en H , constituyen a su vez un espacio de Hilbert, también llamado espacio de Hilbert-Schmidt (con producto escalar (2.3)). Consecuentemente, las matrices densidad se pueden expandir en términos de cualquier base de $\mathcal{B}(H)$. Una elección particular muy conveniente es la formada por el operador identidad I_N y los generadores hermíticos de traza nula del grupo $SU(N)$ $\{\lambda_i\}$ ($i = 1, 2, \dots, N^2 - 1$),

$$\rho = \frac{1}{N}(I_N + \sum_i r_i \lambda_i). \quad (6.5)$$

Esta base es además ortogonal, ya que los generadores de $SU(N)$ satisfacen

$$\langle \lambda_i, \lambda_j \rangle_{HS} = \text{Tr}(\lambda_i \lambda_j) = 2\delta_{ij}, \quad (6.6)$$

y la ortogonalidad con la identidad se sigue de la nulidad de su traza, i.e.

$$\langle \lambda_i, I \rangle_{HS} = \text{Tr}(\lambda_i) = 0. \quad (6.7)$$

Los generadores de $SU(N)$ se caracterizan por las constantes de estructura del álgebra de Lie correspondiente, f_{ijk} y g_{ijk} , que son, respectivamente, completamente antisimétricas y completamente simétricas,

$$\lambda_i \lambda_j = \frac{2}{N} \delta_{ij} I_N + \sum_k i f_{ijk} \lambda_k + \sum_k g_{ijk} \lambda_k. \quad (6.8)$$

Los generadores pueden ser fácilmente construidos a partir de cualquier base ortonormal $\{|a\rangle\}_{a=0}^{N-1}$ de H [158]. Sean l, j, k índices tales que $0 \leq l \leq N-2$ y $0 \leq j < k \leq N-1$. Entonces, cuando $i = 1, \dots, N-1$

$$\lambda_i = w_l \equiv \sqrt{\frac{2}{(l+1)(l+2)}} \left(\sum_{a=0}^l |a\rangle \langle a| - (l+1)|l+1\rangle \langle l+1| \right), \quad (6.9)$$

mientras que cuando $i = N, \dots, (N+2)(N-1)/2$

$$\lambda_i = u_{jk} \equiv |j\rangle \langle k| + |k\rangle \langle j|, \quad (6.10)$$

y cuando $i = N(N+1)/2, \dots, N^2 - 1$

$$\lambda_i = v_{jk} \equiv -i(|j\rangle \langle k| - |k\rangle \langle j|). \quad (6.11)$$

Las relaciones de ortogonalidad (6.6)-(6.7) implican que los coeficientes de ρ en (6.5) están dados por

$$r_i = \frac{N}{2} \text{Tr}(\rho \lambda_i). \quad (6.12)$$

Nótese que la coordenada con respecto a I_N está fijada por la condición de traza unidad. El vector $\mathbf{r} = (r_1 r_2 \cdots r_{N^2-1})^t \in \mathbb{R}^{N^2-1}$, que caracteriza completamente al operador densidad, se denomina vector de Bloch o de coherencia. Esta representación (6.5) fue introducida por Bloch [159] en el caso bidimensional (donde los generadores son las matrices de Pauli (2.67)) y, posteriormente, fue generalizada a dimensiones arbitrarias en [158]. Además de su conveniencia matemática, esta representación también tiene especial atractivo desde el punto de vista experimental ya que muestra una forma de reconstruir el operador densidad a partir de los valores esperados de los operadores λ_i (que son observables puesto que pertenecen a $\mathcal{B}(H)$),

$$\langle \lambda_i \rangle = \text{Tr}(\rho \lambda_i) = \frac{2}{N} r_i. \quad (6.13)$$

Así pues, cualquier operador densidad admite una representación de la forma (6.5); sin embargo, no ocurre lo mismo a la inversa. Una matriz del tipo (6.5) es hermítica y de traza unidad, pero puede no ser semidefinida positiva, por lo que para garantizar esta propiedad se debe restringir en mayor medida la forma del vector de coherencia. El conjunto de todos los vectores de Bloch que dan lugar a un operador densidad se conoce como el espacio de los vectores de Bloch y lo denotaremos por $B(\mathbb{R}^{N^2-1})$. Es bien conocido que en el caso $N = 2$ este espacio es la bola unidad en \mathbb{R}^3 , correspondiéndose los vectores de la esfera unidad con los estados puros (véase, por ejemplo, [17]). La situación es más complicada cuando $N \geq 3$ y, de hecho, determinar geoméricamente $B(\mathbb{R}^{N^2-1})$ es un problema abierto en el que se investiga en la actualidad (véase [160] y la bibliografía allí citada). No obstante, muchas de las propiedades de este conjunto son conocidas. Por ejemplo, utilizando (6.8), es directo comprobar que los estados puros ($\rho^2 = \rho$) deben satisfacer

$$\|\mathbf{r}\|_2 = \sqrt{\frac{N(N-1)}{2}}, \quad \sum_{ij} r_i r_j g_{ijk} = (N-2)r_k, \quad (6.14)$$

donde $\|\cdot\|_2$ es la norma euclídea en \mathbb{R}^{N^2-1} . Así pues, como en el caso bidimensional los estados puros se caracterizan por vectores de Bloch que tocan una esfera, pero a diferencia de este no todos los vectores de la esfera son vectores de coherencia.

En el caso de estados mezcla, en [161, 162] se ha caracterizado recientemente mediante una lista de ecuaciones las condiciones que los vectores de $B(\mathbb{R}^{N^2-1})$ deben satisfacer. Lamentablemente, esto involucra expresiones matemáticas muy embrolladas, lo que restringe su utilidad. Un aspecto conocido del espacio de Bloch [163] que nos será muy útil más adelante es que $B(\mathbb{R}^{N^2-1})$ es un subconjunto de la bola $D_R(\mathbb{R}^{N^2-1})$ de radio $R = \sqrt{\frac{N(N-1)}{2}}$ (que es la bola más pequeña que lo contiene) y que la bola $D_r(\mathbb{R}^{N^2-1})$ de radio $r = \sqrt{\frac{N}{2(N-1)}}$ está incluida en $B(\mathbb{R}^{N^2-1})$. O sea,

$$D_r(\mathbb{R}^{N^2-1}) \subseteq B(\mathbb{R}^{N^2-1}) \subseteq D_R(\mathbb{R}^{N^2-1}). \quad (6.15)$$

Nótese que como $r = R = 1$ cuando $N = 2$, de aquí se deduce que $B(\mathbb{R}^3) = D_1(\mathbb{R}^3)$.

En el caso de sistemas cuánticos bipartitos de dimensiones $M \times N$ ($H \simeq \mathbb{C}^M \otimes \mathbb{C}^N$) compuestos por los subsistemas A y B , los operadores densidad análogamente se pueden representar como ¹

$$\rho = \frac{1}{MN} (I_M \otimes I_N + \sum_i r_i \lambda_i \otimes I_N + \sum_j s_j I_M \otimes \tilde{\lambda}_j + \sum_{ij} t_{ij} \lambda_i \otimes \tilde{\lambda}_j), \quad (6.16)$$

donde los $\{\lambda_i\}$ ($\{\tilde{\lambda}_j\}$) son los generadores de $SU(M)$ ($SU(N)$). Interesantemente, en esta representación podemos distinguir dos tipos de parámetros. Por un lado están $\mathbf{r} \in \mathbb{R}^{M^2-1}$ y $\mathbf{s} \in \mathbb{R}^{N^2-1}$, que son los vectores de coherencia de los subsistemas

$$\rho_A = \text{Tr}_B \rho = \frac{1}{M} (I_M + \sum_i r_i \lambda_i), \quad \rho_B = \text{Tr}_A \rho = \frac{1}{N} (I_N + \sum_i s_i \tilde{\lambda}_i) \quad (6.17)$$

y constituyen, por consiguiente, parámetros locales,

$$r_i = \frac{M}{2} \text{Tr}(\rho \lambda_i \otimes I_N) = \frac{M}{2} \text{Tr}_A(\rho_A \lambda_i) = \frac{M}{2} \langle \lambda_i \rangle_{\rho_A}, \quad (6.18)$$

y similarmente para \mathbf{s} . Los coeficientes t_{ij} , responsables de las posibles correlaciones entre los subsistemas forman la matriz real $T \in \mathbb{R}^{(M^2-1) \times (N^2-1)}$ que denominaremos matriz de correlación. Al igual que en el caso monocomponente, éstos también son fácilmente determinables en los experimentos puesto que son los valores esperados de observables medibles localmente

$$t_{ij} = \frac{MN}{4} \text{Tr}(\rho \lambda_i \otimes \tilde{\lambda}_j) = \frac{MN}{4} \langle \lambda_i \otimes \tilde{\lambda}_j \rangle_{\rho}. \quad (6.19)$$

6.2. Condiciones de separabilidad a partir de la representación de Bloch

La representación de Bloch ya ha sido utilizada previamente en el contexto del problema de la separabilidad en [99] para caracterizar el entrelazamiento de una clase particular de estados de dos qubits y en [166, 167] para estudiar la separabilidad en el entorno del estado maximalmente mezcla (i.e. la identidad). El acercamiento que presentamos aquí es diferente y más general. Empezamos mostrando que la representación de Bloch de los sistemas cuánticos bipartitos (6.16) nos permite encontrar una caracterización sencilla del entrelazamiento en estados puros.

¹Esta representación también es conocida como forma de Fano (ver por ejemplo [164]), ya que parece que este autor fue el primero en considerarla explícitamente en el caso bipartito [165].

Proposición 6.1. *Un estado bipartito puro con representación de Bloch (6.16) es separable si y sólo si*

$$T = \mathbf{r} \mathbf{s}^t. \quad (6.20)$$

Demostración. Nótese simplemente que (6.16) se puede reescribir como

$$\rho = \rho_A \otimes \rho_B + \frac{1}{MN} \left[\sum_{ij} (t_{ij} - r_i s_j) \lambda_i \otimes \tilde{\lambda}_j \right], \quad (6.21)$$

de forma que para los estados producto el segundo término del miembro de la derecha se tiene que anular. Ahora bien, como los $\{\lambda_i \otimes \tilde{\lambda}_j\}$ son linealmente independientes, $\sum_{ij} (t_{ij} - r_i s_j) \lambda_i \otimes \tilde{\lambda}_j = 0$ si y sólo si $t_{ij} - r_i s_j = 0 \forall i, j$. \square

Nótese que en el caso de estados mezcla, (6.20) es una condición suficiente para la separabilidad ya que entonces $\rho = \rho_A \otimes \rho_B$. Nótese igualmente que la Proposición 6.1 expresa el hecho de que dentro de los estados puros, los estados separables se caracterizan por tener términos de covarianza nulos $\langle \lambda_i \otimes \tilde{\lambda}_j \rangle_\rho = \langle \lambda_i \rangle_{\rho_A} \langle \tilde{\lambda}_j \rangle_{\rho_B}$ en una notación matricial que será muy útil en lo que sigue. Lo más interesante de la Proposición 6.1 es que, puesto que los estados separables en general son mezclas (i.e. combinaciones convexas) de estados producto, permite caracterizar la separabilidad en términos de los parámetros de la representación de Bloch:

Un estado cuántico bipartito con representación de Bloch (6.16) es separable si y sólo si existen vectores $\mathbf{u}_i \in \mathbb{R}^{M^2-1}$ y $\mathbf{v}_i \in \mathbb{R}^{N^2-1}$ que cumplan la ecuación (6.14) y pesos p_i tales que $0 \leq p_i \leq 1$, $\sum_i p_i = 1$ de forma que

$$T = \sum_i p_i \mathbf{u}_i \mathbf{v}_i^t, \quad \mathbf{r} = \sum_i p_i \mathbf{u}_i, \quad \mathbf{s} = \sum_i p_i \mathbf{v}_i. \quad (6.22)$$

Dejando aparte su interés formal, esta caracterización del entrelazamiento es poco útil en el sentido de que discernir la separabilidad o entrelazamiento de un estado arbitrario usando (6.22) o la definición original (2.61) es igual de difícil. Sin embargo, acercarse al problema de la separabilidad con este enfoque alternativo permite, como veremos a continuación, derivar nuevos y potentes criterios operacionales de separabilidad. En este cometido utilizaremos herramientas conocidas del análisis matricial. El lector menos familiarizado con este campo encontrará una breve descripción de ellas en el apéndice A.

Teorema 6.1. *Si un estado bipartito en $M \times N$ dimensiones con representación de Bloch (6.16) es separable, entonces se debe cumplir que*

$$\|T\|_{tr} \leq \sqrt{\frac{MN(M-1)(N-1)}{4}}. \quad (6.23)$$

Demostración. Debido a que T debe admitir una descomposición de la forma (6.22) con

$$\|\mathbf{u}_i\|_2 = \sqrt{\frac{M(M-1)}{2}}, \quad \|\mathbf{v}_i\|_2 = \sqrt{\frac{N(N-1)}{2}}, \quad (6.24)$$

utilizando la desigualdad triangular se ha de tener que

$$\|T\|_{tr} \leq \sum_i p_i \|\mathbf{u}_i \mathbf{v}_i^\dagger\|_{tr} = \sum_i p_i \sqrt{\frac{MN(M-1)(N-1)}{4}} \|\mathbf{n}_i \tilde{\mathbf{n}}_i^\dagger\|_{tr}, \quad (6.25)$$

donde $\mathbf{n}_i, \tilde{\mathbf{n}}_i$ son vectores unitarios. Así, $\|\mathbf{n}_i \tilde{\mathbf{n}}_i^\dagger\|_{tr} = 1 \forall i$ y se llega al resultado deseado. \square

Como analizaremos más adelante, este teorema (de demostración tan sencilla) constituye una potente condición para la detección del entrelazamiento. Además, tiene una clara interpretación física. Como mencionamos antes, T contiene la información acerca de las correlaciones de forma que en cierto sentido $\|T\|_{tr}$ mide la magnitud de estas correlaciones. Así pues, el Teorema 6.1 indica que existe una cota superior para las correlaciones contenidas en los estados separables, de forma que un grado mayor de correlación sólo es accesible a través del entrelazamiento. La consistencia de $\|T\|_{tr}$ como medida de las correlaciones es respaldada por el hecho de que toma el mismo valor para todos los operadores densidad que se relacionan a través de un cambio de base local y representan el mismo estado físico, i.e. es invariante bajo transformaciones unitarias locales. Este hecho se mostró en [99] para $M = N = 2$; en la siguiente proposición damos una demostración general (véase también [168]).

Proposición 6.2. Sean $U_A \in \mathcal{L}(H_A)$ y $U_B \in \mathcal{L}(H_B)$ transformaciones unitarias. Si

$$\rho' = (U_A \otimes U_B) \rho (U_A^\dagger \otimes U_B^\dagger), \quad (6.26)$$

entonces $\|T'\|_{tr} = \|T\|_{tr}$.

Demostración. Sean ρ_A y ρ'_A operadores densidad actuando en $H_A \simeq \mathbb{C}^M$ tales que $\rho'_A = U_A \rho_A U_A^\dagger$. Es directo comprobar (a través de (6.12)) que esta transformación induce una transformación lineal en los vectores de Bloch, lo cual significa que $\mathbf{r}' = A\mathbf{r}$ para una determinada matriz $A \in \mathbb{R}^{(N^2-1) \times (N^2-1)}$ ya que toda transformación lineal en un espacio vectorial admite una expresión matricial. Es más, como $\|\cdot\|_{HS}$ es unitariamente invariante se tiene que $\|\rho_A\|_{HS} = \|\rho'_A\|_{HS}$, y utilizando las relaciones de ortogonalidad (6.6)-(6.7) y la ecuación (6.12) se encuentra que

$$\|\rho_A\|_{HS}^2 = \frac{1}{M} \left(1 + \frac{2}{M} \|\mathbf{r}\|_2^2 \right), \quad (6.27)$$

lo cual lleva a que $\|\mathbf{r}\|_2 = \|\mathbf{r}'\|_2$. Esto implica que la matriz que relaciona los vectores de coherencia de estados unitariamente equivalentes ha de ser ortogonal, i.e. existe una rotación O_A que actúa en \mathbb{R}^{M^2-1} tal que $\mathbf{r}' = O_A \mathbf{r}$. Esto significa que

$$\sum_i U_A r_i \lambda_i U_A^\dagger = \sum_i (O_A \mathbf{r})_i \lambda_i. \quad (6.28)$$

Así, cuando un estado bipartito ρ se transforma de acuerdo a (6.26), habrá rotaciones O_A que actúa en \mathbb{R}^{M^2-1} y O_B que actúa en \mathbb{R}^{N^2-1} tales que

$$\mathbf{r}' = O_A \mathbf{r}, \quad \mathbf{s}' = O_B \mathbf{s}, \quad T' = O_A T O_B^\dagger, \quad (6.29)$$

y, así, el resultado se sigue teniendo en cuenta la invariancia unitaria de $\|\cdot\|_{tr}$. \square

Nótese que el único ingrediente en la demostración del Teorema 6.1 es la desigualdad triangular y en la de la Proposición 6.2 la invariancia unitaria de $\|\cdot\|_{tr}$. Así pues, se podría obtener resultados similares con otras normas. El hecho de que utilicemos la norma en traza se debe a que teniendo en cuenta (A.10), de entre las normas más razonables ésta es la que ofrece una condición más restrictiva (i.e., más potente para la detección de estados entrelazados). No obstante, como veremos en el siguiente capítulo, desde otra perspectiva puede resultar interesante considerar otras normas.

Además de la condición necesaria para la separabilidad obtenida en el Teorema 6.1, la caracterización del problema de la separabilidad dada en (6.22) también sugiere la posibilidad de derivar condiciones suficientes para la separabilidad utilizando una prueba constructiva. Una tal posibilidad se refiere en la siguiente proposición.

Proposición 6.3. *Si un estado bipartito en $M \times N$ dimensiones con representación de Bloch (6.16) satisface*

$$\sqrt{\frac{2(M-1)}{M}} \|\mathbf{r}\|_2 + \sqrt{\frac{2(N-1)}{N}} \|\mathbf{s}\|_2 + \sqrt{\frac{4(M-1)(N-1)}{MN}} \|T\|_{tr} \leq 1, \quad (6.30)$$

entonces es un estado separable.

Demostración. Tenga T la descomposición en valores singulares $T = \sum_i \sigma_i \mathbf{u}_i \mathbf{v}_i^t$ (lo que significa que $\|\mathbf{u}_i\|_2 = \|\mathbf{v}_i\|_2 = 1$). Si definimos

$$\tilde{\mathbf{u}}_i = \sqrt{\frac{M}{2(M-1)}} \mathbf{u}_i, \quad \tilde{\mathbf{v}}_i = \sqrt{\frac{N}{2(N-1)}} \mathbf{v}_i, \quad (6.31)$$

entonces

$$T = \sqrt{\frac{4(M-1)(N-1)}{MN}} \sum_i \sigma_i \tilde{\mathbf{u}}_i \tilde{\mathbf{v}}_i^t. \quad (6.32)$$

Así, si se cumple la condición (6.30), se puede descomponer ρ como la siguiente combinación convexa de matrices densidad ϱ_i , ϱ'_i , ρ_r , ρ_s y $\frac{1}{MN} I_{MN}$,

$$\begin{aligned} \rho &= \sqrt{\frac{4(M-1)(N-1)}{MN}} \frac{1}{2} \sum_i \sigma_i (\varrho_i + \varrho'_i) + \sqrt{\frac{2(M-1)}{M}} \|\mathbf{r}\|_2 \rho_r \\ &+ \sqrt{\frac{2(N-1)}{N}} \|\mathbf{s}\|_2 \rho_s + \left(1 - \sqrt{\frac{2(M-1)}{M}} \|\mathbf{r}\|_2 \right. \\ &\left. - \sqrt{\frac{2(N-1)}{N}} \|\mathbf{s}\|_2 - \sqrt{\frac{4(M-1)(N-1)}{MN}} \|T\|_{tr} \right) \frac{1}{MN} I_{MN}, \quad (6.33) \end{aligned}$$

donde $\varrho_i, \varrho'_i, \rho_r$ y ρ_s son tales que

$$\begin{aligned} \mathbf{r}_i &= \tilde{\mathbf{u}}_i, & \mathbf{s}_i &= \tilde{\mathbf{v}}_i, & T_i &= \tilde{\mathbf{u}}_i \tilde{\mathbf{v}}_i^t, \\ \mathbf{r}'_i &= -\tilde{\mathbf{u}}_i, & \mathbf{s}'_i &= -\tilde{\mathbf{v}}_i, & T'_i &= \tilde{\mathbf{u}}_i \tilde{\mathbf{v}}_i^t, \\ \mathbf{r}_r &= \sqrt{\frac{M}{2(M-1)}} \frac{\mathbf{r}}{\|\mathbf{r}\|_2}, & \mathbf{s}_r &= 0, & T_r &= 0, \\ \mathbf{r}_s &= 0, & \mathbf{s}_s &= \sqrt{\frac{N}{2(N-1)}} \frac{\mathbf{s}}{\|\mathbf{s}\|_2}, & T_s &= 0. \end{aligned}$$

Nótese que en virtud de (6.15) todos estos vectores de coherencia pertenecen a los correspondientes espacios de Bloch y, por tanto, las reducciones de $\varrho_i, \varrho'_i, \rho_r$ y ρ_s constituyen matrices densidad. Además, todas estas matrices satisfacen la condición (6.20), por lo que son iguales al producto tensorial de sus reducciones. Por tanto, $\varrho_i, \varrho'_i, \rho_r$ y ρ_s son efectivamente matrices densidad y, además, separables. Consecuentemente, también lo es ρ . \square

Resulta natural preguntarse si la Proposición 6.3 se puede fortalecer utilizando una condición no tan simple como (6.30). Como veremos en el siguiente teorema, la respuesta resulta ser positiva.

Teorema 6.2. *Sea*

$$c = \max \left\{ \sqrt{\frac{2(M-1)}{M}} \|\mathbf{r}\|_2, \sqrt{\frac{2(N-1)}{N}} \|\mathbf{s}\|_2 \right\}. \quad (6.34)$$

Si un estado bipartito en $M \times N$ dimensiones con representación de Bloch (6.16) tal que $c \neq 0$ satisface

$$c + \sqrt{\frac{4(M-1)(N-1)}{MN}} \left\| T - \frac{\mathbf{r} \mathbf{s}^t}{c} \right\|_{tr} \leq 1, \quad (6.35)$$

entonces es un estado separable.

Demostración. Análogamente a la prueba de la Proposición 6.3, tenga $T - \frac{\mathbf{r} \mathbf{s}^t}{c}$ la descomposición en valores singulares $\sum_i \sigma'_i \mathbf{x}_i \mathbf{y}_i^t$, donde $\|\mathbf{x}_i\|_2 = \|\mathbf{y}_i\|_2 = 1$. Si definimos igualmente

$$\tilde{\mathbf{x}}_i = \sqrt{\frac{M}{2(M-1)}} \mathbf{x}_i, \quad \tilde{\mathbf{y}}_i = \sqrt{\frac{N}{2(N-1)}} \mathbf{y}_i, \quad (6.36)$$

entonces se tiene que

$$T - \frac{\mathbf{r} \mathbf{s}^t}{c} = \sqrt{\frac{4(M-1)(N-1)}{MN}} \sum_i \sigma'_i \tilde{\mathbf{x}}_i \tilde{\mathbf{y}}_i^t. \quad (6.37)$$

Ahora, si la condición (6.35) se satisface, entonces ρ admite la siguiente descomposición como combinación convexa de estados separables

$$\begin{aligned} \rho &= \sqrt{\frac{4(M-1)(N-1)}{MN}} \frac{1}{2} \sum_i \sigma'_i(\varrho_i + \varrho'_i) + c\rho_{rs} \\ &+ \left(1 - c - \sqrt{\frac{4(M-1)(N-1)}{MN}} \left\| T - \frac{\mathbf{r}\mathbf{s}^t}{c} \right\|_{tr} \right) \frac{1}{MN} I_{MN}, \end{aligned} \quad (6.38)$$

donde ϱ_i , ϱ'_i y ρ_{rs} son ahora tales que

$$\mathbf{r}_i = \tilde{\mathbf{x}}_i, \quad \mathbf{s}_i = \tilde{\mathbf{y}}_i, \quad T_i = \tilde{\mathbf{x}}_i \tilde{\mathbf{y}}_i^t,$$

$$\mathbf{r}'_i = -\tilde{\mathbf{x}}_i, \quad \mathbf{s}'_i = -\tilde{\mathbf{y}}_i, \quad T'_i = \tilde{\mathbf{x}}_i \tilde{\mathbf{y}}_i^t,$$

$$\mathbf{r}_{rs} = \frac{\mathbf{r}}{c}, \quad \mathbf{s}_{rs} = \frac{\mathbf{s}}{c}, \quad T_{rs} = \frac{\mathbf{r}\mathbf{s}^t}{c^2}.$$

Como en la anterior demostración, y dado que

$$\frac{\mathbf{r}}{c} \leq \sqrt{\frac{M}{2(M-1)}} \frac{\mathbf{r}}{\|\mathbf{r}\|_2}, \quad \frac{\mathbf{s}}{c} \leq \sqrt{\frac{N}{2(N-1)}} \frac{\mathbf{s}}{\|\mathbf{s}\|_2},$$

todos estos vectores de coherencia pertenecen a sus respectivos espacios de Bloch, y ϱ_i , ϱ'_i y ρ_{rs} cumplen (6.20). \square

Nótese que la aplicación de la desigualdad triangular a (6.35) muestra claramente que el Teorema 6.2 es más fuerte que la Proposición 6.3. No obstante, esta proposición ofrece la forma adecuada de entender el límite $c \rightarrow 0$ en el Teorema 6.2. Destacamos que la demostración de estos dos resultados es constructiva, por lo que para los estados que satisfagan (6.30) y/o (6.35) no sólo se discierne su separabilidad sino que también se obtiene una descomposición como combinación convexa de estados separables. Estos estados pueden no ser puros, pero son el producto tensorial de sus reducciones, por lo que para obtener una descomposición en estados producto como en la definición original de separabilidad basta con aplicar la descomposición espectral a las reducciones de ϱ_i , ϱ'_i , ρ_r , ρ_s y/o ρ_{rs} . Estas condiciones suficientes para la separabilidad son un reflejo del hecho de que cuanto más cercano es un estado al estado de máxima ignorancia (o sea, la identidad) más probable es que sea separable.

Debe subrayarse que es posible obtener otras condiciones suficientes para la separabilidad siguiendo este enfoque. Pero, mientras que las condiciones del Teorema 6.2 y de la Proposición 6.3 dependen sólo de las normas de \mathbf{r} , \mathbf{s} y T , las otras posibilidades llevan a incluir más parámetros. Por ejemplo, es posible obtener la siguiente condición suficiente para la separabilidad, que también

depende de la forma de la descomposición en valores singulares de T ,

$$\begin{aligned} & \left\| \sqrt{\frac{N}{2(N-1)}} \mathbf{r} - \sum_i \sigma_i \mathbf{u}_i \right\|_2 + \left\| \sqrt{\frac{M}{2(M-1)}} \mathbf{s} - \sum_i \sigma_i \mathbf{v}_i \right\|_2 + \|T\|_{tr} \\ & \leq \sqrt{\frac{MN}{4(M-1)(N-1)}}, \quad (6.39) \end{aligned}$$

ya que en este caso ρ admite una descomposición en estados separables como en (6.33) pero con $\varrho'_i = \varrho_i$,

$$\mathbf{r}_r = \sqrt{\frac{M}{2(M-1)}} \frac{\mathbf{r} - \sqrt{\frac{2(N-1)}{N}} \sigma_i \mathbf{u}_i}{\left\| \mathbf{r} - \sqrt{\frac{2(N-1)}{N}} \sigma_i \mathbf{u}_i \right\|_2}, \quad \mathbf{s}_s = \sqrt{\frac{N}{2(N-1)}} \frac{\mathbf{s} - \sqrt{\frac{2(M-1)}{M}} \sigma_i \mathbf{v}_i}{\left\| \mathbf{s} - \sqrt{\frac{2(M-1)}{M}} \sigma_i \mathbf{v}_i \right\|_2}.$$

Sin embargo, no parece que complicar las condiciones de esta forma conduzca a criterios claramente más fuertes.

Una última cuestión que surge llegados a este punto es la posibilidad de combinar por un lado la condición necesaria para la separabilidad desarrollada en el Teorema 6.1 con alguna de las condiciones suficientes que acabamos de ver para poder obtener una caracterización del entrelazamiento al menos para un conjunto particular de estados. Esto es efectivamente posible puesto que (6.23) y (6.30) toman la misma forma en el caso de estados de dos qubits tales que $\mathbf{r} = \mathbf{s} = 0$.

Corolario 6.1. *Los estados de dos qubits ($M = N = 2$) con subsistemas maximalmente mezclados² (i.e. $\rho_A = I_M/M$, $\rho_B = I_N/N$) son separables si y sólo si $\|T\|_{tr} \leq 1$.*

Debe aclararse que este último resultado no es nuevo ya que la separabilidad de este tipo de estados ya había sido caracterizada mediante una condición equivalente en [99], aunque debe destacarse que aquí se obtiene como un simple corolario de los resultados anteriores.

Otro aspecto de la teoría cuántica en el que las correlaciones juegan un papel destacado es en las desigualdades de Bell, que muestran que correlaciones accesibles cuánticamente son irreproducibles por teorías realistas locales. Resulta interesante, por tanto, estudiar la conexión de las condiciones de separabilidad aquí obtenidas con la caracterización de los estados de dos qubits que violan una desigualdad de Bell dada en [169]. En concreto, allí se mostró que estos estados violan la desigualdad CHSH [170] si y sólo si los valores singulares de la matriz de correlación satisfacen $\sigma_1^2 + \sigma_2^2 > 1$. Comparando este resultado con el Teorema 6.1 en el caso $M = N = 2$ que establece que si $\sigma_1 + \sigma_2 + \sigma_3 > 1$ el estado es entrelazado, el bien conocido resultado de que existen estados entrelazados que no violan una desigualdad de Bell [81, 171] se manifiesta claramente. Por otro lado, la caracterización de [169] refuerza la idea de que los valores singulares de la matriz de correlación cuantifican de alguna forma el grado de correlación de un estado cuántico.

²También llamados estados diagonales en la base de Bell.

6.3. Eficacia de los nuevos criterios

6.3.1. Ejemplos

En lo que sigue consideramos varios ejemplos para mostrar la utilidad de los criterios de separabilidad derivados en la anterior sección. Comenzamos mostrando que el Teorema 6.1 es lo suficientemente fuerte como para detectar entrelazamiento PPT.

Ejemplo 1: Consideremos el siguiente estado entrelazado PPT en dimensiones 3×3 construido en [172]:

$$\rho = \frac{1}{4} \left(I_9 - \sum_{i=0}^4 |\psi_i\rangle\langle\psi_i| \right), \quad (6.40)$$

donde

$$\begin{aligned} |\psi_0\rangle &= \frac{1}{\sqrt{2}}|0\rangle(|0\rangle - |1\rangle), & |\psi_1\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)|2\rangle, & |\psi_2\rangle &= \frac{1}{\sqrt{2}}|2\rangle(|1\rangle - |2\rangle), \\ |\psi_3\rangle &= \frac{1}{\sqrt{2}}(|1\rangle - |2\rangle)|0\rangle, & |\psi_4\rangle &= \frac{1}{3}(|0\rangle + |1\rangle + |2\rangle)(|0\rangle + |1\rangle + |2\rangle). \end{aligned} \quad (6.41)$$

Para obtener la representación de Bloch de este estado utilizamos como generadores de $SU(3)$ los habituales operadores de Gell-Mann, que son simplemente una reordenación de los dados en (6.9)-(6.11),

$$\lambda_1 = u_{01}, \lambda_2 = v_{01}, \lambda_3 = w_0, \lambda_4 = u_{02}, \lambda_5 = v_{02}, \lambda_6 = u_{12}, \lambda_7 = v_{12}, \lambda_8 = w_1. \quad (6.42)$$

Así, para el estado (6.40) se encuentra que

$$T = -\frac{1}{4} \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 & \frac{\sqrt{27}}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{9}{4} & 0 & -\frac{9}{8} & 0 & 0 & 0 & 0 & \frac{\sqrt{27}}{8} \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -\frac{9}{4} & 1 & 0 & 1 & 0 & -\frac{\sqrt{27}}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{\sqrt{27}}{4} & 0 & \frac{\sqrt{27}}{8} & 0 & 0 & \frac{\sqrt{27}}{2} & 0 & -\frac{3}{8} \end{pmatrix}, \quad (6.43)$$

de forma que $\|T\|_{tr} \simeq 3,1603$, violándose la condición (6.23). Por tanto, a través del Teorema 6.1 deducimos que este estado es entrelazado.

El ejemplo anterior muestra que existen casos en los que el Teorema 6.1 es más fuerte que el criterio PPT. Aún así, esto no es cierto en general, ni siquiera en el caso 2×2 .

Ejemplo 2: Consideremos el siguiente estado de dos qubits,

$$\rho_{\pm} = p|\psi^{\pm}\rangle\langle\psi^{\pm}| + (1-p)|00\rangle\langle 00|, \quad (6.44)$$

donde $p \in [0, 1]$. El criterio de Peres-Horodecki establece que este estado es separable si y sólo si $p = 0$ [93]. Utilizando como generadores de $SU(2)$ las ya referidas matrices de Pauli $\sigma_x = u_{01}$, $\sigma_y = v_{01}$ and $\sigma_z = w_0$ se encuentra que

$$\rho_{\pm} = \frac{1}{4}(I_2 \otimes I_2 + (1-p)\sigma_z \otimes I_2 + (1-p)I_2 \otimes \sigma_z \pm p\sigma_x \otimes \sigma_x \pm p\sigma_y \otimes \sigma_y + (1-2p)\sigma_z \otimes \sigma_z). \quad (6.45)$$

Por consiguiente, $\|T\|_{tr} = 2p + |1 - 2p|$, lo que supone que $\|T\|_{tr} \leq 1$ si $p \leq 1/2$, por lo que el entrelazamiento sólo es detectado si $p > 1/2$.

Ejemplo 3: Los estados de Werner [81] en dimensiones arbitrarias ($M = N = D$) son aquellos cuyas matrices densidad son invariantes bajo transformaciones de la forma $(U \otimes U)\rho(U^\dagger \otimes U^\dagger)$, donde $U \in \mathbb{C}^{D \times D}$ es una matriz unitaria. Se pueden expresar como

$$\rho_W = \frac{1}{D^3 - D}[(D - \phi)I_D \otimes I_D + (D\phi - 1)V], \quad (6.46)$$

donde $-1 \leq \phi \leq 1$ y V es el operador de “flip” o “swap”, que se define como $V\varphi \otimes \tilde{\varphi} = \tilde{\varphi} \otimes \varphi$. Estos estados son separables si y sólo si $\phi \geq 0$ [81]. Utilizando (6.19) o invirtiendo las ecuaciones (6.9)-(6.11) se encuentra que

$$V = \sum_{ij} |ij\rangle\langle ji| = \frac{1}{D}I_D \otimes I_D + \frac{1}{2} \sum_l w_l \otimes w_l + \frac{1}{2} \sum_{j < k} (u_{jk} \otimes u_{jk} + v_{jk} \otimes v_{jk}), \quad (6.47)$$

por lo que

$$\rho_W = \frac{1}{D^2} \left(I_D \otimes I_D + \frac{D(D\phi - 1)}{2(D^2 - 1)} \sum_i \lambda_i \otimes \lambda_i \right), \quad (6.48)$$

donde $\{\lambda_i\}$ son los generadores de $SU(D)$ definidos en (6.9)-(6.11). Así, $\|T\|_{tr} = D|D\phi - 1|/2$, de forma que el Teorema 6.1 sólo reconoce con total certidumbre al estado como entrelazado cuando $\phi \leq (2 - D)/D$, mientras que la Proposición 6.3 garantiza que el estado es separable si $(D - 2)/[D(D - 1)] \leq \phi \leq 1/(D - 1)$. Cuando esta última condición se satisface se puede obtener la descomposición de estos estados como combinación convexa de estados producto. Para ilustrar el procedimiento, consideremos el estado de Werner en, por simplicidad, dimensiones 2×2 . En este caso $V = I_2 \otimes I_2 - 2|\psi^-\rangle\langle\psi^-|$, y definiendo $p = (1 - 2\phi)/3$ el estado toma la sencilla forma (5.26), que ya vimos en el capítulo anterior,

$$\rho = \frac{1-p}{4}I_2 \otimes I_2 + p|\psi^-\rangle\langle\psi^-| = \frac{1}{4}(I_2 \otimes I_2 - p\sigma_x \otimes \sigma_x - p\sigma_y \otimes \sigma_y - p\sigma_z \otimes \sigma_z). \quad (6.49)$$

Con el Corolario 6.1 comprobamos que ρ es separable si y sólo si $p \leq 1/3$ como cabía esperar. Siguiendo los pasos de la demostración de la Proposición 6.3 encontramos que

$$\rho = \sum_{i=x,y,z} \sum_{j=1}^2 \frac{p}{2} \rho_j^{(i)} + (1 - 3p) \frac{1}{4}(I_2 \otimes I_2), \quad (6.50)$$

donde

$$\rho_1^{(i)} = \frac{1}{4}(I_2 \otimes I_2 + \sigma_i \otimes I_2 - I_2 \otimes \sigma_i - \sigma_i \otimes \sigma_i), \quad \rho_2^{(i)} = \frac{1}{4}(I_2 \otimes I_2 - \sigma_i \otimes I_2 + I_2 \otimes \sigma_i - \sigma_i \otimes \sigma_i). \quad (6.51)$$

En este caso podemos reducir el número de estados producto en la descomposición a 8 observando que $\rho_1^{(i)} = |01\rangle_i \langle 01|$ y $\rho_2^{(i)} = |10\rangle_i \langle 10|$, donde $\{|0\rangle_i, |1\rangle_i\}$ denotan los autovectores de σ_i , de forma que, por ejemplo,

$$\begin{aligned} \rho = & \sum_{i=x,y} \frac{p}{2} (|01\rangle_i \langle 01| + |10\rangle_i \langle 10|) + \frac{1-p}{4} (|01\rangle_z \langle 01| + |10\rangle_z \langle 10|) \\ & + \frac{1-3p}{4} (|00\rangle_z \langle 00| + |11\rangle_z \langle 11|). \end{aligned} \quad (6.52)$$

Se sabe, sin embargo, que un estado separable de dos qubits admite una descomposición con un número de estados producto menor o igual a 4 [87, 88].

Ejemplo 4: Los estados isotrópicos [95] en dimensiones arbitrarias ($M = N = D$) son invariantes bajo transformaciones de la forma $(U \otimes U^*)\rho(U^\dagger \otimes U^{*\dagger})$. Se pueden escribir como mezclas del estado maximalmente mezcla y del estado maximalmente entrelazado

$$|\Psi\rangle = \frac{1}{\sqrt{D}} \sum_{a=0}^{D-1} |aa\rangle, \quad (6.53)$$

de forma que se expresan como³

$$\rho = \frac{1-p}{D^2} I_D \otimes I_D + p|\Psi\rangle\langle\Psi|. \quad (6.54)$$

Estos estados son separables si y sólo si $p \leq (D+1)^{-1}$ [95] (véase también [167]). Su representación de Bloch, que se encuentra fácilmente como en el caso de los estados de Werner, está dada por

$$\rho = \frac{1}{D^2} \left(I_D \otimes I_D + \frac{pD}{2} \sum_{i=1}^{(D+2)(D-1)/2} \lambda_i \otimes \lambda_i - \frac{pD}{2} \sum_{i=D(D+1)/2}^{D^2-1} \lambda_i \otimes \lambda_i \right), \quad (6.55)$$

donde, como antes, los $\{\lambda_i\}$ están dados por (6.9)-(6.11). Ahora, $\|T\|_{tr} = pD(D^2 - 1)/2$. Así, el Teorema 6.1 es lo suficientemente fuerte como para detectar todos los estados entrelazados de esta clase ($\|T\|_{tr} \leq D(D-1)/2 \Leftrightarrow p \leq (D+1)^{-1}$), mientras que la Proposición 6.3 asegura que los estados son separables cuando $p \leq (D+1)^{-1}(D-1)^{-2}$.

³En el caso de dos qubits, los estados de Werner (5.26) ($U \otimes U$ invariantes) y los isotrópicos (6.54) ($U \otimes U^*$ invariantes) son equivalentes en el sentido de que están relacionados por una transformación unitaria local. Por esta razón algunos autores se refieren a los estados isotrópicos como estados de Werner generalizados, lo que puede llevar a confusión.

6.3.2. Comparación con otros criterios importantes

En la sección anterior hemos visto que el criterio dado por el Teorema 6.1 es lo suficientemente fuerte como para detectar entrelazamiento PPT. Esta situación es bastante grata puesto que, como ya hemos mencionado, uno de los objetivos dentro del problema de la separabilidad es encontrar condiciones que complementen las carencias del potente criterio de Peres-Horodecki. Queda por aclarar, sin embargo, una cuestión importante. Sería conveniente saber hasta qué punto permite el Teorema 6.1 rellenar los huecos que deja este criterio, y, principalmente, comparar su capacidad con la del principal criterio operacional capaz también de detectar entrelazamiento PPT, el criterio CCNR. Así mismo, también sería deseable establecer cuál es la relación entre nuestro criterio y el no operacional de las LURs en su versión más fuerte (2.69).

En lo que se refiere al criterio CCNR, al igual que el Teorema 6.1, éste es capaz de detectar todos los estados isotrópicos y reconoce como entrelazados el mismo rango de estados de Werner [102]. Sin embargo, este criterio detecta óptimamente el entrelazamiento del estado del Ejemplo 2 [102], por lo que se podría pensar que es más fuerte que el Teorema 6.1. Para comprobar esta posibilidad y evaluar la capacidad de detección de entrelazamiento PPT de los criterios, siguiendo [173] que proporciona una receta para generar un gran número de estados PPT entrelazados 3×3 denominados “estados de tablero de ajedrez”, hemos generado al azar 10^6 estados de este tipo. Hemos encontrado que nuestro teorema detectó entrelazamiento entorno al 4 % de los estados mientras que el criterio CCNR reconoció como entrelazados aproximadamente al 18 % de ellos. Más aún, todos los estados detectados por el Teorema 1 fueron también detectados por el criterio CCNR. Esto sugiere que el criterio CCNR es más fuerte que el Teorema 6.1. En lo que sigue, demostramos analíticamente que este es efectivamente el caso cuando $M = N$, pero también mostraremos, remarcablemente, que esto no es cierto cuando $M \neq N$. Necesitaremos el siguiente lema:

Lema 6.1.

$$\left\| \begin{pmatrix} A & B \\ C & D \end{pmatrix} \right\|_{tr} \geq \|A\|_{tr} + \|D\|_{tr},$$

donde A, B, C, D son matrices complejas cualesquiera (de dimensión adecuada para que tenga sentido la partición por bloques).

Demostración. Tengan A y D las descomposiciones en valores singulares $A = U_A \Sigma_A V_A^\dagger$ y $D = U_D \Sigma_D V_D^\dagger$. Utilizando la invariancia unitaria de la norma en traza se tiene entonces que

$$\begin{aligned} \left\| \begin{pmatrix} A & B \\ C & D \end{pmatrix} \right\|_{tr} &= \left\| \begin{pmatrix} U_A^\dagger & 0 \\ 0 & U_D^\dagger \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} V_A & 0 \\ 0 & V_D \end{pmatrix} \right\|_{tr} \\ &\geq \text{Tr } \Sigma_A + \text{Tr } \Sigma_D, \end{aligned} \quad (6.56)$$

donde hemos utilizado que $\|X\|_{tr} \geq \text{Tr } X$, que es una consecuencia directa de la siguiente caracterización de la norma en traza (véase la ecuación (3.4.7) en

[174]):

$$\|X\|_{tr} = \max\{|\text{Tr } XU| : U \text{ es unitaria}\}. \quad (6.57)$$

□

Proposición 6.4. *En el caso de estados con subsistemas maximalmente mezcla el Teorema 6.1 es más fuerte que el criterio CCNR cuando $M \neq N$, mientras que cuando $M = N$ ambos son equivalentes.*

Demostración. Cuando $\mathbf{r} = \mathbf{s} = 0$ tenemos que

$$R(\rho) = \frac{1}{MN} (|I_M\rangle\langle I_N| + \sum_{ij} t_{ij} |\lambda_i\rangle\langle \tilde{\lambda}_j^*|). \quad (6.58)$$

Como la matriz asociada al operador $R(\rho)$ es en este caso diagonal por bloques se tiene que

$$\begin{aligned} \|R(\rho)\|_{tr} &= \frac{1}{\sqrt{MN}} \left\| \left\| \frac{|I_M\rangle\langle I_N|}{\sqrt{M}\sqrt{N}} \right\|_{tr} \right\| + \frac{2}{MN} \left\| \left\| \sum_{ij} t_{ij} \frac{|\lambda_i\rangle\langle \tilde{\lambda}_j^*|}{\sqrt{2}\sqrt{2}} \right\|_{tr} \right\| \\ &= \frac{1}{\sqrt{MN}} + \frac{2}{MN} \|T\|_{tr}. \end{aligned} \quad (6.59)$$

Así, en este caso el criterio CCNR es equivalente a

$$\|T\|_{tr} \leq \frac{\sqrt{MN}(\sqrt{MN} - 1)}{2}, \quad (6.60)$$

de lo que se sigue el resultado puesto que utilizando que $M + N \geq 2\sqrt{MN}$ se tiene $\sqrt{MN} - 1 \geq \sqrt{(M-1)(N-1)}$. □

Proposición 6.5. *El criterio CCNR es más fuerte que el Teorema 6.1 cuando $M = N$.*

Demostración. Como en este caso en general $\mathbf{r}, \mathbf{s} \neq 0$, la matriz asociada al operador $R(\rho)$ ya no es diagonal por bloques. Por tanto, utilizando el Lema 6.1, se encuentra ahora que

$$\|R(\rho)\|_{tr} \geq \frac{1}{N} + \frac{2}{N^2} \|T\|_{tr}, \quad (6.61)$$

lo que prueba el resultado teniendo en cuenta que en el caso $M = N$ la condición (6.23) del Teorema 6.1 se puede escribir como

$$\frac{1}{N} + \frac{2}{N^2} \|T\|_{tr} \leq 1. \quad (6.62)$$

□

Por tanto, resulta que el Teorema 6.1 y el criterio CCNR son independientes. Sólo en el caso de subsistemas maximalmente mixtos de iguales dimensiones ambos criterios resultan equivalentes. Esto explica por qué se obtenían los mismos resultados con los dos criterios para los estados de Werner y los isotrópicos. No obstante, debe decirse que los cálculos en nuestro formalismo resultan mucho más simples que en los del criterio CCNR en estos casos [102].

Así pues, si bien hemos comprobado que el criterio CCNR es más fuerte que nuestro criterio para subsistemas con las mismas dimensiones, hemos verificado que lo contrario puede ocurrir cuando las dimensiones de los subsistemas son diferentes, lo cual sirve para demostrar el interés práctico en la detección de entrelazamiento del Teorema 6.1 ya que, en conclusión, nuestro criterio es capaz de detectar por un lado estados entrelazados que se le escapan al criterio PPT y, por otro lado, estados entrelazados que no reconoce el criterio CCNR. Sin embargo, queda en el aire comprobar si el Teorema 6.1 mejora efectivamente nuestra capacidad para detectar estados entrelazados reconociendo estados de este tipo que no son identificados como tales por ninguno de los dos criterios anteriores a la vez. A continuación, comprobaremos que esto es efectivamente posible construyendo ejemplos de estados cuyo entrelazamiento es únicamente revelado por el Teorema 6.1. Para esta tarea nos tendremos que centrar en estados PPT entrelazados tales que $M \neq N$, ya que los de subsistemas de dimensiones iguales, como los considerados hasta ahora, que sean reconocidos por el Teorema 6.1 también lo son por el criterio CCNR debido a la Proposición 6.5. Desafortunadamente, aunque se conocen bastantes ejemplos de estados entrelazados PPT cuando $M = N$ (véase por ejemplo [175, 176, 172, 173, 177]), la situación en el caso asimétrico no es tan rica. En [175] se construye un conjunto de estados entrelazados PPT para sistemas 2×4 (véase el artículo para su forma explícita). Sin embargo, se puede comprobar que tanto el Teorema 6.1 como el criterio CCNR son incapaces de identificar estos estados como entrelazados igualmente. En [177], a partir del formalismo de las bases producto no extendibles (UPBs) ideado en [172], se encuentra un estado PPT entrelazado con subsistemas de diferentes dimensiones ($M, N \geq 3$), el cual es obtenido a partir de la UPB conocida como GenTiles2 (su forma explícita se encuentra más abajo). No obstante, mientras que encontramos que el criterio CCNR detecta entrelazamiento en este estado cuando $M = 3, 4 \leq N \leq 8$ y cuando $M = 4, N = 5$, el Teorema 6.1 resulta ser capaz de reconocer entrelazamiento sólo en el caso 3×4 . Por tanto, desafortunadamente para nuestro propósito el criterio CCNR parece más fuerte para estos estados. El entrelazamiento en los estados con invariancia rotacional se ha estudiado en detalle en la literatura reciente. En particular, se ha determinado el conjunto de estados PPT con invariancia rotacional en dimensiones $3 \times N$ [178] y $4 \times N$ [179], se ha mostrado que el criterio de Peres-Horodecki no caracteriza al entrelazamiento en este escenario y muchos estados entrelazados PPT se han identificado. Aunque esta situación parezca prometedora para nuestro objetivo puesto que los estados con invariancia rotacional tienen subsistemas maximalmente mezclados y, en consecuencia, la Proposición 6.4 nos garantiza que el Teorema 6.1 es más fuerte que el criterio CCNR, desgraciadamente, nuestras exploraciones numéricas

indican que el criterio PPT es más fuerte que el Teorema 6.1 para este conjunto de estados.

Los estados considerados arriba son, hasta donde llega nuestro conocimiento, los únicos ejemplos disponibles en la literatura de entrelazamiento PPT con subsistemas de diferente dimensionalidad. Por consiguiente, para alcanzar nuestro objetivo nos vemos obligados a, o bien construir nuevos ejemplos, o bien considerar operaciones que preserven la positividad de la traspuesta parcial y que transformen los estados anteriores en estados diferentes en los que el Teorema 6.1 sea más fuerte que el criterio CCNR, como, por ejemplo, el conjunto de los estados con subsistemas maximalmente mezcla. Resulta que este tipo de transformaciones en concreto ha sido muy bien estudiado en la literatura previa. En [180] (véase también [181]) se muestra que cualquier estado cuya matriz densidad sea invertible se puede transformar bajo la acción de filtros locales (SLOCC) en un estado con subsistemas maximalmente desordenados que se denomina forma normal filtro (FNF). Esta forma es única salvo transformaciones unitarias locales. Matemáticamente, esta operación de SLOCC se representa por operadores de Kraus locales dados por matrices invertibles F_A, F_B (sin pérdida de generalidad se escogen, además, de determinante unidad) que transforman ρ en su FNF $\tilde{\rho}$ de acuerdo a

$$\tilde{\rho} = \frac{(F_A \otimes F_B)\rho(F_A \otimes F_B)^\dagger}{\text{Tr}(F_A \otimes F_B)\rho(F_A \otimes F_B)^\dagger}. \quad (6.63)$$

Existen varios algoritmos constructivos para esta transformación que, dado ρ , obtienen las matrices F_A, F_B necesarias para llevarlo a su forma normal [180, 181] (véase también [182]). En particular, el lector puede hallar un código matlab en [183]. Además, aunque la existencia de la FNF sólo se garantiza para matrices densidad invertibles, los estados con matrices densidad de determinante nulo pueden ser transformados en estados con reducciones arbitrariamente cerca del estado maximalmente mezcla mediante los algoritmos anteriormente referidos. Nótese que, por construcción, la operación dada por (6.63) transforma estados separables en estados separables. Es más, puesto que la operación es invertible, sólo transforma en estados separables a los estados separables, y, por tanto, esta operación preserva la separabilidad o entrelazamiento de los estados. Análogamente, se puede verificar que la transformación (6.63) también preserva la positividad de la traspuesta parcial [181]. En conclusión, la transformación a FNF lleva estados PPT entrelazados a estados PPT entrelazados con subsistemas maximalmente mezclados, por lo que el Teorema 6.1 será más fuerte que los criterios PPT y CCNR para los estados $\tilde{\rho}$ obtenidos de esta forma.

Comenzamos entonces considerando al estado PPT entrelazado ρ_{GT2} , construido a partir de la UPB GenTiles2 [177, 184], en dimensiones $M \times N$

tales que $M \geq 3$, $N > 3$ y $M \leq N$,

$$\rho_{GT2} = \frac{1}{2M-1} \left(I_{MN} - |F\rangle\langle F| - \sum_{j=0}^{M-1} |S_j\rangle\langle S_j| - \sum_{j=0}^{M-1} \sum_{k=1}^{N-3} |L_{jk}\rangle\langle L_{jk}| \right), \quad (6.64)$$

donde

$$\begin{aligned} |F\rangle &= \frac{1}{\sqrt{MN}} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} |ij\rangle, \\ |S_j\rangle &= \frac{1}{\sqrt{2}} (|j\rangle - |j+1 \bmod M\rangle) |j\rangle, \\ |L_{jk}\rangle &= \frac{1}{\sqrt{N-2}} |j\rangle \left(\sum_{l=0}^{M-3} \exp\left(i \frac{2\pi lk}{N-2}\right) |l+j+1 \bmod M\rangle \right. \\ &\quad \left. + \sum_{l=M-2}^{N-3} \exp\left(i \frac{2\pi lk}{N-2}\right) |l+2\rangle \right). \end{aligned} \quad (6.65)$$

Por simplicidad nos restringiremos a las dimensiones más bajas posibles $M = 3$ y $N = 4$. En este caso se obtiene que $\|T_{\rho_{GT2}}\|_{tr} = 4,3428$ y $\|R(\rho_{GT2})\|_{tr} = 1,0315$ y, por tanto, como ya mencionamos, los dos criterios (Teorema 6.1 y CCNR) detectan el entrelazamiento de ρ_{GT2} . Se puede comprobar considerando alguno de los algoritmos mencionados más arriba que los filtros

$$F_A = \begin{pmatrix} -0,2586 - 0,4251i & -0,2586 - 0,4251i & -0,2586 - 0,4251i \\ 0,3421 - 0,3842i & 0,4402 + 0,2817i & -0,7824 + 0,1025i \\ 0,2784 - 0,6568i & -0,5774 + 0,4086i & 0,2990 + 0,2482i \end{pmatrix},$$

$F_B =$

$$\begin{pmatrix} -0,3118 - 0,3092i & -0,3118 - 0,3092i & -0,3118 - 0,3092i & -0,3118 - 0,3092i \\ 0,5499 - 0,2805i & 0,6414 - 0,0813i & -0,3307 + 0,0334i & -0,4303 + 0,1642i \\ -0,3932 - 0,1066i & 0,3198 - 0,3909i & -0,0427 - 0,7619i & 0,0580 + 0,6297i \\ 0,5358 + 0,3605i & 0,1113 - 0,5279i & 0,5169 - 0,0640i & -0,5820 + 0,1157i \end{pmatrix},$$

transforman ρ_{GT2} en su forma normal $\tilde{\rho}_{GT2}$. Ahora, se encuentra que $\|T_{\tilde{\rho}_{GT2}}\|_{tr} = 4,5751$ y $\|R(\tilde{\rho}_{GT2})\|_{tr} = 1,0512$, por lo que, de nuevo, ambos criterios identifican a este estado PPT como entrelazado. Sin embargo, ahora el Teorema 6.1 es más fuerte y, en consecuencia, más robusto al ruido. Así, si consideramos el estado anterior mezclado con ruido blanco, i.e.

$$\rho(p) = p\tilde{\rho}_{GT2} + (1-p)\frac{I_{12}}{12}, \quad (6.66)$$

se encuentra que el Teorema 6.1 identifica como entrelazados a los estados $\rho(p)$ tales que $p \geq 0,9274$ mientras que el criterio CCNR sólo los reconoce como entrelazados cuando $p \geq 0,9330$. Nótese que, por construcción, el criterio PPT es incapaz de identificar entrelazamiento en los estados $\rho(p)$. Así, este ejemplo

muestra que el criterio dado por el Teorema 6.1 puede detectar entrelazamiento en estados para los que tanto el criterio PPT como el CCNR fallan. Se pueden encontrar más ejemplos de esta situación considerando la FNF de ρ_{GT2} para otros valores de M y/o N . Sin embargo, debe destacarse que la FNF de los estados entrelazados PPT 2×4 de [175] no se detecta como entrelazada por ninguno de los criterios y que, de hecho, $\|T_{\tilde{\rho}}\|_{tr} < \|T_{\rho}\|_{tr}$ y $\|R(\tilde{\rho})\|_{tr} < \|R(\rho)\|_{tr}$ para muchos de estos estados.

En el desarrollo anterior hemos utilizado la transformación a la FNF simplemente como un medio para construir estados con unas determinadas características en nuestro objetivo de encontrar ejemplos de estados entrelazados detectados por el Teorema 6.1 pero no por los criterios PPT y CCNR. Sin embargo, como se puso por primera vez de manifiesto en [182], visto desde otro enfoque la transformación a la FNF resulta ser una herramienta muy poderosa en el marco del problema de la separabilidad que no queremos dejar de mencionar aquí. Esto se debe a la observación, ya señalada más arriba, de que ρ es entrelazado si y sólo si $\tilde{\rho}$ es entrelazado, de forma que se puede deducir el entrelazamiento de ρ verificando el de $\tilde{\rho}$. La utilidad de esta idea proviene de que para muchos criterios resulta más probable detectar el entrelazamiento en $\tilde{\rho}$ que en ρ . Esto se debe a que como se demuestra en [180] la transformación a la FNF maximiza una medida de entrelazamiento, lo que significa que $\tilde{\rho}$ está más entrelazado que ρ , lo cual intuitivamente resulta en una mayor probabilidad de violar una condición de separabilidad. Notablemente, este no es el caso de la condición dada por el criterio PPT ya que ρ es PPT si y sólo si $\tilde{\rho}$ es PPT. Sin embargo, la transformación a la FNF sí suele resultar en un aumento sensible de los valores de $\|T\|_{tr}$ y $\|R(\rho)\|_{tr}$ (aunque esto no es una regla, sirva como ejemplo de lo contrario el ya mencionado caso de los estados PPT entrelazados 2×4 de [175]), lo cual permite en muchas ocasiones detectar el entrelazamiento de ρ a partir del de $\tilde{\rho}$ aun cuando los criterios eran originalmente incapaces de hacerlo. Así, si se permiten hacer transformaciones a la FNF (lo cual implica implementar algoritmos numéricos en un ordenador) la eficacia de criterios como el CCNR o el Teorema 6.1 se ve considerablemente aumentada. Baste como ejemplo la simulación llevada a cabo en [182] con los estados del tablero de ajedrez en la que se halló que el número de estados detectados aumentaba hasta cerca del 98 % utilizando este truco⁴. En conclusión, si a uno le es permitido implementar algoritmos numéricos, el problema de la separabilidad se reduce a caracterizar el entrelazamiento en estados con subsistemas maximalmente mezclados, por lo que resulta entonces prioritario encontrar criterios de separabilidad que sean especialmente fuertes para este tipo de estados. Esto redundará, vía la Proposición 6.4, en una mayor importancia de la condición hallada en el Teorema 6.1, ya que para estados en la FNF ésta desbanca completamente al criterio CCNR al ser más potente. Debemos decir también que en [182]⁵, además de la idea de fortalecer los criterios de separabilidad transformando a la FNF, también

⁴Nótese que, como estos estados son 3×3 -dimensionales, en la FNF el Teorema 6.1 y el criterio CCNR son equivalentes.

⁵Nótese que este artículo es posterior a la publicación de los resultados presentados en este capítulo.

se desarrolla un criterio operacional a partir de las LURs que en la FNF es equivalente al Teorema 6.1 cuando $M = N$ y que es más fuerte cuando $M \ll N$ pero más débil cuando M y N tienen valores cercanos. Así, por ejemplo, en la detección de entrelazamiento en los estados $\rho(p)$ de (6.66), el Teorema 6.1 también es más fuerte que este nuevo criterio que reconoce entrelazamiento para $p \geq 0,9290$. Este hecho permite conjeturar la independencia del Teorema 6.1 del de las LURs.

Hasta aquí, hemos discutido y contrastado la potencia para la identificación del entrelazamiento del Teorema 6.1 como condición matemática. Nos gustaría terminar esta sección dedicando unas palabras a esta faceta desde el punto de vista de la implementación experimental. Destacablemente, la condición del Teorema 6.1 no depende de todos los parámetros que caracterizan al operador densidad sino que se descartan los parámetros locales \mathbf{r} y \mathbf{s} y sólo se tienen en cuenta los que forman la matriz de correlación. Esto implica que mientras criterios como el PPT y el CCNR exigen un conocimiento completo de la matriz densidad del estado ($M^2N^2 - 1$ parámetros por especificar), el criterio del Teorema 6.1 sólo necesita determinar $(M^2 - 1)(N^2 - 1)$ parámetros, que además se corresponden con los valores esperados de observables medibles localmente a través de (6.19), por lo que la implementación de este último requiere en un principio un menor esfuerzo experimental. Como hemos visto, los criterios basados en relaciones de incertidumbre también pueden requerir menos medidas que la determinación de ρ . Sin embargo, éstos suelen adolecer de otra desventaja tal y como se analiza en [185]. Resulta que el valor de la suma de desviaciones estándar no es invariante bajo transformaciones unitarias locales del operador densidad. Esto se traduce en que los observadores de A y B deben ser sumamente cuidadosos al alinear sus sistemas de referencia en sus correspondientes dispositivos experimentales ya que un cambio de base a nivel local altera la magnitud que están midiendo. Sin embargo, esto no ocurre en la implementación del Teorema 6.1 puesto que, como se demostró en la Proposición 6.2, el valor de $\|T\|_{tr}$ no se altera por cambios de base locales.

Capítulo 7

Estimación de medidas de entrelazamiento a partir de criterios de separabilidad

Como ya hemos analizado en la introducción en la sección 2.4.3, si la caracterización del entrelazamiento resulta ser un problema muy complicado, más lo es su cuantificación. Esto se debe, ya no a la construcción de medidas, que, como hemos visto, da lugar a una fenomenología bastante rica, sino a la evaluación de las mismas, que, a excepción de E_F , C y τ para sistemas de dos qubits, involucra la resolución de problemas de optimización enormemente complejos. Por ello, en la literatura reciente se viene investigando la posibilidad de, dejando aparte los resultados exactos obtenidos para estados muy particulares, encontrar formas de estimar medidas de entrelazamiento para estados generales. Atendiendo a la forma de (2.79) y (2.83)-(2.85), nótese que, por construcción, es directo encontrar cotas superiores para una medida de entrelazamiento dado el operador densidad del estado, mientras que desarrollar cotas inferiores plantea un problema más duro. Sin embargo, este último tipo de cotas es más deseable ya que desde un punto de vista práctico permitirían contrastar si un estado generado en el laboratorio tiene la cantidad de entrelazamiento necesaria para llevar a cabo un determinado protocolo de información cuántica. Las primeras cotas de este tipo se obtuvieron en [186] para la concurrencia de sistemas cuánticos $2 \times N$. Estas ideas se elaboraron en más detalle en [187] para desarrollar una forma de obtener cotas inferiores para la concurrencia para estados bipartitos en dimensiones arbitrarias. Este método proporciona muchas veces cotas restrictivas, capaces incluso de dar una estimación del entrelazamiento de estados entrelazados PPT (para los que la medida computable de la negatividad no sirve).

No obstante, en este capítulo vamos a seguir una línea alternativa que tronca directamente con los resultados provenientes de los criterios de separabilidad. Éstos se suelen formular, como hemos comprobado a lo largo

de esta tesis, mediante una desigualdad que violan los estados entrelazados. Resulta, entonces, bastante intuitivo pensar que a mayor violación de la desigualdad por parte de un estado, mayor será la cantidad de entrelazamiento contenida en él, ya que mayor es su distancia a los estados separables en el sentido de la medida matemática que expresa la desigualdad. Esta idea tomó cuerpo por primera vez en [188] donde se mostró que en estados $M \times N$ (recuérdese que siempre tomamos $M \leq N$)

$$C(\rho) \geq \sqrt{\frac{2}{M(M-1)}} (\Gamma - 1), \quad (7.1)$$

donde

$$\Gamma = \max(\|\rho^{TA}\|_{tr}, \|R(\rho)\|_{tr}), \quad (7.2)$$

de forma que, no sólo se puede deducir la presencia de entrelazamiento a partir de violaciones de los criterios PPT y CCNR, sino que también se pueden utilizar para obtener una estimación inferior de la cantidad de entrelazamiento a través de la concurrencia. Nótese, además, que (7.1) también se puede utilizar para acotar inferiormente el tangle puesto que $\tau(\rho) \geq C^2(\rho)$. Este resultado puede considerarse complementario al de [187] puesto que según los estados uno u otro ofrece las cotas más restrictivas, aunque debe destacarse que (7.1) tiene una expresión matemática mucho más sencilla y manejable. Así mismo, en [189] se extendieron estos resultados para obtener cotas inferiores para el entrelazamiento de formación,

$$E_F(\rho) \geq \text{co}[F(\Gamma)], \quad (7.3)$$

donde

$$F(x) = h(\gamma(x)) + (1 - \gamma(x)) \log(M - 1), \quad (7.4)$$

$$\gamma(x) = \frac{1}{M^2} \left(\sqrt{x} + \sqrt{(M-1)(M-x)} \right), \quad (7.5)$$

h denota la entropía binaria y $\text{co}[f(x)]$ la envoltura convexa de f , o sea, la mayor función convexa menor o igual que f . De hecho, en este caso se conoce que [128, 190]

$$\text{co}[F(\Gamma)] = \begin{cases} 0 & \Gamma = 1 \\ h(\gamma(\Gamma)) + (1 - \gamma(\Gamma)) \log(M - 1) & 1 < \Gamma < \frac{4(M-1)}{M} \\ \frac{\log(M-1)}{M-2} (\Gamma - M) + \log M & \frac{4(M-1)}{M} \leq \Gamma \leq M \end{cases}. \quad (7.6)$$

Posteriormente, también se ha considerado la posibilidad de estimar la cantidad de entrelazamiento utilizando el criterio no operacional de los testigos de entrelazamiento [191, 192, 193] en general, o bien utilizando testigos específicos en [194, 195] (que en este caso particular pueden mejorar las estimaciones de (7.1) y (7.3) para algunos estados de sistemas cuánticos $N \times N$ con $N \geq 4$ y par).

Nuestro propósito en este capítulo es doble. Por un lado, hemos visto que el criterio de LURs y nuestro Teorema 6.1 ofrecen condiciones extraordinariamente potentes para la identificación de entrelazamiento a veces superiores a los criterios PPT y CCNR. Por ello, merece la pena desarrollar resultados similares a los anteriores con estos criterios, de forma que se pueda obtener cotas inferiores más restrictivas que las dadas por (7.1) y (7.3) para la cantidad de entrelazamiento presente en muchos estados. Por otro lado, dejando aparte el problema de la evaluación de medidas de entrelazamiento, resulta interesante investigar la conexión de estos criterios con la cuantificación del entrelazamiento. Por ejemplo, ¿podemos cuantificar este fenómeno utilizando la desviación estándar? Así mismo, hemos visto que el Teorema 6.1 tiene una clara interpretación física ya que muestra que los estados entrelazados pueden tener, considerando una medida particular, una cantidad de correlación inalcanzable para los estados separables cuyo nivel de correlación está acotado superiormente. En consecuencia, queremos explorar aquí la relación entre esta medida de correlación y las medidas de entrelazamiento, así como valorar qué información nos da de la cantidad de entrelazamiento presente en un estado.

7.1. Cuantificación del entrelazamiento y matriz de correlación

La matriz de correlación T contiene información sobre las correlaciones entre los subsistemas de un estado cuántico bipartito. Para medir el tamaño de estas correlaciones resulta natural considerar una norma de T y en la Proposición 6.2 hemos visto que, por consistencia, ésta ha de ser una norma matricial unitariamente invariante. Por otro lado, en el Teorema 6.1 hemos visto que existe una conexión entre la información contenida en la matriz de correlación y la propiedad de entrelazamiento/separabilidad de un estado: la cantidad de correlación, medida a través de cualquier norma matricial de T , no puede sobrepasar una determinada barrera para los estados separables, de forma que un grado mayor de correlación sólo es posible en virtud del entrelazamiento. Como ya aclaramos entonces, escogimos la norma en traza porque era la que ofrecía una condición más poderosa a la hora de detectar estados entrelazados. Vista esta conexión entre matriz de correlación y entrelazamiento, resulta lógico preguntarse si una medida de la correlación expresada a través de una norma de T también puede expresar de alguna forma la cantidad de entrelazamiento de un estado, o, más concretamente si existe una relación directa entre $\|T\|$ y alguna medida de entrelazamiento, con la libertad ahora de que $\|\cdot\|$ sea una norma matricial cualquiera unitariamente invariante. A continuación mostramos que esto es efectivamente así, ya que para estados puros la concurrencia y el tangle tienen una correspondencia uno a uno con, no la norma en traza de T , sino con su norma de Hilbert-Schmidt.

Proposición 7.1. *Para cualquier estado puro ψ $M \times N$ con representación de*

Bloch (6.16) se tiene que

$$\tau(\psi) = C^2(\psi) = \frac{8}{MN(M+N)} \left(\|T\|_{HS}^2 - \frac{MN(M-1)(N-1)}{4} \right). \quad (7.7)$$

Demostración. Recordando la ecuación (6.27) y que la matriz densidad reducida de un estado arbitrario ρ con representación de Bloch (6.16) es $\rho_A = 1/M(I + \sum_i r_i \lambda_i^A)$, se tiene que

$$\text{Tr}(\rho_A^2) = \frac{M + 2\|\mathbf{r}\|_2^2}{M^2}, \quad (7.8)$$

Por tanto, (2.81) y (2.82) se pueden reescribir como

$$\tau(\psi) = C^2(\psi) = \frac{2(M^2 - M - 2\|\mathbf{r}\|_2^2)}{M^2}. \quad (7.9)$$

Veremos que en el caso de estados puros \mathbf{r} , \mathbf{s} y T se relacionan de una forma determinada. Para estos estados las reducciones ρ_A y ρ_B tienen los mismos autovalores y, en consecuencia, $\text{Tr}(\rho_A^2) = \text{Tr}(\rho_B^2)$. Así que, recordando (7.8), esto implica que

$$\frac{M + 2\|\mathbf{r}\|_2^2}{M^2} = \frac{N + 2\|\mathbf{s}\|_2^2}{N^2}. \quad (7.10)$$

Más aún, los estados puros satisfacen $\text{Tr}(\rho^2) = 1$. Utilizando (6.6)-(6.7) y llevando a cabo una cuenta sin complicaciones, se encuentra que esto significa

$$N\|\mathbf{r}\|_2^2 + M\|\mathbf{s}\|_2^2 + 2\|T\|_{HS}^2 = \frac{MN(MN-1)}{2}, \quad (7.11)$$

donde hemos tenido en cuenta (A.3). Las ecuaciones (7.10) y (7.11) conllevan que el valor de uno de los parámetros $\{\|\mathbf{r}\|_2, \|\mathbf{s}\|_2, \|T\|_{HS}\}$ determina unívocamente el valor de los otros para los estados puros. Así, introduciendo (7.10) en (7.11) llegamos a

$$\|\mathbf{r}\|_2^2 = \frac{M}{M+N} \left(\frac{N(M^2-1)}{2} - \frac{2}{N}\|T\|_{HS}^2 \right). \quad (7.12)$$

Por último, combinando esta ecuación con (7.9) se llega al resultado enunciado. \square

Así pues, al igual que $\|T\|_{tr}$, $\|T\|_{HS}$ también caracteriza el entrelazamiento de los estados puros, pero con la añadidura de que también mide rigurosamente su cantidad de entrelazamiento puesto que, al ser $\tau(\psi)$ y $C(\psi)$ funciones monótonas crecientes de este parámetro, tanto tangle y concurrencia como norma de Hilbert-Schmidt de la matriz de correlación son términos intercambiables. En resumen, sorprendentemente quizá, para un estado puro la cantidad de correlación medida a través de $\|T\|_{HS}$ es equivalente a la cantidad de entrelazamiento medida a través del tangle o la concurrencia. Esta observación es, por supuesto, extendible a los estados mezcla en el sentido del techo convexo: la concurrencia y el tangle vienen dados por el techo convexo sobre descomposiciones en estados puros de los valores que éstos toman para

$\|T\|_{HS}$. Sin embargo, como resultará obvio dada la complejidad computacional de evaluación de estas medidas, en el caso general de estados mezcla no es posible expresar el tangle o la concurrencia como función de esta norma de la matriz de correlación del estado. No obstante, sigue existiendo una relación entre la cantidad de entrelazamiento del estado y $\|T\|_{HS}$, y, de hecho, ésta ofrece una cota inferior para la estimación del tangle.

Teorema 7.1. *Para cualquier estado ρ $M \times N$ con representación de Bloch (6.16) se tiene que*

$$\tau(\rho) \geq \frac{8}{MN(M+N)} \left(\|T\|_{HS}^2 - \frac{MN(M-1)(N-1)}{4} \right). \quad (7.13)$$

Demostración. Sea $\sum_n p_n |\psi_n\rangle\langle\psi_n|$ la descomposición óptima de ρ en (2.85) para la cual se obtiene el valor del tangle. Entonces,

$$\begin{aligned} \tau(\rho) &= \sum_n p_n \tau(\psi_n) \\ &= \frac{8}{MN(M+N)} \left(\sum_n p_n \|T_{\psi_n}\|_{HS}^2 - \frac{MN(M-1)(N-1)}{4} \right) \\ &\geq \frac{8}{MN(M+N)} \left(\left(\sum_n p_n \|T_{\psi_n}\|_{HS} \right)^2 - \frac{MN(M-1)(N-1)}{4} \right) \\ &\geq \frac{8}{MN(M+N)} \left(\left\| \sum_n p_n T_{\psi_n} \right\|_{HS}^2 - \frac{MN(M-1)(N-1)}{4} \right) \\ &= \frac{8}{MN(M+N)} \left(\|T_\rho\|_{HS}^2 - \frac{MN(M-1)(N-1)}{4} \right), \end{aligned}$$

donde en la primera desigualdad hemos utilizado la convexidad de la función $f(x) = x^2$ y en la segunda desigualdad la convexidad de $\|\cdot\|_{HS}$ (i.e., la desigualdad triangular). \square

Aparte del interés teórico de esta desigualdad (de la que daremos una aplicación más adelante), cabe preguntarse acerca de su capacidad para la estimación del tangle en la práctica. Nótese que (7.13) implica el siguiente criterio de separabilidad: todos los estados separables cumplen que $\|T\|_{HS} \leq \sqrt{MN(M-1)(N-1)}/2$. Sin embargo, esta condición es más débil que el Teorema 6.1 y puede ser deducida trivialmente a partir de él ya que $\|T\|_{tr} \geq \|T\|_{HS}$. De hecho, la práctica muestra que esta condición es así mismo, en general, más débil que los criterios PPT y CCNR. Por tanto, el Teorema 7.1 no dará lugar a nuevas cotas no triviales donde (7.1), y las nuevas desigualdades que desarrollemos más adelante en esta sección, fallan. Al revés, éstas últimas ponen cotas no triviales donde (7.13) falla. Sin embargo, a diferencia de los anteriores resultados, la nueva cota se corresponde con el valor exacto del tangle para estados puros, por lo que, aunque más débil para la detección

de entrelazamiento, es más ajustada al valor real del tangle para estados de alta pureza. Esta situación es común en los experimentos cuando el estado puro entrelazado ψ que se intenta preparar es contaminado por diferentes tipos de ruido. Como resultado, se obtiene un estado débilmente mezclado: $\rho = p|\psi\rangle\langle\psi| + (1-p)\rho_{\text{noise}}$ con p cercano a 1. Para comprobar la fortaleza de la cota (7.13), hemos considerado mezclas en dimensiones 3×3 de estados puros entrelazados arbitrarios y ruido blanco (i.e. $\rho_{\text{noise}} = I_9/9$). Hemos encontrado que, en general, la cota del Teorema 7.1 mejora las estimaciones de (7.1) cuando $p \gtrsim 0,94$.

Como prometimos, mostraremos otra aplicación del Teorema 7.1 demostrando una relación entre medidas de entrelazamiento previamente conjeturada recientemente. La medida basada en la norma minkowskiana (MNB) es una medida de entrelazamiento para estados de dos qubits que se define como [196]

$$E_{MNB}(\rho) = \text{máx}\{\text{Tr}\rho^2 - 1 + \text{Tr}\rho(\sigma_y \otimes \sigma_y)\rho^*(\sigma_y \otimes \sigma_y), 0\}, \quad (7.14)$$

donde la conjugación de la matriz densidad se toma, como es habitual, en la base de autoestados de σ_z . Aunque $E_{MNB}(\rho) = 0$ para muchos estados entrelazados, esta medida puede ser calculada analíticamente (de hecho, ni siquiera se necesita diagonalizar una matriz) y se ha demostrado que cumple con el requisito de no crecimiento bajo LOCC en [197]. En este mismo artículo se muestra que $E_{MNB}(\rho)$ acota inferiormente a $C(\rho)$ para algunos estados. Basándose en la evidencia numérica los autores sugieren que esto podría ser cierto para todos los estados. Efectivamente, mostramos a continuación que esta conjetura resulta ser cierta. Utilizando la ecuación (5) en [197] y la (A.3) de esta memoria, se deduce claramente que la medida MNB se puede escribir en términos de la matriz de correlación como

$$E_{MNB}(\rho) = \text{máx}\left\{\frac{1}{2}(\|T\|_{HS}^2 - 1), 0\right\}, \quad (7.15)$$

que es precisamente la cota para el tangle obtenida en el Teorema 7.1 para el caso de estados de dos qubits. Por tanto, se tiene que $\tau(\rho) \geq E_{MNB}(\rho)$. Por otro lado, puesto que para dos qubits se satisface $0 \leq C(\rho) \leq 1$, se cumple que $C(\rho) \geq C^2(\rho)$. Ahora, recordando que en este caso $\tau(\rho) = C^2(\rho)$ llegamos al resultado deseado: $C(\rho) \geq E_{MNB}(\rho)$. Así, se observa que la medida de entrelazamiento MNB está directamente relacionada con la matriz de correlación y que acota inferiormente tanto a la concurrencia como al tangle.

Por último, merece la pena mencionarse que es posible obtener cotas superiores para el tangle en términos de $\|\mathbf{r}\|_2$ y $\|\mathbf{s}\|_2$ usando un razonamiento análogo al del Teorema 7.1 pero comenzando a partir de (7.9) (o su equivalente en función de N y $\|\mathbf{s}\|_2$),

$$\tau(\rho) \leq 2 \text{mín} \left\{ \frac{M^2 - M - 2\|\mathbf{r}\|_2^2}{M^2}, \frac{N^2 - N - 2\|\mathbf{s}\|_2^2}{N^2} \right\}. \quad (7.16)$$

Hasta aquí hemos encontrado interesantes conexiones entre el grado de correlación de un estado medido a través de la norma de Hilbert-Schmidt de

la matriz de correlación y su cantidad de entrelazamiento medida a través de la concurrencia y/o el tangle, de las que, desde el punto de vista práctico, destacan las agudas estimaciones de estas medidas que se obtienen para estados de elevada pureza. Sin embargo, como hemos analizado, el uso de $\|T\|_{HS}$ lleva asociado un criterio de separabilidad relativamente débil, lo que limita la utilidad de estas estimaciones fuera de los estados cercanos a los puros. Por ello, resulta lógico plantearse la posibilidad de obtener relaciones similares con $\|T\|_{tr}$, beneficiándose de la potencia del Teorema 6.1 para estados generales. Como mostramos a continuación esto es posible a pesar de no existir en este caso una correspondencia uno a uno entre $\|T\|_{tr}$ y las medidas de entrelazamiento de estados puros a diferencia de lo que sucedía con $\|T\|_{HS}$ en la Proposición 7.1.

Teorema 7.2. *Para cualquier estado ρ $M \times N$ con representación de Bloch (6.16) se tiene que*

$$C(\rho) \geq \sqrt{\frac{8}{M^3 N^2 (M-1)}} \left(\|T\|_{tr} - \sqrt{\frac{MN(M-1)(N-1)}{4}} \right). \quad (7.17)$$

Demostración. Comencemos relacionando la $\|T_\psi\|_{tr}$ de un estado puro con sus coeficientes de Schmidt. Tenga el estado puro ψ la siguiente descomposición de Schmidt (véase (2.10)),

$$|\psi\rangle = \sum_{j=0}^{M-1} \sqrt{\mu_j} |j_A j_B\rangle. \quad (7.18)$$

Descompongamos su correspondiente matriz densidad como $\rho_\psi = |\psi\rangle\langle\psi| = \rho^{sep} + \epsilon$, donde

$$\rho^{sep} = \sum_j \mu_j |j_A j_B\rangle\langle j_A j_B|, \quad \epsilon = \sum_{j \neq k} \sqrt{\mu_j \mu_k} |j_A j_B\rangle\langle k_A k_B|. \quad (7.19)$$

Nótese que ρ^{sep} es, por tanto, un estado separable. Utilizando (6.10)-(6.11) es directo verificar que ϵ se puede escribir en términos de los generadores de $SU(M)$ y $SU(N)$ no diagonales (construidos a partir de la base de Schmidt),

$$\epsilon = \frac{1}{2} \sum_{j < k} \sqrt{\mu_j \mu_k} (u_{jk} \otimes u_{jk} - v_{jk} \otimes v_{jk}), \quad (7.20)$$

mientras que, al ser ρ^{sep} diagonal en esta base, su representación de Bloch sólo contiene términos correspondientes a las identidades y a los generadores diagonales w_l . Esto significa que la matriz de correlación de ρ_ψ es diagonal por bloques (no se mezclan los w 's con los u 's y v 's) y, por tanto,

$$\begin{aligned} \|T_\psi\|_{tr} &= \|T_{\rho^{sep}}\|_{tr} + MN \sum_{j < k} \sqrt{\mu_j \mu_k} \\ &\leq \sqrt{\frac{MN(M-1)(N-1)}{4}} + MN \sum_{j < k} \sqrt{\mu_j \mu_k}, \end{aligned} \quad (7.21)$$

donde hemos utilizado el Teorema 6.1. Ahora podemos relacionar $\|T_\psi\|_{tr}$ con $C(\psi)$ puesto que esta es fácilmente expresable en términos de los coeficientes de Schmidt (recuérdese que los coeficientes de Schmidt son los autovalores de las matrices densidad reducidas ρ_A y ρ_B),

$$C^2(\psi_n) = 2 \left(1 - \sum_j \mu_j^2 \right) = 4 \sum_{j < k} \mu_j \mu_k. \quad (7.22)$$

Sin más que tener en cuenta la desigualdad [188]

$$\sum_{j < k} \mu_j \mu_k \geq \frac{2}{M(M-1)} \left(\sum_{j < k} \sqrt{\mu_j \mu_k} \right)^2, \quad (7.23)$$

se obtiene entonces que

$$C(\psi) \geq \sqrt{\frac{8}{M^3 N^2 (M-1)}} \left(\|T_\psi\|_{tr} - \sqrt{\frac{MN(M-1)(N-1)}{4}} \right). \quad (7.24)$$

Una vez establecida la relación para estados puros extenderla al caso general es análogo al caso del Teorema 7.1. Sea $\sum_n p_n |\psi_n\rangle\langle\psi_n|$ la descomposición de ρ para la cual $C(\rho) = \sum_n p_n C(\psi_n)$. Entonces, utilizando la desigualdad (7.24) para todos los ψ_n conjuntamente con la desigualdad triangular ($\|\sum_n p_n T_{\psi_n}\|_{tr} = \|\sum_n p_n T_{\psi_n}\|_{tr} \leq \sum_n p_n \|T_{\psi_n}\|_{tr}$) se llega al resultado final. \square

Es este un resultado completamente análogo a (7.1), que establece que la determinación de $\|T\|_{tr}$ no sólo sirve para detectar la presencia de entrelazamiento a través del Teorema 6.1, sino que también permite estimar la cantidad de entrelazamiento presente mediante una cota inferior a la concurrencia (o al tangle ya que $\tau \geq C^2$). Así pues, la norma en traza de la matriz de correlación también contiene información de la cantidad de entrelazamiento. Queda por estudiar la agudeza de la cota dada por (7.17), especialmente en comparación con (7.1). Por un lado, dada la potencia del Teorema 6.1 estudiada en el anterior capítulo, el Teorema 7.2 es capaz de fijar cotas no triviales para multitud de estados y, además, para algunos de ellos como los estados isotrópicos o los Bell-diagonales de dos qubits reproduce el valor exacto de la concurrencia (lo cual se relaciona con el hecho de que el Teorema 6.1 detecta óptimamente el entrelazamiento de estos estados como se muestra en el Ejemplo 4 y el Corolario 6.1). Por otro lado, debido a la Proposición 6.5, el Teorema 7.2 no puede mejorar la estimación dada por $\|R(\rho)\|_{tr}$ en (7.1) cuando $M = N$, por lo que la única posibilidad de mejorar estrictamente las cotas anteriores es en el caso de subsistemas con dimensiones distintas. Desgraciadamente, nuestras exploraciones muestran que, en general, incluso en este caso (7.1) suele dar cotas más restrictivas que (7.17). Esto se debe a que $\|T_\psi\|_{tr}$ se relaciona con los coeficientes de Schmidt mediante una desigualdad (ver (7.21)) mientras que este tipo de relación para $\|\rho^{TA}\|_{tr}$ y $\|R(\rho)\|_{tr}$ es de igualdad (véase [188]). Así, en general, para estados que detectan los tres

criterios se encuentra que $\|\rho^{TA}\|_{tr}$ y $\|R(\rho)\|_{tr}$ dan mejores estimaciones de la concurrencia que $\|T\|_{tr}$. Sin embargo, esto no es una regla general y se pueden presentar excepciones ya que sabemos que existen situaciones en las que el Teorema 6.1 es más fuerte que los criterios PPT y CCNR conjuntamente. Obviamente para este tipo de estados la cota para la concurrencia dada por la matriz de correlación será la mejor. Así, por ejemplo, si consideramos el estado $\tilde{\rho}_{GT2}$ del capítulo anterior, mientras que (7.1) nos dice que $C(\tilde{\rho}_{GT2}) \geq 0,0296$, tenemos que $C(\tilde{\rho}_{GT2}) \geq 0,0320$ por el Teorema 7.2.

Hasta aquí hemos considerado un conocimiento absoluto de las correlaciones que constituyen la matriz de correlación. Sin embargo, aunque esto no es tan costoso experimentalmente como determinar la matriz densidad al completo, siempre es de agradecer un alivio en el esfuerzo experimental. De hecho, en [198] se ha estudiado recientemente cómo extraer información acerca del entrelazamiento de un sistema a partir de un conocimiento limitado de las correlaciones para minimizar el coste experimental. Estas ideas también se pueden aplicar a nuestros resultados debido a que es posible estimar la norma en traza de una matriz a partir del conocimiento de unas pocas de sus entradas. Por ejemplo, del Lema 6.1 se deduce que $\|A\|_{tr} \geq \sum_i |a_{ii}|$ para cualquier matriz A , lo que permite obtener información (cuantitativa) sobre el entrelazamiento simplemente a través de la medida de las correlaciones correspondientes a las entradas diagonales de la matriz de correlación ($M^2 - 1$ medidas), i.e.

$$C(\rho) \geq \sqrt{\frac{8}{M^3 N^2 (M-1)}} \left(\sum_i |t_{ii}| - \sqrt{\frac{MN(M-1)(N-1)}{4}} \right). \quad (7.25)$$

Por último, reproduciendo las ideas de [189] también es posible obtener una cota para el entrelazamiento de formación.

Teorema 7.3. *Para cualquier estado ρ $M \times N$ con representación de Bloch (6.16) se tiene que*

$$E_F(\rho) \geq \text{co} \left\{ F \left[\frac{2}{MN} \left(\|T\|_{tr} - \sqrt{\frac{MN(M-1)(N-1)}{4}} \right) + 1 \right] \right\}, \quad (7.26)$$

donde $\text{co}[F(x)]$ está dado por (7.6).

Demostración. Nótese que para un estado puro ψ con descomposición de Schmidt (7.18) se tiene que $E_F(\psi) = S(\rho_A) = H(\{\mu_j\})$. En [128] se resuelve el siguiente problema de minimización

$$\min_{\{\mu_j\}} \left\{ H(\{\mu_j\}) : 2 \sum_{j < k} \sqrt{\mu_j \mu_k} + 1 = \lambda \right\} = F(\lambda) \quad (7.27)$$

y se demuestra que $\text{co}[F(\lambda)] \leq H(\{\mu_j\})$ y que $\text{co}[F(\lambda)]$ es monótona creciente (y, por supuesto, convexa). Así, sea $\sum_n p_n |\psi_n\rangle\langle\psi_n|$ la descomposición óptima

de ρ que da el valor del entrelazamiento de formación. Entonces, teniendo en cuenta las propiedades que acabamos de mencionar,

$$\begin{aligned} E_F(\rho) &= \sum_n p_n E_F(\psi_n) = \sum_n p_n H(\{\mu_j^{(n)}\}) \geq \sum_n p_n \text{co}\{F(\lambda^{(n)})\} \\ &\geq \text{co}\{F(\sum_n p_n \lambda^{(n)})\}, \end{aligned} \quad (7.28)$$

considerando además la desigualdad (7.21),

$$E_F(\rho) \geq \text{co} \left\{ F \left[\frac{2}{MN} \left(\sum_n p_n \|T_{\psi_n}\|_{tr} - \sqrt{\frac{MN(M-1)(N-1)}{4}} \right) + 1 \right] \right\} \quad (7.29)$$

y, por último, la desigualdad triangular lleva al resultado que queríamos probar. \square

7.2. Cuantificación del entrelazamiento y relaciones de incertidumbre locales

En la sección previa hemos encontrado interesantes conexiones entre la matriz de correlación y las medidas de entrelazamiento que, entre otras cosas, han permitido obtener cotas inferiores para la estimación de éstas de forma análoga a resultados anteriores que involucraban los criterios PPT y CCNR. Resulta natural preguntarse si es posible obtener un resultado similar para el criterio de las LURs. En primer lugar, merece la pena investigar la posibilidad de cuantificar el entrelazamiento mediante relaciones de incertidumbre locales. De hecho, se ha venido especulando con esta posibilidad desde que se propuso este criterio por primera vez en [111] ya que se comprobó que la violación de la desigualdad (2.66) para los estados de Werner de dos qubits se corresponde exactamente con el valor de su concurrencia, lo que ha dado lugar a una interesante literatura que estudia la adecuación de las LURs para la cuantificación del entrelazamiento [199, 157, 185]. Aquí aclararemos esta situación mostrando que lo que implica una determinada violación de un criterio LUR es una determinada cantidad mínima de entrelazamiento en términos de la concurrencia de forma análoga a los criterios considerados anteriormente. En segundo lugar, desde un punto de vista práctico, esto permite derivar cotas inferiores para la estimación de medidas de entrelazamiento. Además, dada la excepcional fortaleza de este criterio (que recordamos es estrictamente mejor que el CCNR), es de esperar que aporte en muchos casos estimaciones más precisas (i.e., cotas inferiores más restrictivas).

Puesto que la formulación del criterio LUR en términos de un conjunto de LOOs (2.69) es, sin duda, la más fuerte y general, lo más razonable es atenerse a este formalismo para la consecución de nuestro objetivo. Para comodidad del lector reproducimos (2.69) aquí; este criterio establece que si $\rho \in \mathcal{D}(H_A \otimes H_B)$

entonces

$$\sum_{i=1}^{N^2} \Delta_{\rho}^2(G_i^A \otimes I_B + I_A \otimes G_i^B) \geq M + N - 2$$

donde $\{G_i^A\}$ y $\{G_i^B\}$ son conjuntos de LOOs, es decir, bases ortonormales (con respecto al producto escalar Hilbert-Schmidt (2.3)) de $\mathcal{B}(H_A)$ y $\mathcal{B}(H_B)$ respectivamente. Recordamos también que si $M < N$, se entiende que $G_i^A = 0$ para $M^2 + 1 \leq i \leq N^2$. Un ejemplo de LOOs es la identidad y los generadores de $SU(\dim H)$ (debidamente normalizados), que ya hemos considerado con anterioridad. Otro ejemplo, que denominaremos conjunto estándar de LOOs, es $\{G_i^A\} = \{g_j, g_{jk}^+, g_{jk}^-\}$, donde

$$\begin{aligned} g_j &= |j\rangle\langle j| \quad (0 \leq j \leq M-1), \\ g_{jk}^+ &= \frac{1}{\sqrt{2}}(|j\rangle\langle k| + |k\rangle\langle j|) \quad (0 \leq j < k \leq M-1), \\ g_{jk}^- &= -\frac{i}{\sqrt{2}}(|j\rangle\langle k| - |k\rangle\langle j|) \quad (0 \leq j < k \leq M-1), \end{aligned} \quad (7.30)$$

y de forma similar para $\{G_i^B\}$. En nuestro objetivo de relacionar (2.69) con la concurrencia comenzamos con un lema en el que nos restringimos a estados puros, y que, luego, utilizaremos para demostrar la relación buscada en el caso general.

Lema 7.1. *Para cualquier conjunto de LOOs $\{G_i^A\}$ y $\{G_i^B\}$ y cualquier estado puro $M \times N$ ψ con descomposición de Schmidt $|\psi\rangle = \sum_{j=0}^{M-1} \sqrt{\mu_j} |j_A j_B\rangle$, se tiene que*

$$\sum_{i=1}^{N^2} \Delta_{\psi}^2(G_i^A \otimes I_B + I_A \otimes G_i^B) \geq M + N - 2 - 4 \sum_{j < k} \sqrt{\mu_j \mu_k}. \quad (7.31)$$

La cota se alcanza para $\{G_i^A\} = \{g_j, g_{jk}^+, g_{jk}^-\}$ y $\{G_i^B\} = \{-g_j, -g_{jk}^+, g_{jk}^-\}$ (construidos a partir de la correspondiente base de Schmidt).

Demostración. Recordando (6.1) tenemos que

$$\sum_i \Delta_{\psi}^2(G_i^A \otimes I_B + I_A \otimes G_i^B) = \sum_i (\Delta_{\rho_A}^2(G_i^A) + \Delta_{\rho_B}^2(G_i^B)) + 2 \sum_i \kappa_{\psi}(G_i^A, G_i^B). \quad (7.32)$$

Como en la demostración del Teorema 7.2, reescribimos $\rho_{\psi} = |\psi\rangle\langle\psi| = \rho^{sep} + \epsilon$, donde ρ^{sep} y ϵ vienen dados por (7.19). Nótese que, además de ser ρ^{sep} separable, sus reducciones son las mismas que las de ρ_{ψ} ($\rho_A^{sep} = \rho_A$, $\rho_B^{sep} = \rho_B$), por lo que

$$\sum_i \kappa_{\psi}(G_i^A, G_i^B) = \sum_i \text{Tr}(\epsilon G_i^A \otimes G_i^B) + \sum_i \kappa_{\rho^{sep}}(G_i^A, G_i^B). \quad (7.33)$$

La separabilidad de ρ^{sep} implica que

$$2 \sum_i \kappa_{\rho^{sep}}(G_i^A, G_i^B) \geq M + N - 2 - \sum_i (\Delta_{\rho_A}^2(G_i^A) + \Delta_{\rho_B}^2(G_i^B)) \quad (7.34)$$

y así

$$\sum_i \Delta_\psi^2(G_i^A \otimes I_B + I_A \otimes G_i^B) \geq M + N - 2 + 2 \sum_i \text{Tr}(\epsilon G_i^A \otimes G_i^B), \quad (7.35)$$

de manera que resta comprobar que $\sum_i \text{Tr}(\epsilon G_i^A \otimes G_i^B) \geq -2 \sum_{j < k} \sqrt{\mu_j \mu_k}$. Para ello, nótese que

$$\begin{aligned} \sum_i \text{Tr}(\epsilon G_i^A \otimes G_i^B) &= \sum_i \sum_{j \neq k} \sqrt{\mu_j \mu_k} \langle j_A | G_i^A | k_A \rangle \langle j_B | G_i^B | k_B \rangle \\ &\geq - \sum_i \sum_{j \neq k} \sqrt{\mu_j \mu_k} |\langle j_A | G_i^A | k_A \rangle \langle j_B | G_i^B | k_B \rangle| \\ &\geq - \sum_i \sum_{j \neq k} \sqrt{\mu_j \mu_k} (|\langle j_A | G_i^A | k_A \rangle|^2 + |\langle j_B | G_i^B | k_B \rangle|^2) / 2, \end{aligned} \quad (7.36)$$

donde en el último paso hemos utilizado que $a^2 + b^2 \geq 2|ab|$. Ahora, el resultado se sigue porque $\sum_i |\langle j_A | G_i^A | k_A \rangle|^2 = \sum_i |\langle j_B | G_i^B | k_B \rangle|^2 = 1 \forall j, k$ para cualquier conjunto de LOOs $\{G_i^A\}$ y $\{G_i^B\}$. Para ver esto, considérese que $\mathcal{B}(H_A)$ es isomorfo a \mathbb{C}^{M^2} con el producto escalar estándar, de forma que escribiendo los $\{G_i^A\}$ en coordenadas de la base estándar de LOOs (construida a partir de la correspondiente base de Schmidt) se obtienen vectores columna que dan una base ortonormal de \mathbb{C}^{M^2} . Estos vectores columna se pueden unir para dar lugar a la matriz unitaria U , y $\sum_i |\langle j_A | G_i^A | k_A \rangle|^2$ se corresponde con sumar el módulo al cuadrado de los elementos de una determinada fila de U , lo que es igual a 1. Obviamente, el mismo razonamiento sirve para $\sum_i |\langle j_B | G_i^B | k_B \rangle|^2$. Falta comprobar que la cota se alcanza para el conjunto de LOOs mencionado. Utilizando que $\sum_i (G_i^A)^2 = M I_A$ y $\sum_i (G_i^B)^2 = N I_B$ [113], es directo encontrar que

$$\begin{aligned} \sum_{i=1}^{N^2} \Delta_\psi^2(G_i^A \otimes I_B + I_A \otimes G_i^B) &= M + N + 2 \sum_i \langle G_i^A \otimes G_i^B \rangle_\psi \\ &\quad - \sum_i \langle G_i^A \otimes I_B + I_A \otimes G_i^B \rangle_\psi^2. \end{aligned} \quad (7.37)$$

Tomando en cuenta que ρ_ψ alcanza su descomposición de Schmidt en el espacio de operadores para los LOOs estándar, i.e.

$$\rho_\psi = \sum_j \mu_j g_j \otimes g_j + \sum_{j < k} \sqrt{\mu_j \mu_k} (g_{jk}^+ \otimes g_{jk}^+ + g_{jk}^- \otimes (-g_{jk}^-)), \quad (7.38)$$

se comprueba que

$$\sum_i \langle G_i^A \otimes G_i^B \rangle_\psi = -1 - 2 \sum_{j < k} \sqrt{\mu_j \mu_k}, \quad \langle G_i^A \otimes I_B + I_A \otimes G_i^B \rangle_\psi = 0 \forall i \quad (7.39)$$

para el conjunto de LOOs dado en el enunciado del lema. \square

Teorema 7.4. Para cualquier estado ρ $M \times N$ y cualquier conjunto de LOOs $\{G_i^A\}$ y $\{G_i^B\}$, se tiene que

$$C(\rho) \geq \frac{M + N - 2 - \sum_i \Delta_\rho^2(G_i^A \otimes I_B + I_A \otimes G_i^B)}{\sqrt{2M(M-1)}}. \quad (7.40)$$

Demostración. Sea $\sum_n p_n |\psi_n\rangle\langle\psi_n|$ la descomposición de ρ para la cual se obtiene el mínimo en (2.84), de forma que $C(\rho) = \sum_n p_n C(\psi_n)$. Teniendo en cuenta (7.22) y (7.23) se halla que

$$C(\psi_n) \geq \sqrt{2/(M(M-1))} 2 \sum_{j < k} \sqrt{\mu_j \mu_k}. \quad (7.41)$$

El uso del Lema 7.1 y de que $\Delta_\rho^2(M) \geq \sum_n p_n \Delta_{\psi_n}^2(M)$ para cualquier observable M (ver por ejemplo [111]) conduce entonces al resultado deseado. \square

Este teorema es pues el análogo de (7.1) y del Teorema 7.2 y establece que la violación del criterio LURs no sólo da información cualitativa del entrelazamiento sino también cuantitativa. Dada la potencia de este criterio, (7.40) da lugar a cotas no triviales para la concurrencia de un enorme número de estados siendo para algunos óptima (i.e., reproduce el valor exacto de la concurrencia). Este es el caso, por ejemplo, de los estados isotrópicos (véase el Ejemplo 4 de la sección 6.3.1) al utilizar los LOOs $\{G_i^A\} = \{I/\sqrt{M}, w_l/\sqrt{2}, u_{jk}/\sqrt{2}, v_{jk}/\sqrt{2}\}$ y $\{G_i^B\} = \{-I/\sqrt{N}, -w_l/\sqrt{2}, -u_{jk}/\sqrt{2}, v_{jk}/\sqrt{2}\}$, lo que explica la coincidencia de concurrencia y violación de (2.66) apuntada en [111]. En lo que se refiere a la fortaleza del Teorema 7.4 para establecer cotas para la concurrencia en comparación con los resultados anteriormente considerados, éste adolece de la desventaja de la falta de operacionalidad, ya que no está claro cuál es la elección de observables más apropiada. Sin embargo, como mostramos a continuación, una elección intuitiva basta para mejorar las predicciones de todas las desigualdades anteriores en muchos casos. Para no abrumar al lector nos limitamos a reseñar dos ejemplos significativos.

Ejemplo 1: Consideremos el estado en dimensiones 2×3

$$\rho = p|\Psi\rangle\langle\Psi| + (1-p)|01\rangle\langle 01|, \quad |\Psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}. \quad (7.42)$$

La figura 7.1 muestra que el Teorema 7.4 da una cota más restrictiva para la concurrencia que (7.1). Este hecho es remarcable, al mostrar que el criterio de las LURs ofrece una estimación más poderosa de la cantidad de entrelazamiento que el criterio PPT a pesar de que este último caracteriza óptimamente el entrelazamiento en estas dimensiones.

Ejemplo 2: Consideremos el siempre esquivo caso de los estados PPT entrelazados. Sea ρ el estado 3×3 -dimensional de (6.40) estudiado en el Ejemplo 1 de la sección 6.3.1. En este caso la cota dada por (7.1) es no trivial únicamente

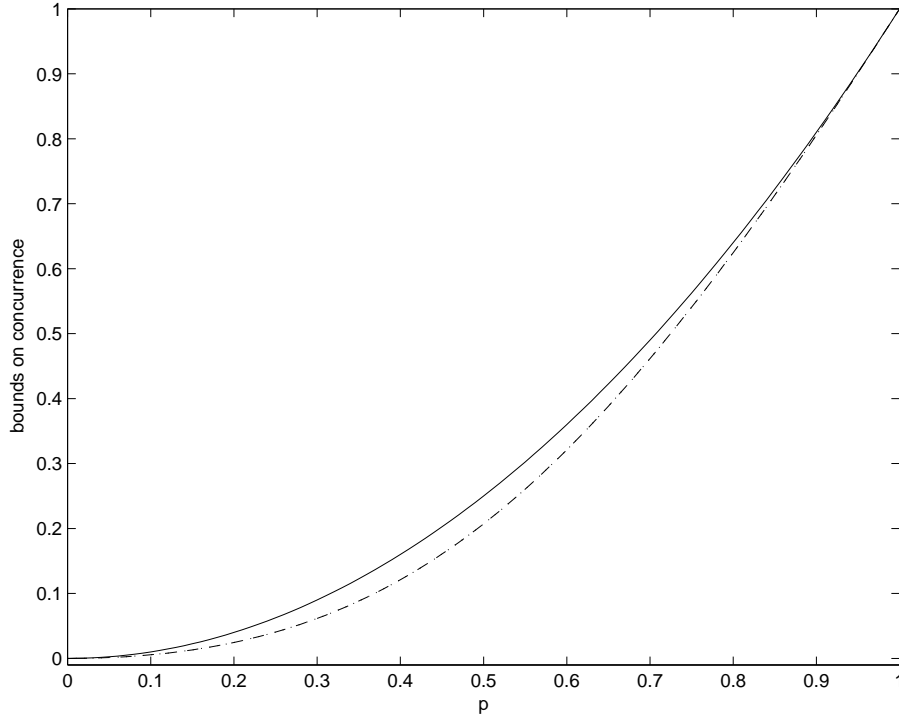


Figura 7.1: Cotas inferiores para la concurrencia del estado ϱ . Línea sólida: Cota dada por el Teorema 7.4 para los LOOs que alcanzan la cota de (7.31) para el estado Ψ . Línea punteada: Cota dada por el criterio PPT. Línea discontinua: Cota dada por el criterio CCNR. (Las dos últimas cotas se solapan). Nótese que debido a la convexidad de la concurrencia se tiene que $C(\varrho) \leq p$.

a través de la estimación dada por $\|R(\rho)\|_{tr}$, obteniéndose $C(\rho) \geq 0,050$. La matriz de correlación vía el Teorema 7.2 nos dice que $C(\rho) \geq 0,021$. Sin embargo, con los LOOs utilizados en [113] para mejorar la detección de ρ mezclado con ruido blanco se obtiene que $C(\rho) \geq 0,052$.

Por último, señalar que, como en la sección anterior, estos resultados son extendibles al entrelazamiento de formación.

Teorema 7.5. *Para cualquier estado ρ $M \times N$ y cualquier conjunto de LOOs $\{G_i^A\}$ y $\{G_i^B\}$ se tiene que*

$$E_F(\rho) \geq co \left\{ F \left[M + N - 1 - \sum_i \Delta_\rho^2(G_i^A \otimes I_B + I_A \otimes G_i^B) \right] \right\}, \quad (7.43)$$

donde $co[F(x)]$ está dado por (7.6).

Demostración. Repítanse los pasos de la demostración del Teorema 7.3 hasta (7.28). Úsese ahora el Lema 7.1 para comprobar que

$$E_F(\rho) \geq \text{co} \left\{ F \left[M + N - 1 - \sum_n p_n \sum_i \Delta_{\psi_n}^2(G_i^A \otimes I_B + I_A \otimes G_i^B) \right] \right\} \quad (7.44)$$

y, por último, vuelva a tenerse en cuenta que $\Delta_\rho^2(M) \geq \sum_n p_n \Delta_{\psi_n}^2(M)$ para cualquier observable M . \square

Parte IV

Conclusiones y perspectivas

En esta tesis hemos considerado cómo las diferentes medidas matemáticas que se conocen para medir la incertidumbre asociada a una distribución de probabilidad pueden ayudar en el estudio de dos de los fenómenos más característicos de la Mecánica Cuántica: el principio de incertidumbre y el entrelazamiento, que, como hemos visto, están interrelacionados. Más en concreto, hemos abordado problemas matemáticos relacionados con la cuantificación rigurosa de la incertidumbre en sistemas atómicos, la derivación de relaciones de incertidumbre rigurosas y restrictivas y la caracterización cualitativa y cuantitativa del entrelazamiento utilizando éstas y otras herramientas relacionadas. A continuación, resumimos la contribución de esta tesis a estos problemas así como las perspectivas abiertas para una investigación futura.

- En el capítulo 3, motivados por su conexión con la expresión de la incertidumbre en sistemas atómicos y moleculares, hemos abordado el problema matemático del cálculo de las entropías de polinomios ortogonales en el caso concreto de los polinomios de Gegenbauer. Utilizando una nueva estrategia basada en el uso de representaciones trigonométricas de los polinomios e integración en el plano complejo, hemos obtenido de una forma unificada fórmulas cerradas para $E(C_n^{(\lambda)})$ cuando $\lambda = 1, 2, 3$, siendo el resultado en este último caso nuevo. Más aún, cuando $\lambda \geq 4$, $\lambda \in \mathbb{N}$, hemos obtenido una expresión completamente analítica para la entropía en términos de sumas finitas, lo que lleva fácilmente a valores exactos de la entropía mediante computación simbólica. El aumento de la complejidad de esta fórmula cuando λ crece (ver (3.97)) ilustra claramente la dificultad de expresar en forma cerrada las entropías de polinomios ortogonales, por lo que creemos que el uso computacional de (3.97) en conjunción con el acercamiento numérico de [64] pueden ser de gran utilidad en este problema. En este mismo sentido, resulta natural preguntarse por la posibilidad de extender nuestras ideas al caso de polinomios de Gegenbauer de parámetro no entero. En este caso, todos nuestros desarrollos siguen siendo válidos (recuérdese el Teorema 3.2) con la particularidad de que la representación de Szegő (3.27) contiene infinitos sumandos, lo cual heredarían las expresiones (3.35) y (3.41) para la integral entrópica $E(C_n^{(\lambda)})$. Queda como un interesante problema abierto para investigación futura el estudio del comportamiento y convergencia de estas series. Recuérdese que sería particularmente interesante complementar los resultados obtenidos para polinomios de Gegenbauer de parámetro entero con los de parámetro semientero ya que, como ya hemos comentado, esto permitiría obtener la entropía de los armónicos hiperesféricos, responsables de la parte angular de todos los sistemas cuánticos sometidos a potenciales centrales.

Así mismo, hemos extendido nuestro método al caso más general de los polinomios de Jacobi, lo que conduce a fórmulas cerradas para $E(P_n^{(1/2, 3/2)})$ y $E(P_n^{(1/2, 5/2)})$, si bien éstas, como ya mencionamos, también

pueden obtenerse por otros medios [145]. De igual forma, se ha podido expresar la entropía de estos polinomios en el caso de parámetros semienteros en términos de sumas finitas, para cuya utilidad y relevancia se aplica lo ya comentado en el caso Gegenbauer.

- En el capítulo 4 se investiga el problema de obtener cotas agudas para las relaciones de incertidumbre entrópicas para pares de observables de un espacio de Hilbert finito-dimensional. Partiendo de la desigualdad de Landau-Pollak, que, a pesar de padecer algunos defectos como relación de incertidumbre, tiene una fortaleza destacable, se ha obtenido una nueva cota para una EUR, que supera a la mejor cota conocida hasta ahora (cota de Maassen-Uffink (2.39)) para el considerable conjunto de observables caracterizado por (4.1). Además de este resultado, como consecuencia del desarrollo llevado a cabo en este capítulo, se han caracterizado las condiciones que fijan cuando la relación de incertidumbre de Landau-Pollak es más fuerte que la de Maassen-Uffink y viceversa ($c \geq 1/\sqrt{2}$ y $c \leq 1/\sqrt{2}$, respectivamente). Adicionalmente, se ha obtenido que la condición de Landau-Pollak es óptima en el caso bidimensional. Es notable que una desigualdad basada en una medida de la incertidumbre tan simple, que ignora todos los valores de una distribución de probabilidad excepto uno, exhiba esta fortaleza cuando $c \geq 1/\sqrt{2}$. No sorprende, por contra, su mayor debilidad en el caso contrario, ya que la desigualdad de Landau-Pollak admite todos los pares de distribuciones de probabilidad de la forma $(1/2, 1/2, 0, \dots, 0)$, lo cual es demasiado permisivo para dimensiones altas y solapamientos pequeños, al menos si no se restringe la forma de las demás probabilidades. Por ello, queda como una interesante opción para investigación futura el añadir este tipo de restricciones u otras desigualdades relacionadas con el principio de incertidumbre (véase [8]) para, desarrollando una estrategia análoga a la presentada aquí, obtener mejores cotas para las EURs.
- En el capítulo 5, siguiendo la línea de investigación abierta en trabajos anteriores [111, 118, 119] que habían explorado la posibilidad de detección de estados entrelazados mediante relaciones de incertidumbre, hemos obtenido condiciones necesarias para la separabilidad (i.e., suficientes para el entrelazamiento) a partir de la relación de incertidumbre de Landau-Pollak. Estas condiciones, aparte de ofrecer un nuevo acercamiento al problema de la separabilidad, tienen un mayor atractivo desde el punto de vista experimental puesto que se basan en valores esperados de observables (por tanto, pueden ser implementadas ignorando el estado del sistema), lo que puede requerir menos medidas que la determinación explícita de ρ . Hemos estudiado en detalle el caso de sistemas de dos qubits donde hemos encontrado que las condiciones derivadas de esta forma son en general más fuertes que las deducidas a partir de relaciones de incertidumbre entrópicas [118, 119]. Así, paradójicamente quizá, aunque la relación de Landau-Pollak débil es menos adecuada para la expresión del principio de incertidumbre que las relaciones de incertidumbre entrópicas

(matemáticamente es incluso más débil en este sentido si consideramos la EUR derivada en el capítulo anterior), ésta resulta más potente para la identificación de estados entrelazados, lo que se debe a la estructura interna de las medidas. Sin embargo, nuestras condiciones han resultado más débiles que la condición (2.66) basada en relaciones de incertidumbre con desviación estándar [111]. Así mismo, también hemos abordado el caso de sistemas bipartitos de dimensión arbitraria donde, a pesar de que la estrategia empleada en el caso de dos qubits es ahora inviable, el Corolario 5.1 ofrece una receta para construir condiciones necesarias para la separabilidad utilizando dos observables, directamente a partir de la relación de incertidumbre de Landau-Pollak (2.43) que satisfagan. Sin embargo, a diferencia de los casos entrópicos y de desviación estándar (véase, por ejemplo, (2.73) y (2.68)), al no existir generalizaciones de la relación de incertidumbre de Landau-Pollak para conjuntos de más de dos observables estas condiciones no se pueden fortalecer considerando la medida de más de dos observables, lo que restringe en cierto modo su capacidad para la detección de entrelazamiento. Por ello, ya no sólo desde el punto de vista de expresión del principio de incertidumbre, sería interesante poder construir tales generalizaciones. Remarcablemente, esto se ha hecho recientemente en [200] con posterioridad a la publicación de nuestros resultados. En este artículo se extiende la relación de incertidumbre de Landau-Pollak en su versión débil (2.43) a medidas generalizadas y conjuntos de más de dos observables. Siguiendo nuestro análisis, sus autores estudian de forma preliminar la relevancia de este resultado para el problema de la separabilidad, prometiendo un análisis detallado para una futura publicación.

- Motivados por la potencia de las condiciones de separabilidad basadas en relaciones de incertidumbre locales con desviación estándar, en el capítulo 6 hemos analizado el problema de la separabilidad al nivel de las correlaciones inherentes a los subsistemas de los estados cuánticos bipartitos $M \times N$. Hemos encontrado que la representación de Bloch ofrece un formalismo adecuado para este análisis, lo que nos ha permitido obtener de forma operacional condiciones necesarias y condiciones suficientes para la separabilidad en términos de los vectores de coherencia locales $\mathbf{r} \in \mathbb{R}^{M^2-1}$ y $\mathbf{s} \in \mathbb{R}^{N^2-1}$ y de la matriz de correlación $T \in \mathbb{R}^{(M^2-1) \times (N^2-1)}$. Adicionalmente, las condiciones suficientes ofrecen una descomposición en estados producto para los estados separables que satisfacen sus hipótesis. Por otro lado, merece destacarse la condición necesaria para la separabilidad expresada en el Teorema 6.1, que, al mostrar que existe una cota superior al grado de correlación presente en los estados separables, ofrece una condición suficiente para la identificación de estados entrelazados operacional, potente y con una interpretación física, además de original. De hecho, hemos comprobado que existen situaciones en las que esta condición es más fuerte que otros criterios (aunque en otras puede ser más débil). Es más, hemos verificado que existen estados entrelazados

reconocidos por nuestra condición para los que los criterios anteriores más potentes fallan. Así mismo, hemos aportado evidencias que indican una mayor facilidad, en comparación con otros criterios, a la hora de la implementación experimental de nuestra condición, ya que ésta depende exclusivamente de la matriz de correlación a través de $\|T\|_{tr}$ y, por ello, no es necesario determinar todos los parámetros que caracterizan al operador densidad, tratándose, además, de una medida invariante bajo cambios de base locales. Por todo esto, podemos concluir que, aunque no se alcance el complicado objetivo final de caracterizar el entrelazamiento, los resultados de este capítulo, y en especial el Teorema 6.1, contribuyen a mejorar nuestra capacidad para reconocer y comprender este fenómeno. De hecho, como muestra la Proposición 6.4, el Teorema 6.1 es especialmente potente, en comparación a otros criterios, en el caso de estados con subsistemas maximalmente mezclados. Destacablemente, en [182], con posterioridad a la publicación de nuestros resultados, se ha mostrado que el problema de la separabilidad se reduce a la caracterización del entrelazamiento en estos estados, lo que redonda en la importancia de nuestro resultado. Como ya hemos mencionado, en este artículo se obtiene un nuevo criterio operacional a partir de las LURs que es más potente que el nuestro en este caso cuando $M \ll N$, pero más débil cuando $M \sim N$. Así mismo, merece la pena mencionarse que en [201] se han extendido las ideas desarrolladas en este capítulo al caso de sistemas cuánticos multipartitos. Por otro lado, creemos que un conocimiento más profundo del carácter geométrico del espacio de Bloch podría llevar a una mejora de las condiciones de separabilidad aquí obtenidas.

- Dada la eficacia de la matriz de correlación para la detección cualitativa del entrelazamiento estudiada en el capítulo 6, en el capítulo 7 hemos ido un paso más lejos y hemos analizado la posibilidad de obtener a través de esta matriz información cuantitativa acerca del entrelazamiento de un estado. Hemos encontrado que para estados puros existe una correspondencia directa entre cantidad de entrelazamiento, medido a través de la concurrencia y/o el tangle, y la información contenida en la matriz de correlación, medida a través de la norma de Hilbert-Schmidt, que, además, sirve para acotar inferiormente el tangle de estados generales (mezcla). Así mismo, hemos demostrado que la condición de separabilidad dada por $\|T\|_{tr}$ en el Teorema 6.1 se puede reformular para establecer una cota inferior para las medidas de entrelazamiento C , τ y E_F para todos los estados entrelazados identificados por esta condición. Por tanto, el criterio desarrollado en el capítulo anterior no sólo contribuye a mejorar nuestra capacidad para comprender e identificar el entrelazamiento, sino que también nos ayuda a cuantificarlo. De esta forma, además de aclarar la relación entre cantidad de entrelazamiento y cantidad de correlación, dada la complejidad de su evaluación, estos resultados permiten estimar medidas de entrelazamiento en analogía a los resultados previos de [188] y [189] obtenidos en relación con los criterios PPT y CCNR. Desde

este punto de vista práctico, hemos analizado que los resultados basados en la matriz de correlación mejoran las estimaciones anteriores para estados de gran pureza y para aquellos estados en los que el criterio del Teorema 6.1 es más fuerte que estos dos últimos. Además, se ha señalado cómo, a diferencia de estos dos criterios, es posible obtener información cuantitativa acerca del entrelazamiento a partir de estos resultados en situaciones en las que el contexto experimental limite significativamente el número de medidas posibles.

En la segunda parte de este capítulo hemos estudiado desde la misma perspectiva el criterio de separabilidad de las LURs basado en desviaciones estándar, ya que, como hemos visto entre otras cosas en el capítulo 5, es el más fuerte de todos los basados en relaciones de incertidumbre. Análogamente a los resultados anteriores hemos establecido que las violaciones de este criterio fijan una cota inferior para la cantidad de entrelazamiento de un estado en términos de C , τ y E_F . Además, hemos aportado ejemplos que muestran que las estimaciones obtenidas de esta forma son en general más fuertes que los resultados anteriores. Sin embargo, la aplicación de este criterio es más complicada ya que depende de la elección del conjunto de observables que se va a utilizar. Aunque, como hemos visto, una elección intuitiva basta para cosechar buenos resultados (de hecho, suficientes para mejorar las estimaciones anteriores incluso en las situaciones en que el criterio PPT es óptimo para la detección del entrelazamiento), sería deseable desarrollar un mecanismo para obtener el conjunto de observables óptimo en cada caso (i.e., el que da la cota más restrictiva). Interesantemente, este problema se ha resuelto recientemente en [202] con posterioridad a la publicación de nuestros resultados, lo cual tiene además el interés de destilar un criterio operacional del de las LURs formulado en (2.69). Notablemente, de los resultados de [202] se deduce que nuestro Teorema 6.1 es más fuerte que las LURs con LOOs (2.69) para estados con subsistemas maximalmente mezclados de diferente dimensionalidad (y son equivalentes si las dimensiones son iguales), lo que vuelve a atestiguar la potencia del Teorema 6.1 en esta situación tan relevante así como la independencia de ambos criterios. Así pues, dependiendo del estado en cuestión bien el resultado anterior (7.1) o bien alguno de los nuevos desarrollados en este capítulo (Teoremas 7.1, 7.2 y 7.4) da la estimación más fuerte para la cantidad de entrelazamiento. Queda pendiente para investigación futura el clarificar en mayor grado la potencia de estas estimaciones en relación a la obtenida por un método totalmente diferente en [187]. Así mismo, de igual forma que se puede mejorar la capacidad de detectar el entrelazamiento de muchos estados transformando a la FNF como se propuso en [182], sería interesante encontrar una forma de beneficiarse de esta idea para obtener mejores estimaciones de la cantidad de entrelazamiento.

Apéndice A

Descomposición en valores singulares y algunas normas matriciales

Para comodidad del lector, en esta sección presentamos brevemente algunos conceptos de análisis matricial que serán muy útiles en algunas partes de esta tesis. El lector interesado en una exposición más detallada y exhaustiva es remitido a las referencias estándar en estos temas, por ejemplo [174] y [203].

La descomposición en valores singulares es una factorización matricial que recuerda en algunos aspectos a la diagonalización pero que tiene la ventaja de existir para cualquier matriz (incluso no cuadrada). Dada una matriz cualquiera $A \in \mathbb{C}^{m \times n}$ se puede factorizar en la forma

$$A = U\Sigma V^\dagger, \quad (\text{A.1})$$

donde $U \in \mathbb{C}^{m \times m}$, $V \in \mathbb{C}^{n \times n}$ son matrices unitarias y $\Sigma \in \mathbb{R}_+^{m \times n}$ es una matriz pseudodiagonal, o sea,

$$(\Sigma)_{ii} = \sigma_i, \quad (\Sigma)_{ij} = 0 \quad (\forall i \neq j),$$

donde los $\{\sigma_i\}_{i=1}^{\min\{m,n\}}$ ($\sigma_i \geq 0 \forall i$), son los valores singulares de A , que por conveniencia se ordenan en forma no creciente. Una vez demostrado que esta descomposición existe, es directo comprobar que los valores singulares son la raíz cuadrada de los autovalores de la matriz semidefinida positiva $A^\dagger A$ (o AA^\dagger). Alternativamente, la factorización (A.1) se puede escribir como

$$A = \sum_{i=1}^{\min\{m,n\}} \sigma_i u_i v_i^\dagger, \quad (\text{A.2})$$

donde los $\{u_i\}$ y $\{v_i\}$ son los vectores columna de U y V .

Como cualquier norma, una norma matricial debe satisfacer la no-negatividad ($\|A\| \geq 0$ con $\|A\| = 0 \Leftrightarrow A = 0$), la homogeneidad ($\|cA\| = |c|\|A\| \forall c \in \mathbb{C}$) y la desigualdad triangular ($\|A + B\| \leq \|A\| + \|B\|$). Adicionalmente, se suele requerir la submultiplicatividad ($\|AB\| \leq \|A\|\|B\|$). Algunos ejemplos importantes son la norma de Hilbert-Schmidt (normalmente denominada de Frobenius en el contexto matricial):

$$\|A\|_{HS} = \sqrt{\text{Tr}(A^\dagger A)} = \sqrt{\sum_{ij} |a_{ij}|^2}, \quad (\text{A.3})$$

que es la norma inducida por el producto escalar de Hilbert-Schmidt (2.3), la norma en traza o de Ky Fan:

$$\|A\|_{tr} = \text{Tr}(\sqrt{A^\dagger A}), \quad (\text{A.4})$$

y la norma espectral:

$$\|A\|_2 = \max_{x \in \mathbb{C}^n} \frac{\|Ax\|_2}{\|x\|_2}, \quad (\text{A.5})$$

que es la norma inducida por la norma euclídea en \mathbb{C}^n . Se puede demostrar que todas estas normas son unitariamente invariantes, i.e.,

$$\|UAV\| = \|A\| \quad (\text{A.6})$$

para cualesquiera matrices unitarias U y V (de dimensiones apropiadas para que tenga sentido la multiplicación). En virtud de (A.1), cualquier norma unitariamente invariante es función exclusivamente de sus valores singulares. Así, se puede comprobar que

$$\|A\|_{HS} = \sqrt{\sum_i \sigma_i^2}, \quad (\text{A.7})$$

$$\|A\|_{tr} = \sum_i \sigma_i, \quad (\text{A.8})$$

$$\|A\|_2 = \max_i \sigma_i. \quad (\text{A.9})$$

De aquí se deduce que cualquier matriz cumple que

$$\|A\|_{tr} \geq \|A\|_{HS} \geq \|A\|_2. \quad (\text{A.10})$$

Apéndice B

Una prueba alternativa de la identidad (3.86)

Lema B.1.

$$\sum_{k=\lceil \frac{m+1}{2} \rceil}^m \frac{(k-1)!}{(2k-m)!(m-k)!} x^k = \frac{(x + \sqrt{x(4+x)})^m + (x - \sqrt{x(4+x)})^m}{m2^m} \quad (m \neq 0).$$

Demostración. Por un lado tenemos que

$$S_m(x) = \sum_{k=\lceil \frac{m+1}{2} \rceil}^m \frac{(k-1)!}{(2k-m)!(m-k)!} x^k = \sum_{j=0}^{\lfloor \frac{m}{2} \rfloor} \frac{(m-j-1)!}{j!(m-2j)!} x^{m-j}. \quad (\text{B.1})$$

Puesto que para cualquier $j \in \mathbb{N}$,

$$\lim_{\epsilon \rightarrow 0} \frac{(m + \epsilon - j - 1)!}{(m - 2j)!} = \begin{cases} \frac{(m-j-1)!}{(m-2j)!} & 0 \leq j \leq \lfloor m/2 \rfloor \\ \frac{(m-j-1)!}{(m-2j)!} \rightarrow 0 & \lfloor m/2 \rfloor + 1 \leq j \leq m-1 \\ \frac{\Gamma(2j-m)}{\Gamma(j+1-m)} \lim_{\epsilon \rightarrow 0} \frac{\sin(2j-m)\pi}{\sin(j+1-m-\epsilon)\pi} = 0 & j \geq m \end{cases} \quad (\text{B.2})$$

donde hemos utilizado que $(-z)! = \pi / (\Gamma(z) \sin \pi z)$, podemos extender la suma en (B.1) hasta el infinito,

$$S_m(x) = \lim_{\epsilon \rightarrow 0} \sum_{j=0}^{\infty} \frac{(m + \epsilon - j - 1)!}{j!(m-2j)!} x^{m-j}, \quad (\text{B.3})$$

lo cual, utilizando

$$(-k)_j = \frac{(-1)^j k!}{(k-j)!} \quad (\text{B.4})$$

y la fórmula de duplicación

$$\Gamma(2z) = 2^{2z-1} \frac{\Gamma(z)\Gamma(z+1/2)}{\Gamma(1/2)}, \quad (\text{B.5})$$

podemos transformar en una función hipergeométrica

$$\begin{aligned} S_m(x) &= x^m \lim_{\epsilon \rightarrow 0} \frac{(m + \epsilon - 1)!}{m!} \sum_{j=0}^{\infty} \frac{(-m/2)_j ((1-m)/2)_j}{(1-m-\epsilon)_j j!} \left(-\frac{4}{x}\right)^j \\ &= x^m \lim_{\epsilon \rightarrow 0} \frac{(m + \epsilon - 1)!}{m!} {}_2F_1 \left(\begin{matrix} -m/2, (1-m)/2 \\ 1-m-\epsilon \end{matrix}; -\frac{4}{x} \right). \end{aligned} \quad (\text{B.6})$$

Nótese que la introducción del parámetro ϵ evita la situación patológica de una función hipergeométrica con parámetro inferior entero negativo. Usando la fórmula de Pfaff [146],

$${}_2F_1 \left(\begin{matrix} -n, b \\ c \end{matrix}; x \right) = \frac{(c-b)_n}{(c)_n} {}_2F_1 \left(\begin{matrix} -n, b \\ b+1-n-c \end{matrix}; 1-x \right) \quad (n \in \mathbb{N}), \quad (\text{B.7})$$

se observa que

$$S_m(x) = x^m \lim_{\epsilon \rightarrow 0} \frac{(m + \epsilon - 1)!}{m!} \frac{(1/2 - n - \epsilon)_n}{(1 - m - \epsilon)_n} {}_2F_1 \left(\begin{matrix} -m/2, (1-m)/2 \\ 1/2 + \epsilon \end{matrix}; \frac{4+x}{x} \right) \quad (\text{B.8})$$

donde hemos usado que cuando m es par $-m/2 = -n$ y cuando m es impar $(1-m)/2 = -n$. Ahora bien, utilizando que $(a-n)_n = (-1)^n (1-a)_n$,

$$\lim_{\epsilon \rightarrow 0} \frac{(1/2 - n - \epsilon)_n}{(1 - m - \epsilon)_n} = \begin{cases} \lim_{\epsilon \rightarrow 0} \frac{(1/2 + \epsilon)_n}{(n + \epsilon)_n} = \frac{(1/2)_n}{(n)_n} = 2^{1-2n} & m \text{ par} \\ \lim_{\epsilon \rightarrow 0} \frac{(1/2 + \epsilon)_n}{(n + 1 + \epsilon)_n} = \frac{(1/2)_n}{(n+1)_n} = 2^{-2n} & m \text{ impar} \end{cases} = 2^{1-m}, \quad (\text{B.9})$$

donde hemos aplicado la fórmula de duplicación (B.5). Así, se llega a

$$S_m(x) = \frac{2^{1-m} x^m}{m} {}_2F_1 \left(\begin{matrix} -m/2, (1-m)/2 \\ 1/2 \end{matrix}; \frac{4+x}{x} \right). \quad (\text{B.10})$$

Por otro lado,

$$\begin{aligned} (A+B)^k + (A-B)^k &= \sum_{i=0}^k \binom{k}{i} A^{k-i} B^i + \sum_{j=0}^k (-1)^j \binom{k}{j} A^{k-j} B^j \\ &= 2 \sum_{i=0}^{\lfloor k/2 \rfloor} \binom{k}{2i} A^{k-2i} B^{2i}, \end{aligned} \quad (\text{B.11})$$

lo que significa que

$$\frac{(x + \sqrt{x(4+x)})^m + (x - \sqrt{x(4+x)})^m}{m 2^m} = \frac{2}{m} \left(\frac{x}{2}\right)^m \sum_{i=0}^{\lfloor m/2 \rfloor} \binom{m}{2i} \left(\frac{4+x}{x}\right)^i. \quad (\text{B.12})$$

Como $\binom{m}{2i} = 0$ para $i > \lfloor m/2 \rfloor$ podemos extender la suma hasta el infinito. Expresando entonces (B.12) como una función hipergeométrica obtenemos (B.10), lo que concluye la prueba. \square

Bibliografía

- [1] A.W. Marshall e I. Olkin, *Inequalities: Theory of Majorization and Its Applications* (Academic Press, New York, 1979).
- [2] G. Birkhoff, Univ. Nac. Tucumán Rev. Ser. A **5**, 147 (1946).
- [3] G.H. Hardy, J.E. Littlewood, y G. Pólya, *Inequalities* (Cambridge University Press, Cambridge, 1952).
- [4] C.E. Shannon, Bell Syst. Tech. J. **27**, 379; 623 (1948) [Reimpreso en: *The Mathematical Theory of Communication*, eds. C.E. Shannon y W. Weaver (Univ. Illinois Press, Urbana, 1949)].
- [5] T.M. Cover y J.A. Thomas, *Elements of Information Theory* (Wiley-Interscience, New York, 1991).
- [6] H. Hatori, Kodai Math. Semi. Rep. **10**, 172 (1958).
- [7] A. Rényi, *Probability Theory* (North-Holland, Amsterdam, 1970).
- [8] J.B.M. Uffink, *Measures of Uncertainty and the Uncertainty Principle*, Tesis doctoral, Universidad de Utrecht (1990).
- [9] M. Ben-Bassat y J. Raviv, IEEE Trans. Inform. Theory **24**, 324 (1978).
- [10] H. Maassen y J.B.M. Uffink, Phys. Rev. Lett. **60**, 1103 (1988).
- [11] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
- [12] R.J.V. Santos, J. Math. Phys. **38**, 4104 (1997).
- [13] M. Gell-Mann y C. Tsallis, eds., *Nonextensive entropy: Interdisciplinary Applications* (Oxford University Press, New York, 2004).
- [14] C. Cohen-Tannoudji, B. Diu, y F. Laloë, *Quantum Mechanics* (Wiley, New York, 1977).
- [15] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, New York, 1993).

- [16] L.E. Ballentine, *Quantum Mechanics: A Modern Development* (World Scientific, Singapore, 1998).
- [17] M.A. Nielsen e I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
- [18] M.A. Neumark, *Izv. Akad. Nauk SSSR, Ser. Mat.* **4**, 53;277 (1940); *C. R. (Doklady) Acad. Sci. URSS* **41**, 359 (1943).
- [19] W. Heisenberg, *Z. Phys.* **43**, 172 (1927).
- [20] E.H. Kennard, *Z. Phys.* **44**, 326 (1927); *Phys. Rev.* **31**, 344 (1928).
- [21] H.P. Robertson, *Phys. Rev.* **34**, 163 (1929).
- [22] J.B.M. Uffink y J. Hilgevoord, *Found. Phys.* **15**, 925 (1985); *Microphysical Reality and Quantum Description*, eds. F. Selleri, A. van der Merwe, y G. Tarozzi (Reidel, Dordrecht, 1988).
- [23] D. Deutsch, *Phys. Rev. Lett.* **50**, 631 (1983).
- [24] H.P. Robertson, *Phys. Rev.* **46**, 794 (1934).
- [25] O. Gühne, *Phys. Rev. Lett.* **92**, 117903 (2004).
- [26] A. Serafini, *Decoherence and Entanglement in Continuous Variable Quantum Information*, Tesis doctoral, Universidad de Salerno (2004).
- [27] M. Ohya y D. Petz, *Quantum Entropy and Its Use* (Springer, Berlin, 1993).
- [28] D. Bouwmeester, A. Ekert, y A. Zeilinger (eds.), *The Physics of Quantum Information* (Springer, Berlin, 2000).
- [29] V. Vedral, *Introduction to Quantum Information Science* (Oxford University Press, New York, 2006).
- [30] M.J.W. Hall, *Phys. Rev. Lett.* **74**, 3307 (1995).
- [31] M.J.W. Hall, *Phys. Rev. A* **55**, 100 (1995).
- [32] M. Koashi, *J. Phys.: Conference Series* **36**, 98 (2006).
- [33] I. B. Damgaard, S. Fehr, R. Renner, L. Salvail, y C. Schaffner, *Advances in Cryptology - CRYPTO 2007*, Lecture Notes in Computer Science **4622**, 360 (Springer, Berlin, 2007).
- [34] J. M. Renes y J.-C. Boileau, arXiv:quant-ph/0702187v1.
- [35] R. Matsumoto y S. Watanabe, arXiv:0711.1731v1.

- [36] D. P. DiVincenzo, M. Horodecki, D. W. Leung, J. A. Smolin, y B. M. Terhal, *Phys. Rev. Lett.* **92**, 067902 (2004).
- [37] M. A. Ballester y S. Wehner, *Phys. Rev. A* **75**, 022319 (2007).
- [38] Reimpresión en *The Many-Worlds Interpretation of Quantum Mechanics*, eds. B.S. DeWitt y N. Graham (Princeton University Press, Princeton, 1973).
- [39] I.I. Hirschman Jr., *Amer. J. Math.* **79**, 152 (1957).
- [40] I. Białynicki-Birula y J. Mycielski, *Commun. Math. Phys.* **44**, 129 (1975).
- [41] W. Beckner, *Ann. Math.* **102**, 159 (1975).
- [42] J. Sánchez-Ruiz, *Entropic Uncertainty Relations in Quantum Mechanics*, Tesis doctoral, Universidad de Barcelona (1997).
- [43] K. Kraus, *Phys. Rev. D* **35**, 3070 (1987).
- [44] M. Krishna y K. R. Parthasarathy, *Sankhya Series A* **64**, 842 (2002).
- [45] J. Schwinger, *Proc. Natl. Acad. Sci. U.S.A.* **46**, 570 (1960); reimpreso en: *Quantum Kinematics and Dynamics* (Benjamin, New York, 1970).
- [46] J. Sánchez-Ruiz, *Phys. Lett. A* **244**, 189 (1998).
- [47] A. J. M. Garrett y S. F. Gull, *Phys. Lett. A* **151**, 453 (1990).
- [48] H.J. Landau y H.O. Pollak, *Bell Syst. Tech. J.* **40**, 65 (1961).
- [49] A.F. Nikiforov y V.B. Uvarov, *Special Functions of Mathematical Physics* (Birkhäuser, Basilea, 1988).
- [50] R.G. Parr y W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
- [51] S.R. Gadre, *Phys. Rev. A* **30**, 620 (1984).
- [52] S.R. Gadre, S.B. Sears, S.J. Chakravorty, y R.D. Bendale, *Phys. Rev. A* **32**, 311 (1985).
- [53] S.R. Gadre y R.D. Bendale, *Phys. Rev. A* **36**, 1932 (1987).
- [54] J.C. Angulo y J.S. Dehesa, *J. Chem. Phys.* **97**, 6485 (1992).
- [55] E. Romera, J.S. Dehesa, y R.J. Yáñez, *Int. J. Quantum Chem.* **56**, 627 (1995)

- [56] R.J. Yáñez, *Entropía de información, polinomios ortogonales y sistemas multielectrónicos*, Tesis Doctoral, Universidad de Granada (1996).
- [57] S.R. Gadre, *Information theoretical approaches to quantum chemistry*, en *Reviews of Modern Quantum Chemistry: A Celebration of the contributions of R.G. Parr*, ed. K.D. Sen (World Scientific, Singapur, 2003).
- [58] J. Sánchez-Ruiz y J. S. Dehesa, *J. Comput. Appl. Math.* **118**, 311 (2000).
- [59] J.S. Dehesa, A. Martínez-Finkelshtein, y J. Sánchez-Ruiz, *J. Comput. Appl. Math.* **133**, 23 (2001).
- [60] A.I. Aptekarev, V.S. Buyarov, y J.S. Dehesa, *Russian Acad. Sci. Sbornik Math.* **185**, 3 (1994) [en ruso]; *Russian Acad. Sci. Sbornik Math.* **82**, 373 (1995) [traducción al inglés].
- [61] W. Van Assche, R.J. Yáñez, y J.S. Dehesa, *J. Math. Phys.* **36**, 4106 (1995).
- [62] J. Sánchez-Ruiz, *Phys. Lett. A* **226**, 7 (1997).
- [63] J.S. Dehesa, R.J. Yáñez, A.I. Aptekarev, y V. Buyarov, *J. Math. Phys.* **39**, 3050 (1998).
- [64] V.S. Buyarov, J.S. Dehesa, A. Martínez-Finkelshtein, y J. Sánchez-Lara, *SIAM J. Scientific Computing* **26**, 488 (2004).
- [65] S.L. Braunstein y P. van Loock, *Rev. Mod. Phys.* **77**, 513 (2005).
- [66] E. Schrödinger, *Proc. Cambridge Philos. Soc.* **31**, 555 (1935).
- [67] A. Einstein, N. Podolski, y N. Rosen, *Phys. Rev.* **47**, 777 (1935).
- [68] J.S. Bell, *Physics* **1**, 195 (1964) [reimpreso en *Speakable and Unsayable in Quantum Mechanics* (Cambridge University Press, Cambridge, 1987)].
- [69] D. Deutsch, *Proc. R. Soc. London A* **400**, 97 (1985).
- [70] D. Deutsch y R. Jozsa, *Proc. R. Soc. London A* **439**, 553 (1992).
- [71] P.W. Shor, *SIAM J. Comp.* **26**, 1484 (1997).
- [72] L.K. Grover, *Phys. Rev. Lett.* **79**, 325 (1997).
- [73] R. Jozsa y N. Linden, *Proc. R. Soc. London A* **459**, 2011 (2003).
- [74] V.M. Kendon y W.J. Munro, *Quantum Inf. Comput.* **6**, 630 (2006).

- [75] A. Datta y G. Vidal, *Phys. Rev. A* **75**, 042310 (2007).
- [76] C.H. Bennet y G. Brassard, *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* (IEEE, New York, 1984) p 175.
- [77] A.K. Ekert, *Phys. Rev. Lett.* **67**, 661 (1991).
- [78] H. Buhrman, R. Cleve, y W. van Dam, *SIAM J. Comput.* **30**, 1829 (2001).
- [79] C.H. Bennet y S.J. Wiesner, *Phys. Rev. Lett.* **69**, 2881 (1992).
- [80] C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, y W.K. Wootters, *Phys. Rev. Lett.* **70**, 1895 (1993).
- [81] R.F. Werner, *Phys. Rev. A* **40**, 4277 (1989).
- [82] K. Życzkowski, P. Horodecki, A. Sanpera, y M. Lewenstein, *Phys. Rev. A* **58**, 883 (1998).
- [83] L. Gurvits, *Proceedings of the 35 Annual ACM Symposium on Theory of Computing (San Diego)* (ACM Press, New York, 2003) p 10.
- [84] L.M. Ioannou, *Quantum Inf. Comput.* **7**, 335 (2007).
- [85] L. Gurvits y H. Barnum, *Phys. Rev. A* **66**, 062311 (2002).
- [86] S.L. Braunstein, C.M. Caves, R. Jozsa, N. Linden, S. Popescu, y R. Schack, *Phys. Rev. Lett.* **83**, 1054 (1999).
- [87] W.K. Wootters, *Phys. Rev. Lett.* **80**, 2245 (1998).
- [88] A. Sanpera, R. Tarrach, y G. Vidal, *Phys. Rev. A* **58**, 826 (1998).
- [89] P. Horodecki, M. Lewenstein, G. Vidal, e I. Cirac, *Phys. Rev. A* **62**, 032310 (2000).
- [90] M. Lewenstein, D. Bruß, J.I. Cirac, B. Kraus, M. Kuś, J. Samsonowicz, A. Sanpera, y R. Tarrach, *J. Mod. Opt.* **47**, 2841 (2000).
- [91] D. Bruß, *J. Math. Phys.* **43**, 4237 (2002).
- [92] R. Horodecki, P. Horodecki, M. Horodecki, y K. Horodecki, arXiv:quant-ph/0702225v2 (enviado a *Rev. Mod. Phys.*).
- [93] A. Peres, *Phys. Rev. Lett.* **77**, 1413 (1996).
- [94] M. Horodecki, P. Horodecki, y R. Horodecki, *Phys. Lett. A* **223**, 1 (1996).

- [95] M. Horodecki y P. Horodecki, Phys. Rev. A **59**, 4206 (1999).
- [96] N. Cerf, C. Adami, y R.M. Gingrich, Phys. Rev. A **60**, 898 (1999).
- [97] M.A. Nielsen y J. Kempe, Phys. Rev. Lett. **86**, 5184 (2001).
- [98] K.G.H. Vollbrecht y M.M. Wolf, J. Math. Phys. **43**, 4299 (2001).
- [99] R. Horodecki y M. Horodecki, Phys. Rev. A **54**, 1838 (1996).
- [100] M.A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).
- [101] M.A. Nielsen y G. Vidal, Quantum Inf. Comput. **1**, 76 (2001).
- [102] O. Rudolph, arXiv:quant-ph/0202121v1.
- [103] K. Chen y L.-A. Wu, Quantum Inf. Comput. **3**, 193 (2003).
- [104] M. Horodecki, P. Horodecki, y R. Horodecki, Open Syst. Inf. Dyn. **13**, 103 (2006).
- [105] M. Lewenstein, B. Kraus, J.I. Cirac, y P. Horodecki, Phys. Rev. A **62**, 052310 (2000).
- [106] A. Jamiolkowski, Rep. Math. Phys. **3**, 275 (1972).
- [107] M.D. Choi, Linear Algebra Appl. **12**, 95 (1975).
- [108] E. Størmer, Acta Math. **110**, 233 (1963).
- [109] S.L. Woronowicz, Rep. Math. Phys. **10**, 165 (1976).
- [110] L.-M. Duan, G. Giedke, J. I. Cirac, y P. Zoller, Phys. Rev. Lett. **84**, 2722 (2000).
- [111] H. F. Hofmann y S. Takeuchi, Phys. Rev. A **68**, 032103 (2003).
- [112] H.F. Hofmann, Phys. Rev. A **68**, 034307 (2003).
- [113] O. Gühne, M. Mechler, G. Tóth, y P. Adam, Phys. Rev. A **74**, 010301(R) (2006).
- [114] S. Yu y N.L. Liu, Phys. Rev. Lett. **95**, 150504 (2005).
- [115] J. Sánchez-Ruiz, Phys. Lett. A **173**, 233 (1993).
- [116] J. Sánchez-Ruiz, J. Phys. A: Math. Gen. **27**, L843 (1994).
- [117] J. Sánchez-Ruiz, Phys. Lett. A **201**, 125 (1995).
- [118] V. Giovannetti, Phys. Rev. A **70**, 012102 (2004).
- [119] O. Gühne y M. Lewenstein, Phys. Rev. A **70**, 022316 (2004).

- [120] V. Vedral, M.B. Plenio, M.A. Rippin, y P.L. Knight, *Physical Review Letters* **78**, 2275 (1997).
- [121] V. Vedral y M.B. Plenio, *Phys. Rev. A* **57**, 1619 (1998).
- [122] M.B. Plenio, *Phys. Rev. Lett.* **95**, 090503 (2005).
- [123] M.B. Plenio y S. Virmani, *Quantum Inf. Comput.* **7**, 1 (2007).
- [124] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin, y W.K. Wootters, *Phys. Rev. A* **54**, 3824 (1996).
- [125] P. Rungta, V. Bužek, C.M. Caves, M. Hillery, y G.J. Milburn, *Phys. Rev. A* **64**, 042315 (2001).
- [126] G. Vidal, *J. Mod. Opt.* **47**, 355 (2000).
- [127] K.M.R. Audenaert, F. Verstraete, y B. De Moor, *Phys. Rev. A* **64**, 052304 (2001).
- [128] B. Terhal y K.G.H. Vollbrecht, *Phys. Rev. Lett.* **85**, 2625 (2000).
- [129] K.G.H. Vollbrecht y R.F. Werner, *Phys. Rev. A* **64**, 062307 (2001).
- [130] P. Rungta y C.M. Caves, *Phys. Rev. A* **67**, 012307 (2003).
- [131] K. Chen, S. Albeverio, y S.-M. Fei, *Rep. Math. Phys.* **58**, 325 (2006).
- [132] T.J. Osborne, *Phys. Rev. A* **72**, 022309 (2005).
- [133] G. Vidal y R.F. Werner, *Phys. Rev. A* **65**, 032314 (2002)
- [134] J. Avery, *Hyperspherical Harmonics. Applications to Quantum Theory* (Kluwer, Dordrecht, 1990).
- [135] A. Berkdemir, C. Berkdemir, y R. Sever, *Mod. Phys. Lett. A* **21**, 2087 (2006).
- [136] S.M. Ikhdair y R. Sever, *J. Math. Chem.* **42**, 461 (2007).
- [137] R.J. Yáñez, W. Van Assche, y J.S. Dehesa, *Phys. Rev. A* **50**, 3065 (1994).
- [138] J.S. Dehesa, W. Van Assche, y R.J. Yáñez, *Meth. Appl. Anal.* **4**, 91 (1997).
- [139] V.S. Buyarov, *Vesnt. Mosk. Univ. (Ser. 1)* **6**, 8 (1997) [en ruso].
- [140] V.S. Buyarov, P. López-Artés, A. Martínez-Finkelshtein, y W. Van Assche, *J. Phys. A: Math. Gen.* **33**, 6549 (2000).

- [141] J. Sánchez-Ruiz, J. Phys. A: Math. Gen. **36**, 4857 (2003).
- [142] R.J. Yáñez, W. Van Assche, R. González-Férez, y J.S. Dehesa, J. Math. Phys. **40**, 5675 (1999).
- [143] J. Sánchez-Ruiz, J. Math. Phys. **38**, 5031 (1997).
- [144] J. Arias (sin publicar) [véase P. López-Artés, *Entropías, potenciales y linealización de polinomios hipergeométricos*, Tesis doctoral, Universidad de Almería (2002)].
- [145] J.F. Sánchez-Lara y J. Sánchez-Ruiz (sin publicar).
- [146] G.E. Andrews, R. Askey, y R. Roy, *Special Functions* (Cambridge University Press, Cambridge, 1999).
- [147] G. Szegő, Math. Ann. **110**, 501 (1934) [en alemán].
- [148] G. Szegő, *Orthogonal Polynomials* (Am. Math. Soc. Colloq. Publ., vol. 23) (American Mathematical Society, Providence, RI, 1975).
- [149] S. Roman, Amer. Math. Monthly **87**, 805 (1980).
- [150] R.L. Graham, D.E. Knuth, y O. Patashnik, *Concrete Mathematics* (Addison-Wesley, Reading, Massachusetts, 1994).
- [151] E.R. Hansen, *A Table of Series and Products* (Prentice-Hall, Englewood Cliffs, New Jersey, 1975).
- [152] M. Feder and N. Merhav, IEEE Trans. Inf. Theory **40**, 259 (1994).
- [153] R.E. Miller, *Optimization, Foundations and Applications* (Wiley-Interscience, New York, 2000).
- [154] O. Gühne y M. Lewenstein, AIP Conf. Proc. **734**, 230 (2004).
- [155] O. Gühne y N. Lütkenhaus, Phys. Rev. Lett. **96**, 170502 (2006).
- [156] J. Sánchez-Ruiz, Phys. Lett. A **181**, 193 (1993).
- [157] S. Samuelsson y G. Björk, Phys. Rev. A **73**, 012319 (2006).
- [158] F.T. Hioe y J.H. Eberly, Phys. Rev. Lett. **47**, 838 (1981).
- [159] F. Bloch, Phys. Rev. **70**, 460 (1946).
- [160] G. Kimura y A. Kossakowski, Open Sys. Inf. Dyn. **12**, 207 (2005).
- [161] G. Kimura, Phys. Lett. A **314**, 339 (2003).
- [162] M.S. Byrd y N. Khaneja, Phys. Rev. A **68**, 062322 (2003).
- [163] J.E. Harriman, Phys. Rev. A **17**, 1249 (1978).

- [164] I. Bengtsson y K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, 2005).
- [165] U. Fano, *Rev. Mod. Phys.* **55**, 855 (1983).
- [166] C.M. Caves y G.J. Milburn, *Opt. Commun.* **179**, 439 (2000).
- [167] P. Rungta, W.J. Munro, K. Nemoto, P. Deuar, G.J. Milburn, y C.M. Caves, *Directions in Quantum Optics: A collection of papers dedicated to the memory of Dan Walls*, eds. H.J. Carmichael, R.J. Glauber, y M.O. Scully (Springer, Berlin, 2001).
- [168] J. Schlienz y G. Mahler, *Phys. Rev. A* **52**, 4396 (1995).
- [169] R. Horodecki, P. Horodecki, y M. Horodecki, *Phys. Lett. A* **200**, 340 (1995).
- [170] J.F. Clauser, M.A. Horne, A. Shimony, y R.A. Holt, *Phys. Rev. Lett.* **23**, 880 (1969).
- [171] B.M. Terhal, *Phys. Lett. A* **271**, 319 (2000).
- [172] C.H. Bennett, D.P. DiVincenzo, T. Mor, P.W. Shor, J.A. Smolin, y B.M. Terhal, *Phys. Rev. Lett.* **82**, 5385 (1999).
- [173] D. Bruß y A. Peres, *Phys. Rev. A* **61**, 030301(R) (2000).
- [174] R.A. Horn y C.R. Johnson, *Topics in Matrix Analysis* (Cambridge University Press, Cambridge, 1991).
- [175] P. Horodecki, *Phys. Lett. A* **232**, 333 (1997).
- [176] P. Horodecki, M. Horodecki, y R. Horodecki, *Phys. Rev. Lett.* **82**, 1056 (1999).
- [177] D.P. DiVincenzo, T. Mor, P.W. Shor, J.A. Smolin, y B.M. Terhal, *Comm. Math. Phys.* **238**, 379 (2003).
- [178] H.-P. Breuer, *J. Phys. A: Math. Gen.* **38**, 9019 (2005).
- [179] R. Augusiak y J. Stasińska, *Phys. Lett. A* **363**, 182 (2007).
- [180] F. Verstraete, J. Dehaene, y B. De Moor, *Phys. Rev. A* **68**, 012103 (2003)
- [181] J.M. Leinaas, J. Myrheim, y E. Ovrum, *Phys. Rev. A* **74**, 012313 (2006).
- [182] O. Gühne, P. Hyllus, O. Gittsovich, y J. Eisert, *Phys. Rev. Lett.* **99**, 130504 (2007).

- [183] F. Verstraete, *A Study of Entanglement in Quantum Information Theory*, Tesis Doctoral, Katholieke Universiteit Leuven (2002).
- [184] D.P. DiVincenzo y B.M. Terhal, *Proc. of the XIII Int. Cong. on Mathematical Physics (London 2000)*, 399 (International Press, Boston, 2001).
- [185] C. Kothe y G. Björk, *Phys. Rev. A* **75**, 012336 (2007).
- [186] A. Loziński, A. Buchleitner, K. Zyczkowski, y T. Wellens, *Europhys. Lett.* **62**, 168 (2003).
- [187] F. Mintert, M. Kuś, y A. Buchleitner, *Phys. Rev. Lett.* **92**, 167902 (2004).
- [188] K. Chen, S. Albeverio, y S.-M. Fei, *Phys. Rev. Lett.* **95**, 040504 (2005).
- [189] K. Chen, S. Albeverio, y S.-M. Fei, *Phys. Rev. Lett.* **95**, 210501 (2005).
- [190] S.-M. Fei y X. Li-Jost, *Phys. Rev. A* **73**, 024302 (2006).
- [191] J. Eisert, F.G.S.L. Brandão, y K.M.R. Audenaert, *New J. Phys.* **9**, 46 (2007).
- [192] O. Gühne, M. Reimpell, y R.F. Werner, *Phys. Rev. Lett.* **98**, 110502 (2007).
- [193] F. Mintert, *Phys. Rev. A* **75**, 052302 (2007).
- [194] H.-P. Breuer, *J. Phys. A: Math. Gen.* **39**, 11847 (2006).
- [195] A. Datta, S.T. Flammia, A. Shaji, y C.M. Caves, *Phys. Rev. A* **75**, 062117 (2007).
- [196] J. Zhang, C.-W. Li, J.-W. Wu, R.-B. Wu, y T.J. Tarn, *Phys. Rev. A* **73**, 022319 (2006).
- [197] J. Zhang, C.-W. Li, T.J. Tarn y J.-W. Wu, *Phys. Rev. A* **76**, 032306 (2007).
- [198] K.M.R. Audenaert y M.B. Plenio, *New J. Phys.* **8**, 266 (2006).
- [199] I. Ali Khan y J.C. Howell, *Phys. Rev. A* **70**, 062320 (2004).
- [200] T. Miyadera y H. Imai, *Phys. Rev. A* **76**, 062108 (2007).
- [201] A.S.M. Hassan y P.S. Joag, arXiv:0704.3942v4 (próxima publicación en *Quantum Inf. Comput.*).
- [202] C.J. Zhang, Y.S. Zhang, S. Zhang, y G.C. Guo, *Phys. Rev. A* **76**, 012334 (2007)
- [203] R. Bhatia, *Matrix Analysis* (Springer-Verlag, New York, 1997).