“Desigualdades para la entropía de von Neumann”

Tesis que presenta:

.

Leopoldo Pantaleón Martínez.

Para obtener el grado de

Maestro en Ciencias (Matemáticas).

Asesor: Dr. Roberto Quezada Batalla.

Septiembre del 2003.

Universidad Autónoma Metropolitana-Iztapalapa.
División de Ciencias Básicas e Ingeniería.

Desigualdades para la entropía de von Neumann

Leopoldo Pantaleón Martínez

8 de septiembre de 2003

Índice General

Reconocimientos

Agradezco a la UAM-I todas las facilidades que obtuve para realizar mis estudios de Maestría y para la realización de esta tesis.

Particularmente agradezco a los **EXCELENTES** PROFESORES, que como alumno tuve la buena suerte de tener:
Jesús Chargoy Corona, Gabriel Villa Salvador, Carlos Signoret Poillon, Richard Wilson Roberts, y muy en especial les doy las gracias a los profesores Julio García Corte y a Roberto Quezada Batalla.

Finalmente, deseo agregar que mis estudios de Maestría fueron realizados con el apoyo del CONACYT mediante una beca-crédito y mediante el proyecto número 37491E.

Introducción

Entropía clásica.

Un concepto central de la teoría de la información es la entropía. La entropía es una medida de la cantidad de información (ó incertidumbre) contenida en una fuente de información. También puede ser usada para cuantificar los recursos necesarios para almacenar la información ver [11]. La fuente puede ser modelada como un proceso estocástico y esto motiva el que se defina la entropía de *Shannon*, $H(X)$ de una variable aleatoria X en (Ω, P), un espacio de probabilidad finito, como

$$H\left(X\right) \equiv -\sum_{x\in X(\Omega)} p_x \log p_x,$$

donde $\{p_x\}$ es la distribución de probabilidad de X, y por definición se acepta que $0 \log 0 = 0$.

Entropía cuántica.

Shannon [17] introdujo su definición en 1948 pero mucho antes, en 1927 *von Neumann* [19,20], queriendo extender la teoría clásica de la mecánica estadística (desarrollada por *Gibbs*, *Boltzman* et al.) al dominio cuántico, mas que buscar desarrollar una teoría de la comunicación cuántica, introdujo las nociones de estado mezclado, representado por una matriz ρ de densidad (i.e. $\rho \geq 0$, $tr\rho = 1$, donde $tr\rho$ denota la traza de la matriz ρ) y su entropía definida como

$$S\left(\rho\right) \equiv -tr\left(\rho \log \rho\right),$$

expresión que, en términos de los valores propios de ρ, se puede escribir como $S\left(\rho\right) = -\sum_{x\in\Omega} \lambda_x \log \lambda_x$, donde nuevamente por definición $0 \log 0 = 0$.

Como las matrices de densidad diagonales (llamadas estados clásicos) corresponden a distribuciones de probabilidad, de la igualdad anterior se sigue que la entropía de *Shannon* es un caso especial de la de *von Neumann*.

Consideremos un sistema compuesto, cuyo espacio de *Hilbert* es un producto tensorial $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ de los espacios de *Hilbert* de dos sistemas A y B. Cuando el estado del sistema compuesto está descrito por la matriz de densidad ρ_{AB} los estados de los subsistemas están dados por las matrices de densidad,

$\rho_A = tr_B(\rho_{AB})$, y $\rho_B = tr_A(\rho_{AB})$ obtenidas al tomar la traza parcial (ver la Sección 2.1 para la definición de traza parcial).

El objetivo principal de este trabajo es discutir las interrelaciones que hay en un conjunto importante de desigualdades relacionadas con la entropía de los subsistemas de un sistema compuesto y algunas aplicaciones en la teoría cuántica de la información; algunas de las desigualdades son las siguientes:
La subaditividad

$$S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B),$$

demostrada en [7];
la concavidad de la entropía,

$$S(\lambda\rho + (1-\lambda)\rho') \geq \lambda S(\rho) + (1-\lambda) S(\rho'),$$

demostrada en [9, 21], que como veremos puede obtenerse usando la desigualdad de subaditividad.

Y cuando el sistema compuesto consta de tres subsistemas se cumple la siguiente desigualdad más fuerte conocida como subaditividad fuerte que se suele simbolizar, por sus siglas en inglés como **(SSA)**

$$S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC}),$$

esta desigualdad fue probada en [9].

Entropía relativa cuántica.

Un concepto útil en la discusión de la entropía que además es interesante por derecho propio es el de la *entropía relativa cuántica*; se define para matrices de densidad (i.e., matrices positivo semi-definidas de traza 1) tales que $\ker\sigma \subseteq \ker\rho$, como

$$S(\rho \parallel \sigma) \equiv tr\rho\log\rho - tr\rho\log\sigma.$$

La definición anterior todavía funciona si eliminamos la condición de que los operadores sean de traza 1.

Veremos en el capítulo cuatro, que la **(SSA)** puede ser reenunciada como una propiedad de monotonía de la entropía relativa bajo el operador traza parcial como

$$S(\rho_{AB} \parallel I_A \otimes \rho_B) \leq S(\rho_{ABC} \parallel I \otimes \rho_{BC}).$$

Mas aún la entropía relativa es monótona bajo operadores Φ completamente positivos **(CP)** que preservan la traza, (de los cuales la traza parcial es un caso particular) es decir se cumple que:

$$S(\Phi(\sigma) \parallel \Phi(\gamma)) \leq S(\sigma \parallel \gamma).$$

Panorama general.

En este trabajo restringiremos nuestra discusión a espacios de *Hilbert* complejos de **dimensión finita** y asumiremos, salvo que se especifique otra cosa, que **las matrices de densidad consideradas son estrictamente positivas**.

Lo que resta del escrito está estructurado de la siguiente manera:
En el capítulo 1, recordamos algunos conceptos básicos y se discuten los resultados de la teoría de los espacios de *Hilbert* necesarios en la tesis, entre otros: el importante teorema espectral para operadores normales y con base en él presentamos el cálculo funcional para esta clase de operadores; las propiedades básicas de la traza de un operador lineal y los productos tensoriales tanto de espacios de *Hilbert* como de operadores.

En el capítulo 2, introducimos la terminología y nociones básicas relacionadas con la *entropía de von Neumann*: la noción de traza parcial; el proceso de purificación; la descomposición de *Schmidt*; el concepto de operador completamente positivo (**CP**) y una generalización de la representación de *Lindblad* [10].

En el capítulo 3, discutimos los antecedentes de las nociones de *von Neumann*: propiedades básicas de la entropía de *Shanonn*, la entropía condicional e información conjunta; la desigualdad de *Klein* junto con una condición necesaria y suficiente para que se cumpla la igualdad; terminamos el capítulo con las propiedades básicas de la entropía y la entropía relativa de *von Neumann.*

En el capítulo 4, que es el núcleo de este trabajo aparece inicialmente el bagaje matemático necesario para el resto del capítulo; luego demostramos las desigualdades siguientes junto con una condición necesaria y suficiente para que se dé la igualdad (excepto de CON) :

(SSA) La **subaditividad fuerte**

(MONO) La monotonía de la entropía relativa bajo operadores CP que preservan la traza.

(MPT) La monotonía de la entropía relativa bajo la traza parcial.

(CVX) La convexidad conjunta de la entropía relativa.

(CON) La concavidad de la entropía condicional.

Terminamos el capítulo cuatro discutiendo la equivalencia entre estas cinco desigualdades.

En el capítulo 5, presentamos algunas aplicaciones y perspectivas: la desigualdad de subaditividad junto con una condición de igualdad; la concavidad de la entropía con una condición de igualdad; la desigualdad del procesamiento de datos y presentamos una prueba completa de la importante desigualdad conocida como la cota de *Holevo.*

En el apéndice abarcamos algunos temas clásicos necesarios en el trabajo: las nociones de convexidad, concavidad y el teorema de *Jensen*; la notación O;

la derivada y la integral de *Riemann* de funciones con dominio en un intervalo de números reales y con valores en $\mathcal{B}(\mathcal{H})$.

Finalizamos esta introducción con los siguientes comentarios:

a) El trabajo es auto-contenido en el sentido de que los resultados relevantes presentados que se escapan del ámbito de los conocimientos generales están demostrados en la tesis.

b) Este trabajo está basado en algunas de las ideas principales del artículo [16] de *Mary Beth Ruskai* (MBR) a saber, la demostración de las desigualdades SSA, MONO, MPT y CVX de la lista escrita líneas arriba, todas ellas con una condición necesaria y suficiente de igualdad, así como la equivalencia entre ellas.

Nuestros principales aportes y diferencias en relación al mencionado artículo, son:

En el capítulo 2, desarrollamos las propiedades del operador traza parcial y damos una prueba completa de una versión general de la representación de *Lindblad*. En el artículo de **MBR** se bosqueja la prueba del resultado para el caso en que el dominio y codominio del operador **CP** involucrado son iguales y se menciona que la representación propuesta es válida en el caso en que esta condición es eliminada. Nosotros demostramos esto último.

En el capítulo 3, demostramos completamente la desigualdad de *Klein* junto con una condición necesaria y suficiente para que se cumpla la igualdad. Esta desigualdad y su condición de igualdad resultan muy útiles para demostrar las condiciones de igualdad de las desigualdades mencionadas en los capítulos 4 y 5; también agregamos las propiedades básicas de la entropía de *von Neumann*.

En el capítulo 4, demostramos la fórmula de *Lie* y la desigualdad de *Golden-Thompson*; añadimos la desigualdad CON y diseñamos una ruta diferente para demostrar las equivalencias, si bien para probar que MONO implica JCX, desarrollamos el bosquejo que da **MBR**. También presentamos una condición necesaria (Corolario 4.11), al parecer nueva, para que se cumpla la igualdad en MONO.

En el capítulo 5, sin ser completamente original, la demostración de la desigualdad de la concavidad junto con su condición de igualdad, no es la que aparece en [16]; aunque la prueba de la cota de *Holevo* se basa en el esbozo que da **MBR**, nosotros desarrollamos una demostración que tiene las ventajas de ser poco conocida, muy simple y que utiliza la teoría que hemos desarrollado.

Capítulo 1

Elementos de los espacios de Hilbert

1.1 Nociones básicas

Por un espacio de *Hilbert* $\mathcal{H}$ siempre se entenderá un espacio vectorial **complejo** dimensionalmente **finito** con un producto interno $\langle \cdot, \cdot \rangle$ lineal conjugado en la primera variable y lineal en la segunda.

Usamos el símbolo $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ (ó $\mathcal{B}(\mathcal{H}_1)$ cuando $\mathcal{H}_1 = \mathcal{H}_2$) para el espacio de Banach de operadores lineales T de $\mathcal{H}_1$ en $\mathcal{H}_2$ con la norma

$$\|T\| = \sup_{x \neq 0} \frac{\|Tx\|_2}{\|x\|_1};$$

siendo el espacio $\mathcal{H}_1$ de dimensión finita se sabe que $\|T\| < \infty$. La norma de T también se puede calcular como $\|T\| = \sup_{\|x\|=1} \|Tx\| = \sup_{\|x\| \leq 1} \|Tx\|$.

Antes de proseguir **convenimos en usar como sinónimos las palabras matriz y operador lineal.**
Para $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ se define su adjunto como el único operador $T^* \in \mathcal{B}(\mathcal{H}_2, \mathcal{H}_1)$ tal que

$$\forall x \in \mathcal{H}_2, \ \forall y \in \mathcal{H}_1, \ \langle x, Ty \rangle_2 = \langle T^* x, y \rangle_1 .$$

Si $S, T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ y $\lambda \in \mathbb{C}$ se tienen las siguientes propiedades:

a) $(\lambda S + T)^* = \overline{\lambda} S^* + T^*$

b) $(ST)^* = T^* S^*$

c) $(S^*)^* = S$

d) $\|S^*\| = \|S\|$.

Con relación a la operación de adjunción se definen las siguientes clases de operadores.

Definición
a) Un operador $A \in \mathcal{B}(\mathcal{H})$ es **Hermitiano o auto-adjunto** si $A^* = A$.

b) Cuando un operador $N \in \mathcal{B}(\mathcal{H})$ satisface $NN^* = N^*N$ se dice que es **normal**.

c) $U \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ es **unitario** si es un isomorfismo (i.e. un operador lineal biyectivo) isométrico (i.e. $\|Ux\|_2 = \|x\|_1$) ó equivalentemente si $U^*U = I_{\mathcal{H}_1}$ y $UU^* = I_{\mathcal{H}_2}$.

d) $V \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ es una **isometría parcial**, si para $h \in [\ker V]^{\perp}$, $\|Vh\| = \|h\|$. El espacio $[\ker V]^{\perp}$ se llama espacio inicial de V y el espacio $ranV$ es llamado espacio final de V.

Como veremos en las próximas secciones, los operadores *auto-adjuntos* y normales tienen propiedades importantes.

Supongamos que φ es un vector de $\mathcal{H}_1$ y ψ es vector de $\mathcal{H}_2$, **definimos** el operador lineal $|\psi\rangle\langle\varphi|$ de $\mathcal{H}_1$ en $\mathcal{H}_2$ mediante la correspondencia

$$|\psi\rangle\langle\varphi|\,\zeta = \langle\varphi, \zeta\rangle\,\psi, \quad \forall\zeta \in \mathcal{H}_1.$$

Es fácil verificar que la función $(v, u) \longrightarrow |u\rangle\langle v|$ de $\mathcal{H} \times \mathcal{H}$ a $\mathcal{B}(\mathcal{H})$ satisface lo siguiente:

a) $|u\rangle\langle v|$ es lineal en u y lineal conjugada en v.

b) $(|v\rangle\,\langle u|)^* = |u\rangle\langle v|$, en particular $|u\rangle\langle u|$ es auto-adjunto.

c) $|u\rangle\langle v|\,|\psi\rangle\langle\varphi| = \langle v, \psi\rangle\,|u\rangle\langle\varphi|$.

d) $\||u\rangle\langle v|\| = \|u\|\,\|v\|$.

e) Para todo $A \in \mathcal{B}(\mathcal{H})$, $A\,|u\rangle\langle v| = |Au\rangle\langle v|$ y $|u\rangle\langle v|\,A = |u\rangle\langle A^*v|$.

Podemos apreciar la utilidad de la notación anterior en la siguiente propiedad llamada **la relación de completez** para bases ortonormales:

Si $\{|i\rangle\}$ es una base ortonormal de $\mathcal{H}$ todo vector se escribe como $v = \sum_i \alpha_i |i\rangle$, donde $\alpha_i = \langle i, v\rangle$ y por tanto $\left(\sum_j |j\rangle \langle j|\right) v = \left(\sum_j |j\rangle \langle j|\right) \left(\sum_i \alpha_i |i\rangle\right) = \sum_j \sum_i \alpha_i |j\rangle \langle j| |i\rangle = \sum_j \sum_i \alpha_i \langle j, i\rangle |j\rangle = \sum_j \alpha_j |j\rangle = v$, es decir

$$\sum_i |i\rangle \langle i| = I.$$

Dado un subespacio $\mathcal{W}$ de $\mathcal{H}$, con una base ortonormal $|1\rangle, |2\rangle, ..., |k\rangle$ se define la *proyección ortogonal* P sobre $\mathcal{W}$ como

$$P = \sum_{i=1}^{k} |i\rangle \langle i|,$$

se puede demostrar que la suma no depende de la base ortonormal elegida, y es inmediato verificar que las *proyecciones ortogonales* forman parte de la clase de los operadores *auto-adjuntos.*

1.2 Descomposición espectral

Definiciones. Sea $A \in \mathcal{B}(\mathcal{H})$, con $\mathcal{H}$ un espacio de *Hilbert* de dimensión finita. Un vector v *no cero* tal que $Av = \lambda v$ es un **vector propio** del operador lineal A y el número complejo λ es el **valor propio** de A correspondiente a v. Los valores propios también se pueden definir como las raíces del polinomio $p_A(\lambda) = \det(\lambda I - A)$ llamado ***polinomio característico*** de A*;* puesto que nuestros espacios de *Hilbert* son complejos el teorema fundamental del álgebra garantiza que todo operador tiene al menos un valor propio y su correspondiente vector propio.

El conjunto de vectores propios correspondientes a un valor propio λ es el *espacio propio* asociado a λ; el conjunto de valores propios de un operador es su **espectro**; observe que el espectro es un conjunto finito porque está formado por los ceros del polinomio característico.

Se dice que A es **diagonalizable** si existe una base ortonormal de vectores propios suyos $\{w_i\} \subset \mathcal{H}$ con valores propios correspondientes $\{\lambda_i\}$ tales que $A = \sum_i \lambda_i |w_i\rangle \langle w_i|$. También se dice A que se *diagonaliza* con respecto a la base $\{w_i\}$.

Ahora comentaremos algunas propiedades que nos servirán para demostrar la que tal vez, sea la propiedad más importante de los operadores normales: que son diagonalizables.
Si N es operador normal, entonces
a) $\ker N = \ker N^*$, lo que se muestra desarrollando $\langle N^*x, N^*x\rangle$ y usando que $NN^* = N^*N$.

b) Si x es un elemento tal que $Nx = \lambda x$, entonces $N^* x = \overline{\lambda} x$; porque $(N - \lambda I)$ es normal y por la propiedad a), si $(N - \lambda I)\, x = 0$, también $(N - \lambda I)^*\, x = 0$.
c) Si v es un vector propio de N tal que $Nv = \lambda v$ y v, w son vectores ortogonales, entonces también Nw y v son ortogonales, pues $\langle v, Nw \rangle = \langle N^* v, w \rangle = \overline{\lambda} \langle v, w \rangle$.

Teorema 1 *(Descomposición espectral). Cualquier operador normal N en un espacio de Hilbert $\mathcal{H}$ es diagonal con respecto a alguna base ortonormal de $\mathcal{H}$. Recíprocamente, cualquier operador diagonalizable es normal.*

Demostración. La proposición recíproca es inmediata.
Probaremos la proposición directa por inducción en $d = \dim \mathcal{H}$. El caso 1 es trivial. Sea $d = n$, como estamos tratando con espacios complejos, toda matriz tiene n valores propios. Por las propiedades mencionadas arriba, N y N^* tienen un vector propio común w_1, que podemos suponer con norma 1.

Sea $W = \{w_1\}^{\perp}$, por la propiedad c) la restricción de N a W está bien definida y por la hipótesis de inducción, existe $\{w_2, w_3, ..., w_n\}$ una base ortonormal de W constituida por vectores propios de N. Es claro que $\{w_1, w_2, ..., w_n\}$ es una base ortonormal de $\mathcal{H}$, por ende todo $v \in \mathcal{H}$ se expresa como $v = \sum_i a_i w_i$ y entonces $\left(\sum_i \lambda_i \left|w_i\right\rangle \left\langle w_i\right|\right) v = \sum_i \lambda_i a_i w_i = N \left(\sum_i a_i w_i\right)$.

Concluimos que para todo operador normal N existe una base ortonormal de $\mathcal{H}$, $\{w_i\}$ formada por vectores propios de N con valores propios asociados $\{\lambda_i\}$ tales que

$$N = \sum_i \lambda_i \left|w_i\right\rangle \left\langle w_i\right| . \quad \textbf{QED}$$

Observemos que la expresión anterior equivale a decir que $N = \sum_i \lambda_i P_i$ donde P_i es la proyección sobre el *espacio propio* asociado a λ_i.

Aquí tenemos una consecuencia del teorema espectral.

Teorema 2 *(Diagonalización simultánea). Los operadores Hermitianos $A, B \in \mathcal{B}(\mathcal{H})$ conmutan si, y sólo si, existe una base ortonormal de $\mathcal{H}$ con respecto a la cual A y B son diagonalizables simultáneamente.*

Demostración. Demostremos la afirmación directa ya que la recíproca es inmediata.
Sea $A = \sum_a a P_a$ donde los valores propios $\{a\}$ son distintos entre sí y P_a es la proyección sobre el *espacio propio* E_a asociado con a.
Observemos que

$$B\,(E_a) \subseteq E_a$$

pues si $x \in E_a$ se cumple $ABx = BAx = Bax = aBx$, así la restricción $B_a \equiv P_a B P_a$ de B a E_a está bien definida y es *Hermitiana*, por ende E_a tiene una base ortonormal de vectores propios de B_a (note que también son vectores propios de A).

Veamos ahora que cada vector propio $u \in E_a$ de B_a también lo es de B, en efecto, si $B_a u = bu$ tenemos que $bu = B_a u = P_a B P_a u = P_a B u = Bu$ pues

$Bu \in E_a$ por la primera observación. Así, por cada E_a hay una base ortonormal de vectores propios de A y B. La base buscada para $\mathcal{H}$ es la unión de dichas bases. **QED**

Antes de proseguir, conviene destacar que el Teorema 2 no es consecuencia necesariamente del teorema espectral, de hecho es un caso especial del siguiente teorema del Álgebra Lineal, ver [12] :

Sean U y T operadores lineales de V en V, donde V es un espacio vectorial de dimensión finita. Si U y T son diagonalizables, las siguientes condiciones son equivalentes:
a) $TU = UT$
b) T y U son simultáneamente diagonalizables.

1.3 Operadores positivos

Definición. Un operador $P \in \mathcal{B}(\mathcal{H})$ es **positivo** o **positivo semi-definido**, si

$$\forall x \in \mathcal{H}, \langle Px, x\rangle \geq 0$$

Esto se simboliza con $P \geq 0$. Si además $\langle Px, x\rangle = 0$ sólo es posible con $x = 0$ diremos que P es un operador *positivo-definido* ó *estrictamente positivo* y escribiremos $P > 0$. También definimos $A \geq B$ si $A - B \geq 0$.

Ejemplos.

a) $\forall u \in \mathcal{H}, |u\rangle \langle u|$ es positivo.

b) $\forall T \in \mathcal{B}(\mathcal{H}), T^*T$ y TT^* son positivos y $T \geq 0 \Leftrightarrow T^* \geq 0$

c) Si $T, S \in \mathcal{B}(\mathcal{H})$ son operadores positivos, entonces $\lambda S + T$ es un operador positivo si λ es un real no negativo.

Proposición 3 *Si $P \geq 0$, entonces P es Hermitiano.*

Demostración. Todo operador T se expresa como $T = A + iB$ con A,B *Hermitianos* por ejemplo,

$$A = \frac{1}{2}\left(T + T^*\right) \text{ y } B = \frac{1}{2i}\left(T - T^*\right)$$

y en la expresión para P, se tiene que $B = 0$. **QED**

Proposición 4 *Sea $P \in \mathcal{B}(\mathcal{H})$ un operador normal y sea $\{\lambda_i\}$ su espectro, entonces P es un operador positivo si, y sólo, si $\lambda_i \geq 0$ para toda i.*

Demostración. Sea $P = \sum_i \lambda_i |w_i\rangle \langle w_i|$ su descomposición espectral.
La afirmación directa es inmediata de que $\lambda_i = \langle w_i, Pw_i \rangle$. Recíprocamente, sea $x \in \mathcal{H}$ y $x = \sum_i a_i w_i$ su expresión en términos de la base ortonormal $\{w_i\}$, luego $\langle x, Px \rangle = \sum_i \lambda_i |a_i|^2$, por lo cual se cumple la afirmación recíproca. **QED**

Se puede reenunciar la proposición anterior de la siguiente forma :

Proposición 5 *(Descomposición espectral para operadores positivos). Para todo operador positivo $P \in \mathcal{B}(\mathcal{H})$ existe una base ortonormal de $\mathcal{H}$ formada con vectores propios de P, $\{w_i\}$ con valores propios correspondientes $\{\lambda_i\}$ tales que*

$$P = \sum_i \lambda_i |w_i\rangle \langle w_i| \text{ con } \lambda_i \geq 0.$$

1.4 Cálculo funcional

Dada una función f de valores complejos con dominio en algún subconjunto de los complejos que contenga el espectro de un operador normal N usando su *descomposición espectral* se define un operador normal asociado con f como

$$f(N) = \sum_i f(\lambda_i) |w_i\rangle \langle w_i| .$$

Este procedimiento se usa para definir la raíz cuadrada de un operador positivo, la exponencial de un operador normal, el logaritmo de un operador positivo definido (o estrictamente positivo), ó el operador inverso de uno estrictamente positivo.

Por ejemplo, dado un operador A sabemos que $AA^* \geq 0$ y $A^*A \geq 0$ por lo que $\sqrt{AA^*}$ y $\sqrt{A^*A}$ están bien definidos y son operadores positivos. El último recibe el nombre de *valor absoluto* de A y se simboliza con

$$|A| = \sqrt{A^*A}$$

1.5 Descomposición polar de un operador

Teorema 6 *Sea A un operador lineal en el espacio de Hilbert $\mathcal{H}$. Entonces existen un operador **unitario** U y un operador positivo $K \equiv \sqrt{AA^*}$ tales que*

$$A = U|A| = KU,$$

además $|A|$ y K son los únicos operadores positivos que satisfacen estas igualdades, y si A es invertible U es único.

A la primera expresión se le conoce comúnmente como la *descomposición polar*, mientras que a la segunda algunas veces se le llama la "*descomposición polar derecha*".

Demostración. Sea $\sum_i \lambda_i |w_i\rangle \langle w_i|$, $(\lambda_i \geq 0)$ la descomposición espectral de $J \equiv |A|$.
Si definimos $\psi_i = Aw_i$, se cumple que $\langle \psi_i, \psi_i \rangle = \lambda_i^2$.

Para aquellos i tales que $\lambda_i > 0$ defínase $\varphi_i = \psi_i / \lambda_i$. Vemos que satisfacen $\langle \varphi_i, \varphi_j \rangle = \langle w_i, A^* A w_j \rangle / \lambda_i \lambda_j = \langle w_i, J^2 w_j \rangle / \lambda_i \lambda_j = \delta_{ij}$ por lo que son un conjunto ortonormal que podemos extender a una base ortonormal que seguimos simbolizando con φ_i.

Consideremos el operador unitario $U \equiv \sum_i |\varphi_i\rangle \langle w_i|$, luego si $\lambda_i > 0$, tenemos que $UJw_i = \lambda_i \varphi_i = \psi_i = Aw_i$ y cuando $\lambda_i = 0$, $UJw_i = 0 = \psi_i$ lo que prueba que A y UJ coinciden en los básicos w_i y en consecuencia son iguales.

J es único pues al multiplicar por $A^* = JU^*$ en el lado izquierdo de $A = UJ$ obtenemos $J^2 = A^* A$ de donde $J = \sqrt{A^* A}$. Una pequeña reflexión nos convence que J es invertible si A es invertible, así U está determinado por la igualdad $U = AJ^{-1}$.

Para ver la descomposición polar derecha, $A = UJ = UJU^* U = KU$ donde $K \equiv UJU^*$ y puesto que $AA^* = KUU^* K = K^2$ debemos tener $K = \sqrt{AA^*}$.
QED

1.5.1 Descomposición usando valores singulares

Lema 7 *(Descomposición usando valores singulares). Si A es una matriz cuadrada, entonces existen matrices unitarias U, V y una matriz diagonal D con entradas no negativas tales que $A = UDV$.*

Los elementos de la diagonal de D son llamados los *valores singulares* de A .

Demostración. Consideremos la descomposición espectral

$$|A| = \sum_i \lambda_i |w_i\rangle\langle w_i|,$$

donde $\{w_i\}$ es una base ortonormal. Sea V^* la matriz de cambio de la base estándar β a la base $\{w_i\}$, entonces $|A| = V^* DV$ donde D es la representación de $|A|$ en la base $\{w_i\}$. Luego usando la *descomposición polar*, existe una matriz unitaria R tal que $A = R\,|A|$ concluimos que $A = (RV^*)\,DV$ es la descomposición buscada. **QED**

1.6 La traza de un operador

Otra función de operadores importante es la *traza* de un operador. Como sabemos, cuando A es una matriz cuadrada su traza es la suma de los elementos de la diagonal $tr\,(A) = \sum_i A_{ii}$, vamos a reconsiderar esta función en la siguiente,

Definición. La transformación $tr : \mathcal{B}(\mathcal{H}) \longrightarrow \mathbb{C}$ definida por $trA = \sum_n \langle \varphi_n, A\varphi_n \rangle$ donde $\{\varphi_n\}$ es cualquier base ortonormal de $\mathcal{H}$ se llama la traza de A.

Veamos que la definición no depende de la base elegida: sea $\{\psi_m\}$ otra base ortonormal, luego

$$\langle \varphi_n, A\varphi_n \rangle = \sum_m \langle \varphi_n, \psi_m \rangle \langle \psi_m, A\varphi_n \rangle = \sum_m \langle A^*\psi_m, \varphi_n \rangle \langle \varphi_n, \psi_m \rangle$$

y por tanto $\sum_n \langle \varphi_n, A\varphi_n \rangle = \sum_n \sum_m \langle A^*\psi_m, \varphi_n \rangle \langle \varphi_n, \psi_m \rangle$

$$= \sum_m \sum_n \langle A^*\psi_m, \varphi_n \rangle \langle \varphi_n, \psi_m \rangle = \sum_m \langle A^*\psi_m, \psi_m \rangle = \sum_m \langle \psi_m, A\psi_m \rangle\,,$$

el intercambio en el orden de las sumas se puede realizar porque son finitas.

Proposición 8 *La traza de un operador tiene las siguientes propiedades*

a) $tr : \mathcal{B}(\mathcal{H}) \longrightarrow \mathbb{C}$, es una función lineal continua con la norma de operadores.

b) $trT^* = \overline{trT}$

c) $trTS = trST$ cuando $S \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2), T \in \mathcal{B}(\mathcal{H}_2, \mathcal{H}_1)$ (**la traza es cíclica**).

d) $trUTU^{-1} = trT$ para cualquier operador invertible U. En particular cuando U es unitario.

e) Si $0 \leq T$ entonces $0 \leq trT$

f) $trA = \sum_i A_{ii}$, para todo $A \in \mathcal{B}(\mathcal{H})$.

g) trT = suma de valores propios de T incluyendo multiplicidades.

h) $tr\ A\,|u\rangle \langle v| = \langle v, Au \rangle$ para todos $u, v \in \mathcal{H}$.

Demostración.
Las propiedades b), e) y f) son inmediatas.

a) La linealidad es inmediata. Para probar la continuidad, supongamos que $n = \dim \mathcal{H}$ y sea λ_i un valor propio de T asociado al vector propio φ_i, que podemos suponer de norma 1, entonces

$|\lambda_i|^2 = \langle T\varphi_i, T\varphi_i\rangle = \|T\varphi_i\|^2 \leq (\|T\| \, \|\varphi_i\|)^2$ de donde $|\lambda_i| \leq \|T\|$ ahora usamos g) y tenemos $|trT| = |\sum_i \lambda_i| \leq \sum_i |\lambda_i| \leq n \, \|T\|$ de donde se sigue la continuidad.

c) Sean $\{e_n\}, \{f_m\}$ bases ortonormales de $\mathcal{H}_1$ y de $\mathcal{H}_2$ respectivamente, luego $\langle f_m, STf_m\rangle_2 = \ \langle S^* f_m, Tf_m\rangle_1 = \sum_n \ \langle S^* f_m, e_n\rangle_1 \ \langle e_n, Tf_m\rangle_1$ y por tanto

$trST = \sum_m \sum_n \ \langle S^* f_m, e_n\rangle_1 \ \langle e_n, Tf_m\rangle_1 = \sum_n \sum_m \ \langle T^* e_n, f_m\rangle_2 \ \langle f_m, Se_n\rangle_2 =$

$\sum_n \ \langle T^* e_n, Se_n\rangle_2 = \sum_n \ \langle e_n, TSe_n\rangle_1 = trTS.$

d) El resultado es inmediato de c).

g) Se usa f) y por d) podemos usar cualquier representación matricial, pues bien, usemos la descomposición canónica de *Jordan.*

h) Sea v distinto de cero. No perdemos generalidad si suponemos que tiene norma 1. Ahora completamos v a una base ortonormal digamos $\{v = e_1, e_2..., e_n\}$ y entonces $tr \ A\,|u\rangle\,\langle v| = \sum_i \langle e_1, e_i\rangle \langle e_i, Au\rangle = \langle e_1, Au\rangle$. **QED**

1.7 Productos tensoriales

Mediante el producto tensorial podemos usar espacios de *Hilbert* para formar espacios de *Hilbert* mas grandes. Supongamos que V *y* W son espacios de *Hilbert* de dimensión n y m respectivamente entonces $V \otimes W$ es un espacio vectorial de dimensión nm.

Para justificar la existencia del producto tensorial y sus propiedades necesitamos la siguiente discusión.

Proposición 9 *(Construcción del espacio vectorial libre con base $\mathcal{S}$). Dado un conjunto finito $\mathcal{S} = \{s_1,,, s_n\}$, existe un espacio vectorial $\mathcal{T}$ sobre $\mathbb{K}$ que tiene como base a $\{1s_1, ..., 1s_n\}$.*

Demostración. Definamos para cada $s_i \in \mathcal{S}$ y cada $c \in \mathbb{K}$, el símbolo cs_i como la función $cs_i : \mathcal{S} \longrightarrow \mathbb{K}$ tal que $cs_i\,(s_j) = c\delta_{ij}$. De esto es claro que $\forall a \in \mathbb{K}, \ a\,(cs_i) = (ac)\,s_i$ y $(c + a)\,s_i = as_i + as_i$.
Sea $\mathcal{T}$ el espacio vectorial de todas las combinaciones lineales formales

$$c_1 s_1 + c_2 s_2 + ... + c_n s_n \ ,$$

es decir $\mathcal{T}$ consta de todas las funciones de $\mathcal{S}$ en $\mathbb{K}$ que se pueden escribir en la forma $\sum_{i=1}^n c_i s_i$ para algunos $c_i \in \mathbb{K}$ por lo que queda claro que es un espacio vectorial sobre $\mathbb{K}$.

El espacio $\mathcal{T}$ tiene como base a las funciones $\{1s_1, ..., 1s_n\}$ pues éstas forman un subconjunto linealmente independiente ya que si $\sum_{i=1}^{n} c_i 1s_i = 0$ al evaluar en el elemento s_j la función del lado izquierdo obtenemos

$$\sum_{i=1}^{n} c_i 1s_i\,(s_j) = \sum_{i=1}^{n} c_i 1\delta_{ij} = c_j,$$

es decir $c_j = 0$ para cada j. **QED**

1.7.1 El Producto tensorial de espacios vectoriales

Sean V, W y U espacios vectoriales sobre un campo $\mathbb{K}$.

Definición. $G : V \times W \longrightarrow U$ es una **función bilineal**, si son lineales las aplicaciones

$$\forall v \in V\ ,\quad w \longmapsto G\,(v, w) \quad \text{y} \quad \forall w \in V\ ,\quad v \longmapsto G\,(v, w)$$

Teorema 10 *(Existencia del Producto tensorial y propiedades básicas). Si V y W son espacios vectoriales de dimensión n y m respectivamente sobre un campo $\mathbb{K}$, entonces existe un espacio vectorial $V \otimes W$ de dimensión nm y una función bilineal $\tau : V \times W \longrightarrow V \otimes W$, tal que $(v, w) \overset{\tau}{\longmapsto} v \otimes w$ satisface las siguientes propiedades:*

a) El producto de $v \in V$ y $w \in W$ se denota por $v \otimes w$ y satisface las siguientes relaciones:

Si v_1, $v_2 \in V$ y $w \in W$,

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$$

Si w_1, $w_2 \in W$ y $v \in V$,

$$v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$$

Si c es un escalar arbitrario y $w_2 \in W$, $v \in V$,

$$c\,(v \otimes w) = (cv) \otimes w = v \otimes (cw)$$

b) Todo elemento del producto tensorial se puede escribir en la forma $\sum_i c_i v_i \otimes u_i$ para algunos escalares c_i, y algunos $v_i \in V, u_i \in W$.

c) Si $\{v_1, ..., v_n\}$ y $\{w_1, ..., w_m\}$ son bases de V y W respectivamente, entonces los elementos $v_i \otimes w_j$, $i = 1, ..., n$ y $j = 1, ..., m$ constituyen una base de $V \otimes W$.

d) *Propiedad Universal.* Si U es un espacio vectorial sobre $\mathbb{K}$ y $\mathcal{L} : V \times W \longrightarrow U$ es una función bilineal, entonces existe una *única aplicación lineal* $L : V \otimes W \longrightarrow U$ tal que $\mathcal{L} = L \circ \tau$ es decir

$$\mathcal{L}(v, w) = L(u \otimes v) \quad \forall v \in V,\ \forall w \in V.$$

Prueba de a).
Sean $\{v_1, ..., v_n\}$ y $\{w_1, ..., w_m\}$ bases de V y W y sea $\mathcal{S} = \{s_{ij}\}_{ij}$ un conjunto cualquiera con nm elementos. Definimos $V \otimes W$ como el espacio vectorial libre sobre $\mathbb{K}$ con base $\mathcal{S}$. Es decir $V \otimes W$ es un espacio vectorial de dimensión nm y consta de todas las combinaciones lineales formales de elementos s_{ij} con coeficientes en $\mathbb{K}$, o sea expresiones de la forma $\sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} s_{ij}$ con $c_{ij} \in \mathbb{K}$.

Si $v = \sum_{i=1}^{n} x_i v_i$ y $w = \sum_{j=1}^{m} y_j w_j$ son las expresiones de v y w en términos de los básicos se define su producto tensorial como el elemento

$$v \otimes w = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j s_{ij}$$

en particular $v_i \otimes w_j = s_{ij}$

Probemos que $(v, w) \overset{\tau}{\longmapsto} v \otimes w$ es bilineal. Sea $v' = \sum_{i=1}^{n} x_i' v_i$ y sean v y w como antes, entonces $cv +\ v' = \sum_{i=1}^{n} (cx_i + x_i') v_i$, luego por definición $(cv + v') \otimes w = \sum_i \sum_j (cx_i + x_i') y_j s_{ij} = \sum_i \sum_j cx_i y_j s_{ij} + \sum_i \sum_j x_i' y_j s_{ij} = cv \otimes w + v' \otimes w$. La prueba de la linealidad en el otro lado es semejante.

Prueba de b).
Con la notación anterior, sabemos que un elemento arbitrario del producto tensorial es de la forma $\xi = \sum_i \sum_j c_{ij} v_i \otimes w_j$, reescribiéndolo tenemos que $\xi = \sum_i v_i \otimes \left(\sum_j c_{ij} w_j \right)$ lo que muestra el resultado.

Prueba de c).
Sean $\{v_1', ..., v'_n\}$ y $\{w_1', ..., v_m'\}$ bases de V y W respectivamente, si $v = \sum_i x_i v_i'$ y $w = \sum_j y_j w_j'$ son las expresiones en términos de los básicos de v y w, entonces $v \otimes w = \sum_i \sum_j x_i y_j \left(v_i' \otimes w_j' \right)$ es decir los nm elementos $v_i' \otimes w_j'$ generan al espacio vectorial $V \otimes W$ de dimensión nm y por tanto, necesariamente son linealmente independientes lo que prueba que forman una base.

Prueba de d).
Sea $\mathcal{L} : V \times W \longrightarrow U$ una función bilineal. Definiendo $L(v_i \otimes w_j) = \mathcal{L}(v_i, w_j)$, sabemos que existe una *única* trasformación lineal $L : V \otimes W \longrightarrow U$. Además, si $v \in V, w \in W$ son elementos cualesquiera expresados (como antes) en términos de los básicos, entonces

$\mathcal{L}(v, w) = \mathcal{L}\left(\sum_i x_i v_i, \sum_j y_j w_j \right) = \sum_i \sum_j x_i y_j \mathcal{L}(v_i, w_j) = L(v \otimes w)$.

1.7.2 Producto tensorial de operadores

Como aplicación de la propiedad universal del producto tensorial se puede justificar que, salvo isomorfismo de espacios vectoriales, el producto tensorial es único. También se usa para construir el producto tensorial de operadores como sigue.

Supongamos que V, W son espacios vectoriales de dimensión finita y que $A : V \longrightarrow V$, $B : W \longrightarrow W$ son operadores lineales. Definimos $\mathcal{L} : V \times W \longrightarrow V \otimes W$ como $\mathcal{L}(v, w) = Av \otimes Bw$, es inmediato verificar que $\mathcal{L}$ es una función bilineal, entonces existe una única función lineal $L : V \otimes W \longrightarrow V \otimes W$ tal que $L(v \otimes w) = \mathcal{L}(v, w) = Av \otimes Bv$.
Se denota a L como $A \otimes B$ es decir, $A \otimes B$ es el único operador lineal tal que

$$\forall v \in V, \ \forall w \in V, \quad (A \otimes B)(v \otimes w) = Av \otimes Bw.$$

Sea $\mathcal{L}_a(V)$ el espacio vectorial formado por todas las transformaciones lineales de V en V. Observemos que se cumple que $\mathcal{L}_a(V) \otimes \mathcal{L}_a(W) \subset \mathcal{L}_a(V \otimes W)$ y como el subespacio vectorial de la izquierda tiene la misma dimensión que el de la derecha, vemos que en el contexto de dimensionalidad finita se cumple que

$$\mathcal{L}_a(V) \otimes \mathcal{L}_a(W) = \mathcal{L}_a(V \otimes W).$$

Por 2) del teorema anterior esto significa que todo operador lineal $L : V \otimes W \longrightarrow V \otimes W$ se puede expresar como una combinación lineal de productos tensoriales de operadores lineales que van de V en V con operadores lineales de W en W

$$L = \sum_i c_i A_i \otimes B_i.$$

1.7.3 Producto tensorial de espacios de Hilbert

Cuando $(\mathcal{H}_1, \langle \cdot, \cdot \rangle_1)$ y $(\mathcal{H}_1, \langle \cdot, \cdot \rangle_2)$ son espacios de *Hilbert* de dimensión n *y* m respectivamente, la construcción anterior nos proporciona el espacio vectorial $\mathcal{H}_1 \otimes \mathcal{H}_2$. Usando los productos internos de cada espacio se define un producto interno en $\mathcal{H}_1 \otimes \mathcal{H}_2$ de la siguiente forma:

Sean $\{v_1, ..., v_n\}$ y $\{w_1, ..., w_m\}$ bases de $\mathcal{H}_1$ y $\mathcal{H}_2$ respectivamente, si $v = \sum_i x_i v_i$, $v' = \sum_r x'_r v_r$, $w = \sum_j y_j w_j$ y $w' = \sum_s y'_s w_s$ son las expresiones en términos de los básicos, entonces se define el producto interno de $v \otimes w = \sum_i \sum_j x_i y_j (v_i \otimes w_j)$ y $v' \otimes w' = \sum_r \sum_s x'_r y'_s (v_r \otimes w_s)$ como

$$\left\langle \sum_i \sum_j x_i y_j (v_i \otimes w_j), \sum_r \sum_s x'_r y'_s (v_r \otimes w_s) \right\rangle_\otimes \equiv$$
$$\sum_{i,j,r,s} \overline{x_i y_j} x'_r y'_s \langle v_i, v_r \rangle_1 \langle w_j, w_s \rangle_2$$

en particular $\langle c_1(v_1 \otimes w_1), c_2(v_2 \otimes w_2)\rangle_{\otimes} = \overline{c_1}c_2 \langle v_1, v_2\rangle_1 \langle w_1, w_2\rangle_2$. Se puede verificar que esta función satisface los axiomas de un producto interno.

Con respecto a la norma inducida por $\langle \cdot, \cdot\rangle_{\otimes}$, $\mathcal{H}_1 \otimes \mathcal{H}_2$ es un espacio de *Hilbert* ya que por ser un espacio de dimensión finita es completo, además si $\{\varphi_i\}, \{\psi_j\}$ son bases ortonormales de $\mathcal{H}_1$ y $\mathcal{H}_2$ respectivamente, entonces $\{\varphi_i \otimes \psi_j\}$ es una base ortonormal de $\mathcal{H}_1 \otimes \mathcal{H}_2$.

La construcción anterior se generaliza cuando aparecen involucrados más de dos espacios de *Hilbert* y se cumplen las siguientes propiedades del producto tensorial de operadores.

Proposición 11 *Sean $\mathcal{H}_i$, $i = 1, 2, ..., n$ espacios de Hilbert de dimensión finita. Si $S = \otimes_{i=1}^n S_i$ y $T = \otimes_{i=1}^n T_i$, donde $S_i : \mathcal{H}_i \longrightarrow \mathcal{H}_i$, $T_i : \mathcal{H}_i \longrightarrow \mathcal{H}_i$, entonces*

a) $T^* = \otimes_i T_i^*$ y $ST = \otimes_i S_i T_i$

b) Si cada T_i tiene un inverso, T tiene inverso y $T^{-1} = \otimes_i T_i^{-1}$

c) T es auto-adjunto, unitario, normal o proyección si cada T_i es auto-adjunto, unitario, normal o proyección.

d) S es positivo si cada S_i es positivo

f) Si $T_i = |u_i\rangle\langle v_i|$ donde $u_i, v_i \in \mathcal{H}_i$, entonces

$$T = |u_1 \otimes u_2 \otimes \cdots \otimes u_n\rangle\langle v_1 \otimes v_2 \otimes \cdots \otimes v_n|$$

Demostración. Se sigue directamente de las definiciones y la omitimos.

Proposición 12 *Para todos $S \in \mathcal{B}(\mathcal{H}_1)$ y $T \in \mathcal{B}(\mathcal{H}_2)$ se cumple*

$$trS \otimes T = trS\ trT.$$

Demostración. Si $\{v_1, ..., v_n\}$ y $\{w_1, ..., w_m\}$ son bases ortonormales de $\mathcal{H}_1$ y $\mathcal{H}_2$ respectivamente, $\{v_i \otimes w_j\}$ es una base ortonormal, luego

$$\begin{aligned}
\sum_{i,j} \langle v_i \otimes w_j, S \otimes T(v_i \otimes w_j)\rangle &= \sum_{i,j} \langle v_i \otimes w_j, Sv_i \otimes Tw_j\rangle \\
&= \sum_{i,j} \langle v_i, Sv_i\rangle \ \langle w_j, Tw_j\rangle \\
&= \sum_{i} \langle v_i, Sv_i\rangle \sum_{j} \ \langle w_j, Tw_j\rangle . \quad \textbf{QED}
\end{aligned}$$

1.8 Resultados auxiliares

A continuación discutiremos dos lemas que serán de utilidad en los capítulos restantes, la razón de incluirlos aquí es que son similares en contenido y uno de ellos es una propiedad referente al producto tensorial de operadores positivos.

Definición. Sea $\varrho \geq 0$ y $\varrho = \sum_i \lambda_i |u_i\rangle \langle u_i|$ su descomposición espectral, entonces definimos el operador $\log \varrho = \sum_i \widehat{\lambda}_i |u_i\rangle \langle u_i|$ donde

$$\widehat{\lambda}_i = \begin{cases} \log \lambda_i & si \quad \lambda_i > 0 \\ 0 & si \quad \lambda_i = 0 \end{cases}$$

La diferencia fundamental con $\log \rho$ cuando $\rho > 0$ es que este último es invertible mientras que el que acabamos de definir no lo es en general.

Lema 13 *Sea λ un número real y ϱ un operador en $\mathcal{H}$. Si $\lambda > 0$ y $\varrho > 0$, entonces*

$$\log (\lambda \rho) = \log \lambda I + \log \varrho.$$

Demostración. Si $\varrho = \sum_i \mu_i |u_i\rangle \langle u_i|$ es la descomposición espectral de ϱ, es claro que $\lambda \varrho = \sum_i \lambda \mu_i |u_i\rangle \langle u_i|$ es la descomposición espectral de $\lambda \rho$ y por ende

$\log \lambda \varrho = \sum_i \log (\lambda \mu_i) |u_i\rangle \langle u_i| = \sum_i (\log \lambda + \log \mu_i) |u_i\rangle \langle u_i| =$

$\log \lambda \sum_i |u_i\rangle \langle u_i| + \sum_i \log \mu_i |u_i\rangle \langle u_i| = \log \lambda I + \log \varrho.$ **QED**

Lema 14 *Si $\varrho, \sigma > 0$ con $\varrho \in \mathcal{B}(\mathcal{H}_1)$ y $\sigma \in \mathcal{B}(\mathcal{H}_2)$, entonces*

$$\log (\varrho \otimes \sigma) = (\log \varrho) \otimes I_2 + I_1 \otimes (\log \sigma)$$

Demostración. Si $\sum_i \lambda_i |u_i\rangle \langle u_i|$, $\sum_j \eta_j |v_j\rangle \langle v_j|$ son descomposiciones espectrales de ϱ y σ respectivamente, entonces $\sum_{i,j} \lambda_i \eta_j |u_i \otimes v_j\rangle \langle u_i \otimes v_j|$ es la descomposición espectral de $\varrho \otimes \sigma$ y por tanto

$$\log (\varrho \otimes \sigma) = \sum_{i,j} \log (\lambda_i \eta_j) |u_i \otimes v_j\rangle \langle u_i \otimes v_j|,$$

pero esta suma es igual a

$\sum_{i,j} \log \lambda_i |u_i\rangle \langle u_i| \otimes |v_j\rangle \langle v_j| + \sum_{i,j} \log \eta_j |u_i\rangle \langle u_i| \otimes |v_j\rangle \langle v_j| =$

$\left(\sum_i \log \lambda_i |u_i\rangle \langle u_i|\right) \otimes \left(\sum_j |v_j\rangle \langle v_j|\right) + \left(\sum_i |u_i\rangle \langle u_i|\right) \otimes \left(\sum_j \log \eta_j |v_j\rangle \langle v_j|\right) =$

$(\log \varrho) \otimes I_2 + I_1 \otimes (\log \sigma)$. **QED**

Capítulo 2

Representación de Lindblad

2.1 La traza parcial

Advertencia sobre la notación.
Sean $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ espacios de *Hilbert*. Cuando se trata con productos tensoriales, para hacer menos pesada la notación se acostumbra simbolizar $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ con $\mathcal{H}_{123}$, y ρ_{123} simboliza un operador lineal de $\mathcal{H}_{123}$ en $\mathcal{H}_{123}$. Si en un mismo enunciado aparecen $\varrho_{123} \in \mathcal{B}(\mathcal{H}_{123})$, $\varrho_{12} \in \mathcal{B}(\mathcal{H}_{12})$, $\varrho_2 \in \mathcal{B}(\mathcal{H}_2)$ etc, significa que $\varrho_2 = tr_{13}\varrho_{123}$, $\varrho_{12} = tr_3\varrho_{123}$, etc. (el símbolo $tr_{13}(\cdot)$, etc. se definirá en el siguiente párrafo).
También se acostumbra abusar de la notación y suprimir la identidad en productos tensoriales de operadores, *e.g.* $\log \varrho_{12} \in \mathcal{B}(\mathcal{H}_{123})$ significa $(\log \varrho_{12}) \otimes I_3$. Sin embargo, cuando sea pertinente escribiremos explícitamente los símbolos.

Definición. Sean $\mathcal{H}_1$, $\mathcal{H}_2$ espacios de *Hilbert*, y ϱ_{12} un operador en $\mathcal{H}_{12}$ definimos la traza parcial (en el espacio $\mathcal{H}_1$) de ϱ_{12} como el operador $\varrho_1 \in \mathcal{B}(\mathcal{H}_1)$ tal que dada una base ortonormal $\{e_i\}$ de $\mathcal{H}_2$ se cumple que para cualesquiera $u, v \in \mathcal{H}_1$,

$$\langle u, \varrho_1 v\rangle = \sum_j \langle u \otimes e_j, \varrho_{12}(v \otimes e_j)\rangle, \tag{TP1}$$

el operador traza parcial se simboliza con $\varrho_1 \equiv tr_2\varrho_{12}$.
Nota: análogamente se define $\varrho_2 \equiv tr_1\varrho_{12}$.

Demostraremos a continuación la existencia de la traza parcial. Primero veamos que la suma no depende de la base elegida. Sea $\{\varphi_j\}$ otra base ortonomal de $\mathcal{H}_2$ y recordemos que $trA\,|x\rangle\langle y| = \langle y, Ax\rangle$, luego

$$\sum_j \langle u \otimes \varphi_j, \varrho_{12}(v \otimes \varphi_j)\rangle = \sum_j tr\varrho_{12}\,|v \otimes \varphi_j\rangle\langle u \otimes \varphi_j| =$$

$$\sum_j tr\varrho_{12}\left(|v\rangle\langle u| \otimes |\varphi_j\rangle\langle\varphi_j|\right) = tr\varrho_{12}\left(\sum_j |v\rangle\langle u| \otimes |\varphi_j\rangle\langle\varphi_j|\right) =$$

$tr\varrho_{12}\left[|v\rangle\langle u|\otimes\left(\sum_j|\varphi_j\rangle\langle\varphi_j|\right)\right]=tr\varrho_{12}\left(|v\rangle\langle u|\ \otimes I_2\right)$, es decir

$$\sum_j\langle u\otimes\varphi_j,\varrho_{12}\left(v\otimes\varphi_j\right)\rangle=tr\varrho_{12}\left(|v\rangle\langle u|\ \otimes I_2\right).$$

Pongamos provisionalmente $L_u\left(v\right)=\sum_j\langle u\otimes e_j,\varrho_{12}\left(v\otimes e_j\right)\rangle$, y observemos que para u fijo $L_u\left(\cdot\right)\in\mathcal{B}\left(\mathcal{H}_1,\mathbb{C}\right)$ lo que se ve aplicando la desigualdad de *Schwarz* y usando que la dimensión de $\mathcal{H}_2$ es finita. Aplicando el bien conocido *Lema de Riesz* tenemos que existe un $u'\in\mathcal{H}_1$ tal que $\langle u',v\rangle=L_u\left(v\right)$ para todo $v\in\mathcal{H}_1$, ahora haciendo variar a u podemos definir una función $\sigma:\mathcal{H}_1\to\mathcal{H}_1$ tal que $\sigma u=u'$. Se puede verificar que σ es un operador lineal continuo. Por la forma como fue construido tenemos que $L_u\left(v\right)=\langle\sigma u,v\rangle=\langle u,\sigma^*v\rangle$ definimos $\varrho_1=\sigma^*$ y por tanto ϱ_1 es el único operador lineal tal que $\forall u,v\in\mathcal{H}_1$,

$$\sum_i\langle u\otimes e_j,\varrho_{12}\left(v\otimes e_j\right)\rangle=\langle u,\varrho_1v\rangle\ \ .$$

El operador ϱ_1 es único porque de haber otro con la misma propiedad, digamos ρ_1, tendríamos que $\forall u,v\in\mathcal{H}_1$, $\langle u,\varrho_1v\rangle=\langle u,\rho_1v\rangle$ de donde $\varrho_1=\rho_1$. **QED**

Podríamos haber definido la traza parcial de la siguiente manera.

Proposición 1 *Sea* ϱ_{12} *un operador en* $\mathcal{H}_{12}$ *y* $\varrho_1\in\mathcal{B}\left(\mathcal{H}_1\right)$. *Entonces* $\varrho_1=tr_2\varrho_{12}$ *si, y sólo si, para todo* $x\in\mathcal{B}\left(\mathcal{H}_1\right)$ *se cumple que*

$$tr\varrho_1x=tr\varrho_{12}\left(x\otimes I_2\right)\qquad\text{(TP2)}$$

Demostración. Para la implicación directa, sean $\{e_i\},\{f_j\}$ bases ortonormales de $\mathcal{H}_1$ y $\mathcal{H}_2$ respectivamente, luego

$$\begin{aligned}tr\varrho_{12}\left(x\otimes I_2\right)&=\sum_{i,j}\langle e_i\otimes f_j,\varrho_{12}\left(x\otimes I_2\right)\left(e_i\otimes f_j\right)\rangle\\&=\sum_i\left[\sum_j\langle e_i\otimes f_j,\varrho_{12}\left(xe_i\otimes f_j\right)\rangle\right]\\&=\sum_i\langle e_i,\varrho_1\left(xe_i\right)\rangle=tr\varrho_1x.\end{aligned}$$

Recíprocamente, si ϱ_1 satisface (TP2), para $u,v\in\mathcal{H}_1$, se cumple que
$\langle u,\varrho_1v\rangle=tr\varrho_1\,|v\rangle\langle u|=tr\varrho_{12}\left(|v\rangle\langle u|\otimes I_2\right)$, y como ya vimos en la prueba de la independencia de la base,

$$tr\varrho_{12}\left(|v\rangle\langle u|\ \otimes I_2\right)=\sum_j\langle u\otimes f_j,\varrho_{12}\left(v\otimes f_j\right)\rangle,$$

lo que demuestra (TP1). **QED**

Proposición 2 *(Propiedades de la traza parcial). Para ϱ_{12} un operador en $\mathcal{H}_{12}$ se cumple lo siguiente,*

a) La traza parcial es un *operador lineal que preserva la traza,* es decir, si $\rho_{12}, \gamma_{12} \in \mathcal{B}(\mathcal{H}_{12})$ y $\lambda \in \mathbb{C}$, entonces

$$tr_2(\lambda\rho_{12} + \gamma_{12}) = \lambda tr_2\rho_{12} + tr_2\gamma_{12}.$$

y

$$tr\varrho_{12} = tr\varrho_1 = tr\varrho_2.$$

c) Para todos $\rho \in \mathcal{B}(\mathcal{H}_1), \gamma \in \mathcal{B}(\mathcal{H}_2)$ se satisface que

$$\begin{aligned} tr_2(\rho \otimes \gamma) &= \rho\, tr\gamma\ , \text{ y} \\ tr_1(\rho \otimes \gamma) &= (tr\rho)\,\gamma. \end{aligned}$$

En particular si $\rho_A \in \mathcal{B}(\mathcal{H}_A), \gamma_B \in \mathcal{B}(\mathcal{H}_B)$ y $tr\rho_A = tr\gamma_B = 1$, entonces $tr_A(\rho_A \otimes \gamma_B) = \gamma_B$ y $tr_B(\rho_A \otimes \gamma_B) = \rho_A$.

Antes de iniciar las demostraciones merece la pena comentar, que en proposiciones donde aparece involucrada la traza parcial, algunas veces convendrá utilizar la propiedad (TP1) y otras, su equivalente (TP2) como se ilustrará en las siguientes demostraciones.

Demostración.
a) Sean $\rho_1 = tr_2\rho_{12}$ y $\gamma_1 = tr_2\gamma_{12}$ y $x \in \mathcal{B}(\mathcal{H}_1)$. Por la linealidad de la traza y (TP2) tenemos,

$$\begin{aligned} tr[(\lambda\rho_1 + \gamma_1)x] &= \lambda tr(\rho_1 x) + tr(\gamma_1 x) \\ &= \lambda tr\rho_{12}(x \otimes I_2) + tr\gamma_{12}(x \otimes I_2) \\ &= tr(\lambda\rho_{12} + \gamma_{12})(x \otimes I_2). \end{aligned}$$

b) Al usar (TP2) con $x = I_1$, se obtiene $tr(\varrho_1 I_1) = tr[\varrho_{12}(I_1 \otimes I_2)]$.

c) Verificaremos que se cumple (TP1). Sea $\{f_j\}$ una base ortonormal de $\mathcal{H}_2$ y sean $u, v \in \mathcal{H}_1$ y $\lambda \equiv \sum_j \langle f_j, \gamma(f_j)\rangle = tr\gamma$. Calculando tenemos,

$$\begin{aligned} \sum_j \langle u \otimes f_j, (\rho \otimes \gamma)(v \otimes f_j)\rangle &= \sum_j \langle u \otimes f_j, \rho v \otimes \gamma f_j\rangle \\ &= \sum_j \langle u, \rho v\rangle \langle f_j, \gamma f_j\rangle = \left[\sum_j \langle f_j, \gamma f_j\rangle\right] \langle u, \rho v\rangle \\ &= \langle u, \lambda\rho v\rangle, \end{aligned}$$

y por (TP1) concluimos que $tr_2(\rho \otimes \gamma) = \lambda\rho$ es decir, $tr_2(\rho \otimes \gamma) = \rho\ tr\gamma$. **QED**

Del mismo estilo son las pruebas de las siguientes propiedades.

Lema 3 *Si* $\varrho_{23} \in \mathcal{B}(\mathcal{H}_{23})$ *y* $\varrho_2 = tr_3 \varrho_{23}$, *entonces se cumple que*

$$tr_3 (I_1 \otimes \varrho_{23}) = I_1 \otimes \varrho_2.$$

Demostración. Sean $a, a' \in \mathcal{H}_1$ y $b, b' \in \mathcal{H}_2$ y sea $\{e_i\}$ una base ortonormal de $\mathcal{H}_3$. Calculando tenemos

$$\sum_i \langle a \otimes b \otimes e_i, (I_1 \otimes \varrho_{23}) (a' \otimes b' \otimes e_i) \rangle = \sum_i \langle a \otimes b \otimes e_i, a' \otimes \varrho_{23} (b' \otimes e_i) \rangle =$$

$$\sum_i \langle a, a' \rangle \langle b \otimes e_i, \varrho_{23} (b' \otimes e_i) \rangle = \langle a, a' \rangle \langle b, \varrho_2 b' \rangle = \langle a \otimes b, (I_1 \otimes \varrho_2) (a' \otimes b') \rangle,$$

es decir se cumple que

$$\langle a \otimes b, (I_1 \otimes \varrho_2) (a' \otimes b') \rangle = \sum_i \langle a \otimes b \otimes e_i, (I_1 \otimes \varrho_{23}) (a' \otimes b' \otimes e_i) \rangle.$$

Luego dados $u, v \in \mathcal{H}_{12}$ por b) del Teorema 1.10 podemos expresarlos en la forma $u = \sum_r a_r \otimes b_r$, $v = \sum_s a'_s \otimes b'_s$ y por la sesqui-linealidad del producto interno, obtenemos que

$$\sum_i \langle u \otimes e_i, (I_1 \otimes \varrho_{23}) (v \otimes e_i) \rangle =$$

$$\sum_i \langle (\textstyle\sum_r a_r \otimes b_r) \otimes e_i, (I_1 \otimes \varrho_{23}) ((\textstyle\sum_s a'_s \otimes b'_s) \otimes e_i) \rangle =$$

$$\sum_r \sum_s \sum_i \langle a_r \otimes b_r \otimes e_i, (I_1 \otimes \varrho_{23}) (a'_s \otimes b'_s \otimes e_i) \rangle =$$

$$\sum_r \sum_s \langle a_r \otimes b_r, (I_1 \otimes \varrho_2) (a'_s \otimes b'_s) \rangle = \langle \textstyle\sum_r a_r \otimes b_r, (I_1 \otimes \varrho_{23}) (\textstyle\sum_s a'_s \otimes b'_s) \rangle =$$

$\langle u, (I_1 \otimes \varrho_2) v \rangle$, es decir se tiene que para todas $u, v \in \mathcal{H}_{12}$

$$\langle u, (I_1 \otimes \varrho_2) v \rangle = \sum_i \langle u \otimes e_i, (I_1 \otimes \varrho_{23}) (v \otimes e_i) \rangle. \quad \textbf{QED}$$

Desarrollaremos un resultado auxiliar que ocuparemos en el capítulo 4, pero antes, necesitamos observar que si $\varrho_{123} > 0$ es un operador en $\mathcal{H}_{123}$ y $R \equiv (I_1 \otimes (\varrho_2 + uI) \otimes I_3)^{-1}$ con $u \geq 0$, $u \in \mathbb{R}$, entonces $R = I_1 \otimes R_2 \otimes I_3$, donde $R_2 = (\varrho_2 + uI)^{-1} > 0$.

Proposición 4 *Sea* $\varrho_{123} \in \mathcal{H}_{123}$ *con* $\varrho_{123} > 0$. *Si* $\varrho_2 = tr_{13} \varrho_{123}$, $\varrho_{12} = tr_3 \varrho_{123}$ *y* $\varrho_{23} = tr_1 \varrho_{123}$, *entonces*

$$tr [(\varrho_{12} \otimes I_3) R (I_1 \otimes \varrho_{23}) R] = tr [\varrho_2 R_2 \varrho_2 R_2].$$

Demostración. Sean $\sigma_1 \in \mathcal{B}(\mathcal{H}_1)$ y $\sigma_2 \in \mathcal{B}(\mathcal{H}_2)$, se afirma que

$$tr\left[(\sigma_1 \otimes \sigma_2 \otimes I_3) R (I_1 \otimes \varrho_{23}) R\right] = tr\sigma_1 tr\left[\sigma_2 R_2 \varrho_2 R_2\right].$$

En efecto, si $\{a\}, \{b\}, \{c\}$ son bases ortonormales de $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ respectivamente, entonces

$\sum_{a,b,c} \langle a \otimes b \otimes c, [(\sigma_1 \otimes \sigma_2 \otimes I_3) R (I_1 \otimes \varrho_{23}) R] a \otimes b \otimes c\rangle =$

$\sum_{a,b} \sum_c \langle \sigma_1^* a \otimes R_2 \sigma_2^* b \otimes c, (I_1 \otimes \varrho_{23}) a \otimes R_2 b \otimes c\rangle =$

$\sum_{a,b} \langle \sigma_1^* a \otimes R_2 \sigma_2^* b, (I_1 \otimes \varrho_2) a \otimes R_2 b\rangle =$

$\sum_b \sum_a \langle a \otimes R_2 \sigma_2^* b, (\sigma_1 \otimes I_2)(I_1 \otimes \varrho_2) a \otimes R_2 b\rangle =$

$\sum_b \langle R_2 \sigma_2^* b, (tr\sigma_1) \varrho_2 R_2 b\rangle = tr\sigma_1 \sum_b \langle b, \sigma_2 R_2 \varrho_2 R_2 b\rangle$, observe que hemos usado el lema previo al sumar sobre c y la Proposición 2 inciso c) en la penúltima igualdad.

Ahora escribimos $\varrho_{12} = \sum_i \sigma_1^i \otimes \sigma_2^i$ con algunos $\sigma_1^i \in \mathcal{B}(\mathcal{H}_1)$ y $\sigma_2^i \in \mathcal{B}(\mathcal{H}_2)$; no perdemos generalidad si suponemos que cada σ_1^i tiene traza 1. Por argumentos de linealidad y usando lo anterior concluimos que
$tr\left[(\varrho_{12} \otimes I_3) R (I_1 \otimes \varrho_{23}) R\right] = tr\left[\varrho_2 R_2 \varrho_2 R_2\right]$. **QED**

Lema 5 *Sea $\rho_{12} \in \mathcal{B}(\mathcal{H}_{12})$ y $\gamma \in \mathcal{B}(\mathcal{H}_3)$. Si $\varrho_2 = tr_1 \varrho_{12}$, entonces*

a) Cuando $tr\gamma = 1$,

$$tr_{13}(\rho_{12} \otimes \gamma) = \rho_2$$

b)

$$tr_1(\rho_{12} \otimes \gamma) = \rho_2 \otimes \gamma$$

Demostración.
a) Sea $Y \in \mathcal{B}(\mathcal{H}_2)$. Calculando tenemos que

$tr(\rho_{12} \otimes \gamma)(I_1 \otimes Y \otimes I_3) = tr(\rho_{12}(I_1 \otimes Y) \otimes \gamma) =$

$tr\rho_{12}(I_1 \otimes Y)\ tr\gamma = tr\rho_{12}(I_1 \otimes Y) = tr(\rho_2 Y\ )$,

o sea $tr(\rho_2 Y\ ) = tr(\rho_{12} \otimes \gamma)(I_1 \otimes Y \otimes I_3)$

b) Sea $\{a\}$, una base ortonormal de $\mathcal{H}_1$ y tomemos $b, b' \in \mathcal{H}_2$ y $c, c' \in \mathcal{H}_3$, entonces

$$\begin{aligned}
\sum_a \langle a \otimes b \otimes c, (\rho_{12} \otimes \gamma) a \otimes b' \otimes c'\rangle &= \sum_a \langle a \otimes b \otimes c, \rho_{12}(a \otimes b') \otimes \gamma c'\rangle \\
&= \langle c, \gamma c'\rangle \sum_a \langle a \otimes b, \rho_{12}(a \otimes b')\rangle \\
&= \langle c, \gamma c'\rangle \langle b, \rho_2 b'\rangle \\
&= \langle b \otimes c, (\rho_2 \otimes \gamma) b' \otimes c'\rangle,
\end{aligned}$$

o sea

$$\langle b \otimes c, (\rho_2 \otimes \gamma)\, b' \otimes c' \rangle = \sum_a \langle a \otimes b \otimes c, (\rho_{12} \otimes \gamma)\, a \otimes b' \otimes c' \rangle\,,$$

luego usamos la linealidad del producto interno y el hecho de que podemos escribir $u = \sum_i b_i \otimes c_i$ y $v = \sum_i b'_i \otimes c'_i$ con algunos $b_i, b'_i \in \mathcal{H}_2$ y $c_i, c'_i \in \mathcal{H}_3$ para concluir que $\langle u, (\rho_2 \otimes \gamma)\, v \rangle = \sum_a \langle a \otimes u, (\rho_{12} \otimes \gamma)\, a \otimes v \rangle$ para cualesquiera $u, v \in \mathcal{H}_{23}$. **QED**

2.2 Purificación

Lema 6 *(Purificación). Sea $\mathcal{H}_1$ un espacio de Hilbert. Dado $\varrho_1 \in \mathcal{B}(\mathcal{H}_1)$, con $tr\varrho_1 = 1$ y $\varrho_1 \geq 0$, existen un espacio de Hilbert $\mathcal{H}_{12}$ y $\varrho_{12} \in \mathcal{B}(\mathcal{H}_{12})$ tal que $\varrho_{12} = |\varphi\rangle\langle\varphi|$ para algún $\varphi \in \mathcal{H}_{12}$ con $\|\varphi\| = 1$ y que satisface,*

$$\varrho_1 = tr_2 \varrho_{12}\ .$$

Este resultado se suele expresar diciendo que dado un estado cuántico ϱ_1 (ó una matriz de densidad), existe un estado puro ϱ_{12} (i.e. un operador de la forma $\varrho_{12} = |\varphi\rangle\langle\varphi|$ para algún $\varphi \in \mathcal{H}_{12}$ con $\|\varphi\| = 1$) tal que $\varrho_1 = tr_2 \varrho_{12}$.

Demostración. Escribamos a ϱ_1 en su representación espectral

$$\varrho_1 = \sum_{i=1}^{m} \lambda_i\, |\varphi_i\rangle\langle\varphi_i|\,,$$

donde $\{\varphi_i\}$ es una base ortonormal de $\mathcal{H}_1$ de vectores propios de ϱ_1 asociados a los valores propios $\{\lambda_i\}$. Sea $\mathcal{H}_2$ un espacio de dimensión m con una base ortonormal $\{\psi_j\}$.

Si $\varphi \equiv \sum_{i=1}^{m} \lambda_i^{\frac{1}{2}} \varphi_i \otimes \psi_i$, entonces $\varrho_{12} \equiv |\varphi\rangle\langle\varphi|$ es un estado puro tal que $\varrho_1 = tr_2 \varrho_{12}$, en efecto calculando $\|\varphi\|^2$ tenemos

$$\begin{aligned} \langle \varphi, \varphi \rangle &= \sum_{i,j} \lambda_i^{\frac{1}{2}} \lambda_j^{\frac{1}{2}} \langle \varphi_i \otimes \psi_i, \varphi_j \otimes \psi_j \rangle \\ &= \sum_{i,j} \lambda_i^{\frac{1}{2}} \lambda_j^{\frac{1}{2}} \langle \varphi_i, \varphi_j \rangle \langle \psi_i, \psi_j \rangle = \sum_i \lambda_i = 1. \end{aligned}$$

Y usando la sesqui-linealidad de $|\cdot\rangle\langle\cdot|$ obtenemos,

$$|\varphi\rangle\langle\varphi| = \sum_{i,j} \lambda_i^{\frac{1}{2}} \lambda_j^{\frac{1}{2}}\, |\varphi_i\rangle\langle\varphi_j| \otimes |\psi_i\rangle\langle\psi_j|$$

de tal forma que por la linealidad de la traza parcial,

$tr_2\, |\varphi\rangle\langle\varphi| = \sum_{i,j} \lambda_i^{\frac{1}{2}} \lambda_j^{\frac{1}{2}} tr_2 \left[|\varphi_i\rangle\langle\varphi_j| \otimes |\psi_i\rangle\langle\psi_j| \right] = \sum_{i,j} \lambda_i^{\frac{1}{2}} \lambda_j^{\frac{1}{2}}\, |\varphi_i\rangle\langle\varphi_j|\, tr\left(|\psi_i\rangle\langle\psi_j| \right)$ $= \sum_{i,j} \lambda_i^{\frac{1}{2}} \lambda_j^{\frac{1}{2}}\, |\varphi_i\rangle\langle\varphi_j|\, \langle \psi_j, \psi_i \rangle = \sum_i \lambda_i\, |\varphi_i\rangle\langle\varphi_i|\,.$ **QED**

2.3 Descomposición de Schmidt

Teorema 7 *(La descomposición de Schmidt). Para cualquier vector normalizado $\varphi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ existen conjuntos ortonormales $\{\varphi_i\}$,$\{\psi_j\}$ de $\mathcal{H}_1$ y $\mathcal{H}_2$ respectivamente tales que*

$$\varphi = \sum_i \lambda_i \varphi_i \otimes \psi_i,$$

donde $\{\lambda_i\}$ son números reales no negativos que satisfacen $\sum_i \lambda_i^2 = 1$.

Demostración. Sean $\{e_j\}, \{f_k\}$ bases ortonormales de $\mathcal{H}_1$ y $\mathcal{H}_2$, respectivamente entonces $\{e_j \otimes f_k\}$ es una base ortonormal de $\mathcal{H}_1 \otimes \mathcal{H}_2$ y para algunos escalares a_{jk} podemos escribir

$$\varphi = \sum_{j,k} a_{jk} e_j \otimes f_k. \tag{S1}$$

En caso de que las dimensiones de $\mathcal{H}_1$ y $\mathcal{H}_2$ no coincidan, ponemos algunos $a_{rs} = 0$ para que la matriz $A = (a_{jk})$ sea cuadrada, luego por el lema 1.7 existen matrices unitarias U, V tales que $A = UDV$, donde D es una matriz diagonal con entradas no negativas. Por tanto, el elemento j,k-ésimo de A se puede escribir como

$a_{jk} = (UDV)_{jk} = \sum_i u_{ji} (DV)_{ik} = \sum_i u_{ji} \left(\sum_r d_{ir} v_{rk}\right) = \sum_i u_{ji} d_{ii} v_{ik}$,
sustituyendo esto en **(S1)** podemos escribir

$\varphi = \sum_{i,j,k} u_{ji} d_{ii} v_{ik} e_j \otimes f_k = \sum_i d_{ii} \left(\sum_j u_{ji} e_j\right) \otimes \left(\sum_k v_{ik} f_k\right).$
Definiendo

$$\varphi_i \equiv \sum_j u_{ji} e_j, \ \psi_i \equiv \left(\sum_k v_{ik} f_k\right), \ \lambda_i \equiv d_{ii},$$

y usando que U es unitaria (respec. V), por la ortonormalidad de $\{e_j\}$ (respec. $\{f_k\}$) es fácil verificar que $\{\varphi_i\}$ y $\{\psi_i\}$ son conjuntos ortonormales, y con base es esto, al desarrollar $\langle \varphi, \varphi \rangle$ y usar que φ es un vector de norma 1 se obtiene que $\sum_i \lambda_i^2 = 1$. **QED**

2.4 Operadores completamente positivos

Definición. Un operador **completamente positivo (CP)**, es una función $\Phi : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ de la forma

$$\Phi(\varrho) = \sum_{i=1}^{n} E_i \varrho E_i^*,$$

para algunos $E_i \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$.

Comentemos que esta representación de Φ se llama representación de *Krauss* y que existen otras definiciones equivalentes para un operador CP, ver [6].

Definición. Una función lineal $\Psi : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ **preserva la traza** si para todo $\varrho \in \mathcal{B}(\mathcal{H}_1)$, se cumple que

$$tr\Psi(\varrho) = tr\varrho.$$

A continuación discutimos una clase especial de operadores CP que preservan la traza.

Proposición 8 *La traza parcial es un operador completamente positivo que preserva la traza.*

Demostración. Preserva la traza por la Proposición 2 a).
Para probar la primera parte supongamos que tenemos un sistema conjunto $\mathcal{H}_{12}$ y que deseamos descartar $\mathcal{H}_2$.

Sea $\{v_j\}_{j\in J}$ una base ortonormal de $\mathcal{H}_2$.

Para cada $i \in J$, definimos un operador lineal $E_i : \mathcal{H}_{12} \to \mathcal{H}_1$ por

$$E_i\left(\sum_j \lambda_j q_j \otimes v_j\right) = \lambda_i q_i \quad ,$$

donde $\{\lambda_j\} \subseteq \mathbb{C}$ y $\{q_j\} \subseteq \mathcal{H}_1$. Observemos que cada E_i está bien definido pues cada elemento de $\mathcal{H}_{12}$ se puede escribir como combinación lineal de elementos de la forma $q \otimes h$, con $q \in \mathcal{H}_1$ y $h \in \mathcal{H}_2$ pero cada uno de estos se expresa como $q \otimes h = q \otimes \left(\sum_j \mu_j v_j\right) = \sum_j \mu_j q \otimes v_j$.

Ahora verificaremos que para cada i, el operador adjunto $E_i^* : \mathcal{H}_1 \to \mathcal{H}_{12}$, actúa como

$$E_i^*(q) = q \otimes v_i \quad ,$$

para esto, sean $p, q \in \mathcal{H}_1$ y $y \in \mathcal{H}_2$ expresado en términos de la base $\{v_j\}$ como $y = \sum_j \alpha_j v_j$.
Calculando tenemos,
$\langle q \otimes v_i, p \otimes y\rangle = \sum_j \alpha_j \langle q \otimes v_i, p \otimes v_j\rangle = \sum_j \alpha_j \langle q, p\rangle \langle v_i, v_j\rangle = \alpha_i \langle q, p\rangle \langle v_i, v_i\rangle = \alpha_i \langle q, p\rangle$,

y por otra parte
$\langle q, E_i(p \otimes y)\rangle = \left\langle q, E_i\left(\sum_j \alpha_j p \otimes v_j\right)\right\rangle = \langle q, \alpha_i p\rangle = \alpha_i \langle q, p\rangle$, es decir $\forall\ p, q \in \mathcal{H}_1$ y $\forall y \in \mathcal{H}_2$ se cumple que

$$\langle q \otimes v_i, p \otimes y\rangle = \langle q, E_i(p \otimes y)\rangle .$$

Con base en lo anterior abordemos el caso general; sean $q \in \mathcal{H}_1$ y $\sum_j p_j \otimes y_j \in \mathcal{H}_{12}$, entonces

$\left\langle q, E_i\left(\sum_j p_j \otimes y_j\right)\right\rangle = \sum_j \langle q, E_i\left(p_j \otimes y_j\right)\rangle = \sum_j \langle q \otimes v_i, p_j \otimes y_j\rangle =$

$\left\langle q \otimes v_i, \left(\sum_j p_j \otimes y_j\right)\right\rangle$ y podemos concluir que $E_i^*\left(q\right) = q \otimes v_i$.

Ahora definimos $\Phi : \mathcal{B}\left(\mathcal{H}_{12}\right) \to \mathcal{B}\left(\mathcal{H}_1\right)$, tal que para todo $\gamma \in \mathcal{B}\left(\mathcal{H}_{12}\right)$, $\Phi\left(\gamma\right) = \sum_i E_i \gamma E_i^*$.
Verificaremos que para todo $\varrho \in \mathcal{B}\left(\mathcal{H}_1\right)$ se cumple la igualdad

$$\Phi\left(\varrho \otimes |v_j\rangle\langle v_{j'}|\right) = \varrho\delta_{jj'} = tr_2\left(\varrho \otimes |v_j\rangle\langle v_{j'}|\right) \quad , \tag{R1}$$

en efecto, si $q \in \mathcal{H}_1$,

$\Phi\left(\varrho \otimes |v_j\rangle\langle v_{j'}|\right) q = \sum_i E_i\left(\varrho \otimes |v_j\rangle\langle v_{j'}|\right) E_i^*\left(q\right) =$

$\sum_i E_i\left(\varrho \otimes |v_j\rangle\langle v_{j'}|\right)\left(q \otimes v_i\right) = \sum_i E_i\left[\varrho q \otimes |v_j\rangle\langle v_{j'}|\, v_i\right] =$

$\sum_i \langle v_{j'}, v_i\rangle E_i\left[\varrho q \otimes v_j\right] = E_{_{j'}}\left[\varrho q \otimes v_j\right] = \delta_{jj'}\varrho q$. Esto es

$$\Phi\left(\varrho \otimes |v_j\rangle\langle v_{j'}|\right) q = \delta_{jj'}\varrho q.$$

Por otra parte, sea $x \in \mathcal{H}_1$ arbitrario. Usando la definición de traza parcial se tiene que

$\langle x, tr_2\left(\varrho \otimes |v_j\rangle\langle v_{j'}|\right) q\rangle = \sum_r \langle x \otimes v_r, \left(\varrho \otimes |v_j\rangle\langle v_{j'}|\right) q \otimes v_r\rangle =$

$\sum_r \langle x \otimes v_r, \varrho q \otimes |v_j\rangle\langle v_{j'}|\, v_r\rangle = \sum_r \langle v_{j'}, v_r\rangle \langle x \otimes v_r, \varrho q \otimes v_j\rangle =$

$\langle x \otimes v_{j'}, \varrho q \otimes v_j\rangle = \langle x, \varrho q\rangle \langle v_{j'}, v_j\rangle = \langle x, \delta_{j'j}\varrho q\rangle$.

Por tanto, para toda $x \in \mathcal{H}_1$ se cumple que $\langle x, tr_2\left(\varrho \otimes |v_j\rangle\langle v_{j'}|\right) q\rangle = \langle x, \delta_{j'j}\varrho\left(q\right)\rangle$, y podemos concluir que para todo $q \in \mathcal{H}_1$

$$tr_2\left(\varrho \otimes |v_j\rangle\langle v_{j'}|\right) q = \delta_{j'j}\varrho\left(q\right),$$

lo cual muestra que **(R1)** se cumple.

Como Φ y $tr_2\left(\cdot\right)$ son lineales, **(R1)** implica que son iguales como veremos a continuación.
Dado un elemento arbitrario $u \in \mathcal{H}_2$ lo podemos escribir como combinación lineal en términos de la base $\{v_j\}$ y por ende, usando la sesqui-linealidad de $|\cdot\rangle\langle\cdot|$ todo $|u\rangle\langle u| \in \mathcal{B}\left(\mathcal{H}_2\right)$ se puede escribir como combinación lineal de elementos de la forma $|v_j\rangle\langle v_{j'}|$, entonces por la linealidad de Φ y de $tr_2\left(\cdot\right)$ la relación (R1) también es válida para cualquier $\varrho \in \mathcal{B}\left(\mathcal{H}_1\right)$ y para todo $|u\rangle\langle u|$ con $u \in \mathcal{H}_2$ arbitrario.
Esto último implica que Φ y $tr_2\left(\cdot\right)$ coinciden en elementos de la forma $\varrho \otimes \sigma$ donde ϱ es arbitrario y σ es *Hermitiano* porque por el teorema espectral, σ se puede escribir como combinación lineal de elementos de la forma $|u_i\rangle\langle u_i|$ con $\{u_i\} \subseteq \mathcal{H}_2$.

Finalmente, Φ y $tr_2(\cdot)$ coinciden en elementos de la forma $\varrho \otimes \sigma$ donde ϱ, σ son arbitrarios, pues recordemos que en la prueba de la Proposición 1.3 vimos que todo operador σ se puede escribir como combinación lineal de dos operadores *Hermitianos.* **QED**

2.4.1 Adjunto de un operador CP

Dado un espacio de *Hilbert* $\mathcal{H}$, en $\mathcal{B}(\mathcal{H})$ se define un producto escalar $\langle \cdot, \cdot \rangle_{HS}$ llamado producto interno de *Hilbert-Schmidt* de la siguiente manera

$$\forall S, T \in \mathcal{B}(\mathcal{H}), \quad \langle S, T \rangle_{HS} = trS^*T.$$

Resulta que respecto a este producto interno el adjunto del operador **CP** $\Phi : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$ con representación de *Krauss* $\Phi(x) = \sum_{i=1}^n E_i x E_i^*$, es el operador $\widehat{\Phi} : \mathcal{B}(\mathcal{H}_2) \to \mathcal{B}(\mathcal{H}_1)$ tal que $\widehat{\Phi}(y) = \sum_{i=1}^n E_i^* y E_i$, $\forall y \in \mathcal{B}(\mathcal{H}_2)$ como veremos a continuación:

$\langle y, \Phi x \rangle_{HS} = \langle y, \sum_i E_i x E_i^* \rangle_{HS} = \sum_i \langle y, E_i x E_i^* \rangle_{HS} = \sum_i tr\,(y^* E_i x E_i^*) =$

$\sum_i tr\,(E_i^* y^* E_i x) = \sum_i tr\left[(E_i^* y E_i)^* x\right] = tr\left[(\sum_i E_i^* y E_i)^* x\right] =$

$\langle (\sum_i E_i^* y E_i), x \rangle_{HS}$ y puesto que esta igualdad se cumple $\forall y \in \mathcal{B}(\mathcal{H}_2), \forall x \in \mathcal{B}(\mathcal{H}_1)$ se concluye que

$$\widehat{\Phi}(y) = \sum_{i=1}^n E_i^* y E_i, \quad \forall y \in \mathcal{B}(\mathcal{H}_2).$$

Observe que en el desarrollo anterior hemos usado la ciclicidad de la traza y que están en juego dos productos internos de *Hilbert-Schmidt* que podrían ser distintos, pero que no obstante esto, los hemos simbolizado de la misma manera para evitar cargar aún más la notación.

En la proposición que sigue discutimos dos criterios necesarios y suficientes para que un operador CP preserve la traza.

Proposición 9 *Un operador completamente positivo* $\Phi : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$, *con representación de Krauss* $\Phi(\varrho) = \sum_{i=1}^n E_i \varrho E_i^*$, $E_i \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$, *preserva la traza si, y sólo si, su adjunto respecto del producto interno de Hilbert-Schmidt* $\widehat{\Phi} : \mathcal{B}(\mathcal{H}_2) \to \mathcal{B}(\mathcal{H}_1)$ *tal que* $\widehat{\Phi}(y) = \sum_{i=1}^n E_i^* y E_i$ *para todo* $y \in \mathcal{B}(\mathcal{H}_2)$, *es unital i.e.,* $\widehat{\Phi}(I_2) = I_1$ *ó equivalentemente si, y sólo si,*

$$\sum_{i=1}^n E_i^* E_i = I_1$$

Demostración. Supongamos que preserva la traza, entonces $\forall\varrho \in \mathcal{B}(\mathcal{H}_1)$, $tr\varrho = tr\left(\sum_i E_i\varrho E_i^*\right) = \sum_i tr\left(E_i^* E_i \varrho\right) = tr\left[\left(\sum_i E_i^* E_i\right)\varrho\right]$ donde hemos usado la ciclicidad de la traza. Es decir para todo ϱ se cumple,

$$tr\varrho = tr\left[\widehat{\Phi}(I_2)\,\varrho\right].$$

Sea $\widehat{\Phi}(I_2) = \sum_i \lambda_i |u_i\rangle\langle u_i|$ su descomposición espectral donde $\{u_i\}$ es una base ortonormal de $\mathcal{H}_1$. Como $\widehat{\Phi}(I_2)|u_j\rangle\langle u_j| = \lambda_j |u_j\rangle\langle u_j|$, entonces para cada j, $1 = tr|u_j\rangle\langle u_j| = tr\widehat{\Phi}(I_2)|u_j\rangle\langle u_j| = tr\lambda_j |u_j\rangle\langle u_j| = \lambda_j$, por lo tanto $\widehat{\Phi}(I_2) = I_1$.

Recíprocamente, suponiendo que $\sum_{i=1}^n E_i^* E_i = I_1$, inferimos que $\forall\varrho \in \mathcal{B}(\mathcal{H}_1)$, $tr\varrho = tr\left[\left(\sum_i E_i^* E_i\right)\varrho\right] = \sum_i tr\left(E_i^* E_i \varrho\right) = tr\left(\sum_i E_i \varrho E_i^*\right) = tr\Phi(\varrho)$. **QED**

2.5 La representación de Lindblad

En el siguiente teorema discutimos otra forma de representar a un operador Φ CP que preserva la traza, que generaliza una expresión dada por *Lindblad* en [10], la representación original establece que cuando el dominio y codominio de Φ coinciden, se cumplen los incisos **b)** y **c)** es decir, $\Phi(\varrho) = tr_2 U\left(\varrho \otimes |a\rangle\langle a|\right)U^*$, donde U es un operador unitario. En la versión que presentamos U es sólo una isometría parcial pero en cambio $\Phi : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$, o sea, que el dominio y codominio pueden ser distintos.

Terminamos este comentario destacando que en el inciso **d)** enunciamos otra manera de representar el **adjunto** de un operador CP que preserva la traza que será útil en el capítulo 4.

Teorema 10 *(Representación de Lindblad general) Sea $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ un operador* ***CP*** *que preserva la traza, con representación de Krauss $\Phi(\varrho) = \sum_{i=1}^n E_i\varrho E_i^*$, donde $E_i \in \mathcal{B}(\mathcal{H}_A, \mathcal{H}_B)$ para cada i, entonces existen operadores $V : \mathcal{H}_A \longrightarrow \mathcal{H}_B \otimes \mathbb{C}^n$ y $U : \mathcal{H}_A \otimes \mathbb{C}^n \longrightarrow \mathcal{H}_B \otimes \mathbb{C}^n$ tales que:*

a) U, V son isometrías parciales.

b)

$$V\varrho V^* = U\left(\varrho \otimes |a\rangle\langle a|\right)U^*, \quad \varrho \in \mathcal{B}(\mathcal{H}_A),$$

donde $a \in \mathbb{C}^n$ es un vector normalizado.

c)

$$\Phi(\varrho) = tr_2 V\varrho V^*, \quad \varrho \in \mathcal{B}(\mathcal{H}_A).$$

d) Si $\widehat{\Phi} : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_A)$ es el adjunto de Φ respecto del producto interno de *Hilbert-Schmidt* se cumple que

$$\widehat{\Phi}(\varrho) = V^*\left(\varrho \otimes I_2\right)V, \quad \forall\varrho \in \mathcal{B}(\mathcal{H}_B).$$

Demostración Sea $\beta = \{b\}$ una base ortonormal de $\mathbb{C}^n$ y $a \in \beta$, definimos un operador lineal $U : \mathcal{H}_A \otimes \mathbb{C}^n \longrightarrow \mathcal{H}_B \otimes \mathbb{C}^n$ como $U = 0$ en $[\mathcal{H}_A \otimes a]^\perp$ y por

$$U(\psi \otimes a) = \sum_{b\in\beta} E_b\psi \otimes b, \quad \forall\psi \otimes a \in \mathcal{H}_A \otimes a, \tag{L1}$$

como Φ preserva la traza se cumple que $\sum_{i=1}^n E_i^* E_i = I_1$, y de esto resulta que U es una isometría parcial con espacio inicial $\mathcal{H}_A \otimes a$ como veremos a continuación: Sea $\psi \in \mathcal{H}_A$, luego

$\langle U\psi \otimes a, U\psi \otimes a\rangle = \langle \sum_b E_b\psi \otimes b, \sum_{b'} E_{b'}\psi \otimes b'\rangle =$

$\sum_b \sum_{b'} \langle E_b\psi \otimes b, E_{b'}\psi \otimes b'\rangle = \sum_b \sum_{b'} \langle E_b\psi, E_{b'}\psi\rangle \langle b, b'\rangle = \sum_b \langle E_b\psi, E_b\psi\rangle \langle b, b\rangle$

$= \sum_b \langle E_b\psi, E_b\psi\rangle = \sum_b \langle \psi, E_b^* E_b\psi\rangle = \langle \psi, \sum_b E_b^* E_b\psi\rangle = \langle \psi, \psi\rangle .$

Ahora veremos que $\forall\psi \in \mathcal{H}_A, \forall b \in \{b\}$

$$U^*(\psi \otimes b) = E_b^*\psi \otimes a, \tag{L2}$$

para lo cual sea $\{\varphi\}$ una base de $\mathcal{H}_A$, luego

$\langle \psi \otimes b, U(\varphi \otimes a)\rangle = \langle \psi \otimes b, \sum_{b'} E_{b'}\varphi \otimes b'\rangle = \sum_{b'} \langle \psi \otimes b, E_{b'}\varphi \otimes b'\rangle =$

$\sum_{b'} \langle \psi, E_{b'}\varphi\rangle \langle b, b'\rangle = \langle \psi, E_b\varphi\rangle \langle b, b\rangle = \langle E_b^*\psi, \varphi\rangle \langle a, a\rangle = \langle E_b^*(\psi) \otimes a, \varphi \otimes a\rangle$

esto es

$$\begin{aligned} \langle U^*(\psi \otimes b), \varphi \otimes a\rangle &= \langle \psi \otimes b, U(\varphi \otimes a)\rangle \\ &= \langle E_b^*(\psi) \otimes a, \varphi \otimes a\rangle , \end{aligned}$$

también si $b' \neq a$ se cumple que

$\langle U^*(\psi \otimes b), \varphi \otimes b'\rangle = \langle \psi \otimes b, U(\varphi \otimes b')\rangle = \langle \psi \otimes b, 0\rangle =$

$0 = \langle E_b^*(\psi), \varphi\rangle \langle a, b'\rangle = \langle E_b^*(\psi) \otimes a, \varphi \otimes b'\rangle ,$

es decir que para cualquier elemento $\varphi \otimes b$ de la base $\{\varphi \otimes b\}$ tenemos que

$$\langle U^*(\psi \otimes b), \varphi \otimes b\rangle = \langle E_b^*(\psi) \otimes a, \varphi \otimes b\rangle ,$$

por lo tanto se cumple la igualdad (L2).

Para continuar definimos $j : \mathcal{H}_A \longrightarrow \mathcal{H}_A \otimes \mathbb{C}^n$ mediante $j(h) = h \otimes a$ y afirmamos que su adjunto $j^* : \mathcal{H}_A \otimes \mathbb{C}^n \longrightarrow \mathcal{H}_A$ actúa según la correspondencia

$$j^*(h \otimes v) = \langle a, v\rangle h. \tag{L3}$$

En efecto, sea un elemento arbitrario $v \in \mathbb{C}^n$ expresado en términos de los básicos como $v = \sum_b \alpha_b b$, luego para elementos arbitrarios $h, h_2 \in \mathcal{H}_A$, tenemos que

$$\langle h \otimes v, j(h_2)\rangle = \langle h \otimes v, h_2 \otimes a\rangle = \langle h, h_2\rangle \langle v, a\rangle = \langle h, h_2\rangle \left\langle \textstyle\sum_b \alpha_b b, a\right\rangle =$$

$$\langle h, h_2\rangle \textstyle\sum_b \overline{\alpha_b} \langle b, a\rangle = \overline{\alpha_a} \langle h, h_2\rangle \langle a, a\rangle = \langle \alpha_a h, h_2\rangle ,$$

esto es

$$\begin{aligned} \langle j^*(h \otimes v), h_2\rangle &= \langle h \otimes v, j(h_2)\rangle \\ &= \langle \langle a, v\rangle h, h_2\rangle \end{aligned}$$

lo que muestra (L3).

Después definimos la función $V : \mathcal{H}_A \longrightarrow \mathcal{H}_B \otimes \mathbb{C}^n$ como la composición $V = U \circ j$ y verificaremos que tiene las propiedades enunciadas:

a) Como U es una isometría parcial, por la forma en que está definida V se sigue inmediatamente que también es isometría parcial.

b) Sea $\{\varphi\}$ una base ortonormal de $\mathcal{H}_A$. Verificaremos que $V\varrho V^*$ y $U\varrho \otimes |a\rangle \langle a| U^*$ son iguales en la base $\{\varphi \otimes b\}$ de $\mathcal{H}_A \otimes \mathbb{C}^n$.
En efecto:

$$\begin{aligned} V\varrho V^*(\varphi \otimes b) &= Uj\varrho j^* U^*(\varphi \otimes b) = Uj\varrho j^*(E_b^*\varphi \otimes a) \\ &= Uj\varrho(E_b^*\varphi) = Uj(\varrho E_b^*\varphi) = U(\varrho E_b^*\varphi \otimes a) \end{aligned}$$

y por otra parte

$$\begin{aligned} U\varrho \otimes |a\rangle \langle a| U^*(\varphi \otimes b) &= U\varrho \otimes |a\rangle \langle a| (E_b^*\varphi \otimes a) = \\ U(\varrho E_b^*\varphi \otimes |a\rangle \langle a| a) &= \langle a, a\rangle U(\varrho E_b^*\varphi \otimes a) = U(\varrho E_b^*\varphi \otimes a), \end{aligned}$$

lo que demuestra **b)**.

c) Hagamos algunos cálculos: sean $x, y \in \mathcal{H}_A$, luego

$$\begin{aligned} \langle x \otimes b, V\varrho V^*(y \otimes b)\rangle &= \langle j^* U^*(x \otimes b), \varrho j^* U^*(y \otimes b)\rangle \\ &= \langle j^*(E_b^* x \otimes a), \varrho j^*(E_b^* y \otimes a)\rangle \\ &= \langle E_b^* x, \varrho E_b^* y\rangle = \langle x, E_b \varrho E_b^* y\rangle , \end{aligned}$$

y después sumemos para obtener que

$$\begin{aligned} \langle x, tr_2[V\varrho V^*] y\rangle &= \sum_b \langle x \otimes b, V\varrho V^*(y \otimes b)\rangle \\ &= \sum_b \langle x, E_b \varrho E_b^* y\rangle \\ &= \left\langle x, \sum_b E_b \varrho E_b^* y \right\rangle \end{aligned}$$

y por la unicidad de la traza parcial concluimos que $tr_2 V \varrho V^* = \sum_b E_b \varrho E_b^* = \Phi(\varrho)$.

d) Usaremos la definición del producto interno de *Hilbert-Schmidt* $\langle \cdot, \cdot \rangle_{HS}$, la propiedad (TP2) de la traza parcial y la ciclicidad de la traza como sigue: sean $x \in \mathcal{B}(\mathcal{H}_B)$, $y \in \mathcal{B}(\mathcal{H}_A)$ luego

$\langle x, \Phi y \rangle_{HS} = tr x^* \Phi(y) = tr \Phi(y) x^* = tr\left[tr_2 (V y V^*) x^*\right] = tr\left[V y V^* (x^* \otimes I_2)\right] = tr\left[V^* (x^* \otimes I_2) V y\right] = tr (V^* x \otimes I_2 V)^* y = \langle V^* x \otimes I_2 V, y \rangle_{HS}$ y por lo tanto

$$\begin{aligned} \left\langle \widehat{\Phi}(x), y \right\rangle_{HS} &= \langle x, \Phi(y) \rangle_{HS} \\ &= \langle V^* x \otimes I_2 V, y \rangle_{HS}, \end{aligned}$$

de aquí concluimos que $\widehat{\Phi}(x) = V^* (x \otimes I_2) V$. **QED**

Para terminar este capítulo enunciaremos un corolario que conjunta los resultados del Teorema 10 y la Proposición 8.

Corolario 11 *Una función $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ es un operador **CP** que preserva la traza si, y sólo si, existen $n \in \mathbb{N}$, $a \in \mathbb{C}^n$ un vector normalizado y un operador $U : \mathcal{H}_A \otimes \mathbb{C}^n \longrightarrow \mathcal{H}_B \otimes \mathbb{C}^n$ tales que:*

$$\Phi(\varrho) = tr_2 U (\varrho \otimes |a\rangle \langle a|) U^*, \quad \varrho \in \mathcal{B}(\mathcal{H}_A).$$

Capítulo 3

Entropía

3.1 Entropía de Shannon ó clásica

La entropía es una medida de la cantidad de información (o incertidumbre) contenida en una fuente de información. También puede ser usada para cuantificar los recursos necesarios para almacenar la información ver [11, 17]. La fuente puede ser modelada como un proceso estocástico y esto motiva la siguiente definición.

Definición (*Entropía de Shannon*). La *entropía* de *Shannon* $H(X)$ de una variable aleatoria X definida en (Ω, P), un espacio de probabilidad finito es

$$H(X) \equiv -\sum_{x \in X(\Omega)} p(x) \log p(x)$$

donde $p(x) = P(X = x) \equiv P\left(X^{-1}\{x\}\right)$ es la distribución de probabilidad de X, por tanto $p(x) \geq 0$ y $\sum_{x \in X(\Omega)} p(x) = 1$.
Completamos la definición conviniendo en que $0 \log 0 = 0$ lo cual se justifica porque la función $f(x) = \begin{cases} x \log x, & x > 0 \\ 0, & x = 0 \end{cases}$ es continua en cero, pues $x \log x \to 0$ cuando $x \to 0^{+}$.

Observemos que **la entropía de Shannon es una cantidad no negativa**, si la base elegida de los logaritmos es estrictamente mayor que 1. Otro factor que determina la elección de la base, es el de la unidad en que se necesite expresar la entropía. Cuando la base es 2, la entropía está expresada en *bits*; y si se emplean logaritmos naturales la entropía se mide en *nats*. Por ejemplo la entropía de una moneda honesta es 1 *bit* ó log(2) *nat*.
Usaremos logaritmos naturales a lo largo de este trabajo.

Ejemplo (*función de la entropía binaria.*) La entropía de una variable aleatoria X de dos resultados es $H_{bin}(q) = -q \log q - (1-q) \log(1-q)$, donde

q y $1-q$ son las probabilidades de los dos resultados. No es difícil ver que $H_{bin}(q)$ es una función cóncava y que alcanza su máximo en $q=\frac{1}{2}$.

3.1.1 Entropía relativa de Shannon

La entropía relativa es una medida muy útil para medir la cercanía de dos distribuciones de probabilidad $p(x)$ y $q(x)$. Es una cantidad muy útil porque frecuentemente otras cantidades entrópicas pueden ser consideradas como casos especiales de ella .

Definición. Supongamos que $p(x), q(x)$ son dos distribuciones de probabilidad. La entropía relativa de $p(x)$, respecto de $q(x)$ es

$$H(p(x) \parallel q(x)) \equiv \sum_x p(x) \log(p(x)/q(x)) \equiv -H(X) - \sum_x p(x) \log q(x)$$

donde definimos $-0\log 0 = 0$ y $-p(x)\log 0 = +\infty$ si $p(x) > 0$.

Teorema 1 *(No negatividad de la entropía relativa). Sean $p(x), q(x)$ dos distribuciones de probabilidad, entonces*

$$H(p(x) \parallel q(x)) \geq 0$$

con igualdad si, y sólo si, $p(x) = q(x)$ para toda x.

Demostración. La desigualdad de *Jensen* **(Apéndice A)**, nos garantiza que si f es una función estrictamente convexa, entonces $f\left(\sum_i a_i x_i\right) \leq \sum_i a_i f(x_i)$, con tal que cada $a_i \geq 0$ y $\sum_i a_i = 1$. Además, si cada a_i es tal que $0 < a_i < 1$, se obtiene la igualdad sólo si los x_i son iguales entre sí.
En lo anterior, poniendo $a_x = q(x)$ y $f(t) = t\log t$, obtenemos

$$\begin{aligned} H(p(x) \parallel q(x)) &= \sum_x p(x) \log(p(x)/q(x)) \\ &= \sum_x q(x)(p(x)/q(x)) \log(p(x)/q(x)) \\ &\geq \left(\sum_x q(x)(p(x)/q(x))\right) \log\left(\sum_x q(x)(p(x)/q(x))\right) \\ &= \left(\sum_x p(x)\right) \log\left(\sum_x p(x)\right) = 1 \cdot 0. \end{aligned}$$

Si por otra parte $H(p(x) \parallel q(x)) = 0$, en el desarrollo anterior tenemos que

$$\begin{aligned} 0 &= \sum_x q(x)(p(x)/q(x)) \log[p(x)/q(x)] \\ &= \left(\sum_x q(x)(p(x)/q(x))\right) \log\left(\sum_x q(x)(p(x)/q(x))\right) , \end{aligned}$$

y dado que en este caso f es una función estrictamente convexa, por lo ya mencionado concluimos que para todo x se cumple que $p(x)/q(x) = 1$. **QED**

Un ejemplo de cómo la entropía relativa se usa para deducir propiedades es el siguiente.

Teorema 2 *Supongamos que X es una variable aleatoria con d valores, entonces $H(X) \leq \log d$ con igualdad si, y sólo si, X está uniformemente distribuida.*

Demostración. Usaremos la no negatividad de la entropía relativa. Sea $p(x)$ una distribución de probabilidad sobre X y sea $q(x) \equiv 1/d$ la distribución uniforme, entonces $0 \leq H(p(x) \parallel 1/d) = -H(X) - \sum_x p(x) \log(1/d) = -H(X) + \log d \sum_x p(x) = -H(X) + \log d$ y hay igualdad si, y sólo si, $p(x) = 1/d$ para cada x. **QED**

3.2 Entropía conjunta

Supongamos que X, Y son variables aleatorias; para contestar la pregunta de cómo se relaciona la información contenida en X con la información contenida en Y se introducen los conceptos de *entropía condicional* e *información compartida.*

Definición. Sean X, Y variables aleatorias discretas. La **entropía conjunta** de X , Y se define por

$$H(X,Y) = -\sum_{x,y} p(x,y) \log p(x,y) \quad ,$$

donde $p(x,y) \equiv P(X = x, Y = y)$.

Intuitivamente, la entropía conjunta mide nuestra incertidumbre total acerca del par (X,Y).

3.2.1 Entropía condicional

Ahora supongamos que conocemos el valor de Y, por lo que hemos adquirido $H(Y)$ digamos, *bits*, de información sobre el par (X,Y). La incertidumbre remanente acerca del par (X,Y) se asocia con nuestra incertidumbre sobre X dado que conocemos Y.

Definición. Sean X, Y variables aleatorias. La entropía de X **condicionada** a que conocemos Y se define como

$$H(X \mid Y) = H(X,Y) - H(Y)$$

Intuitivamente, la entropía condicional mide la incertidumbre promedio que tenemos sobre el valor de X dado que conocemos el valor de Y.

3.2.2 Información común

Definición. Sean X, Y variables aleatorias. La información **común o compartida** entre X y Y se define como

$$H(X:Y) = H(X) + H(Y) - H(X,Y)$$

Note que $H(X:Y) = H(X) - H(X \mid Y)$.

Intuitivamente, la información común contenida en X y Y mide que tanta información tienen en común X y Y y se obtiene después de quitarle a la suma de las informaciones $H(X)$ y $H(Y)$ la información conjunta $H(X,Y)$ porque en $H(X)+H(Y)$ la *información común* se ha contado dos veces, una por $H(X)$ y otra por $H(Y)$, mientras que la información no común se ha contado una vez.

3.3 Propiedades de la entropía de Shannon

La definición que sigue la usaremos en el Teorema 3.
Definición: Una sucesión de variables aleatorias (X, Y, Z) es una cadena de *Markov* si para toda tripleta x, y, z, se verifica que

$$\frac{P(Z=z, Y=y, X=x)}{P(Y=y, X=x)} = \frac{P(Z=z, Y=y)}{P(Y=y)}.$$

Abreviaremos la condición anterior escribiendo simplemente $P(Z \mid Y, X) = P(Z \mid Y)$ y con $X \longrightarrow Y \longrightarrow Z$, simbolizamos que (X, Y, Z) es una cadena de *Markov.*

Observemos que si, $X \longrightarrow Y \longrightarrow Z$ es una cadena de *Markov,* entonces también $Z \longrightarrow Y \longrightarrow X$ es una cadena de *Markov.*

A continuación enlistaremos algunas propiedades de la entropía de *Shannon.* Volveremos a encontrar la propiedad 6) en una forma más general en el capítulo 4 y la propiedad 4) aparecerá nuevamente en el capítulo 5

Teorema 3 *(Propiedades de la entropía de Shannon)*

1) $H(X,Y) = H(Y,X)$, $H(X:Y) = H(Y:X)$

2) $H(Y \mid X) \geq 0$ y como consecuencia

$$H(X:Y) \leq H(Y),$$

con igualdad si, y sólo si, Y es función de X, $Y = f(X)$.

3) $H(X) \leq H(X,Y)$, con igualdad si, y sólo si, Y es función de X, $Y = f(X)$.

4) Subaditividad:

$$H(X,Y) \leq H(X) + H(Y),$$

con igualdad si, y sólo si, X y Y son variables aleatorias independientes.

5) $H(Y \mid X) \leq H(Y)$ y por ende $H(X:Y) \geq 0$, con igualdad si, y sólo si, X y Y son variables aleatorias independientes.

6) Subaditividad fuerte:

$$H(X,Y,Z) + H(Y) \leq H(X,Y) + H(Y,Z),$$

con igualdad si, y sólo si, $Z \longrightarrow Y \longrightarrow X$ es una cadena de *Markov*.

7) El condicionamiento reduce la entropía:

$$H(X \mid Y,Z) \leq H(X \mid Y)$$

Demostración.
1) Se sigue de las definiciones.

2) Puesto que $p(x,y) = p(x)\,p(y \mid x)$ tenemos que

$$\begin{aligned} H(X,Y) &= -\sum_{x,y} p(x,y) \log p(x)\,p(y \mid x) \\ &= -\sum_{x} p(x) \log p(x) - \sum_{x,y} p(x,y) \log p(y \mid x) \\ &= H(X) - \sum_{x,y} p(x,y) \log p(y \mid x) \end{aligned}$$

por lo tanto $H(Y \mid X) \equiv H(X,Y) - H(X) = -\sum_{x,y} p(x,y) \log p(y \mid x)$, es decir

$$H(Y \mid X) = \sum_{x,y} p(x,y) \left[-\log p(y \mid x)\right], \tag{E1}$$

como este último término es no negativo se sigue la primera parte del resultado.

Para analizar la condición de igualdad, si $H(Y \mid X) = 0$ cada sumando $[-\log p(y \mid x)]\,p(x,y)$ debe ser cero lo que obliga a que $\log p(y \mid x) = 0$ ó bien $p(x,y) = 0$.

Cuando $p(x,y) = 0$ dado que x ocurre, y no ocurre. Ahora supongamos que $p(x,y) \neq 0$, entonces $\log p(y \mid x) = 0$ es decir $p(y \mid x) = 1$, y en todo caso se cumple que la ocurrencia de y depende de la de x, o sea $y = f(x)$. Recíprocamente, usando la misma expresión **(E1)** es claro que $p(y \mid x) = 1\ \forall x$ implica que $H(Y \mid X) = 0$.

3) Es consecuencia del inciso anterior.

4) Usaremos que para todo $t > 0$ se cumple la desigualdad $t - 1 \geq \log t$ con igualdad si, y sólo si, $t = 1$.

$$\begin{aligned}\sum_{x,y} p(x,y) \log \frac{p(x)\,p(y)}{p(x,y)} &\leq \sum_{x,y} p(x,y) \left[\frac{p(x)\,p(y)}{p(x,y)} - 1\right] \\ &= \sum_{x,y} p(x)\,p(y) - p(x,y) \\ &= 1 \cdot 1 - 1 = 0\ ,\end{aligned}$$

con igualdad si, y sólo si, $p(x)\,p(y) = p(x,y)\ \ \forall x, y$.

5) Se sigue de la subaditividad y de las definiciones relevantes.

6) Se usa la no negatividad de la entropía relativa de la distribución de probabilidad $p(x,y,z)$ respecto de la distribución de probabilidad $q(x,y,z) = p(x,y)\,[p(y)]^{-1}\,p(y,z)$ y la siguiente equivalencia,

$$\begin{aligned}H\left(p(x,y,z) \parallel q(x,y,z)\right) &\geq 0 \iff \\ H(X,Y,Z) + H(X,Y) - H(Y) + H(Y,Z) &\geq 0\end{aligned}$$

La equivalencia se verifica usando la definición de entropía relativa y las siguientes igualdades:

$$\begin{aligned}\sum_{x,y,z} p(x,y,z) \log p(x,y) &= \sum_{x,y} \log p(x,y) \sum_{z} p(x,y,z) \\ &= \sum_{x,y} \log p(x,y)\, p(x,y)\ ,\end{aligned}$$

$\sum_{x,y,z} p(x,y,z) \log p(y) = \sum_{y} p(y) \log p(y)\ ,$

y $\sum_{x,y,z} p(x,y,z) \log p(y,z) = \sum_{y,z} p(y,z) \log p(y,z)$

La condición de igualdad para la subaditividad fuerte equivale a que la entropía relativa de $p(x,y,z)$ respecto de $q(x,y,z)$ sea igual a cero, o lo que es lo mismo que para toda tripleta x, y, z, se cumpla que,

$$p(x,y,z) = p(x,y)\,[p(y)]^{-1}\,p(y,z) \quad \forall x,y,z \tag{S1}$$

La igualdad anterior se puede escribir como $p(x,y,z)\,[p(x,y)]^{-1} = p(y,z)\,[p(y)]^{-1}$, y denotando con $p(Z \mid Y)$ la distribución de probabilidad condicional clásica, esta última igualdad la podemos reescribir, como

$$p(Z \mid X, Y) = p(Z \mid Y)\ ,$$

es decir $X \longrightarrow Y \longrightarrow Z$ es una cadena de *Markov.*

Observemos que la igualdad **(S1)** también, se puede escribir en la forma equivalente,

$$\log p(x, y, z) - \log p(x, y) = \log p(y, z) - \log p(y) \quad \forall x, y, z$$

que nos interesa resaltar ahora porque aparecerá nuevamente cuando discutamos las condiciones de igualdad de la subaditividad fuerte de la entropía de *von Neumann.*

7) Intuitivamente, esperamos que la incertidumbre acerca de $X,$ dado que conocemos el valor de Y y Z sea menor que nuestra incertidumbre acerca de $X,$ si sólo conocemos el valor de Y. Para la prueba formal simplemente usamos que por definición se cumplen las igualdades: $H(X \mid Y, Z) = H(X, Y, Z) - H(Y, Z)$ y $H(X \mid Y) = H(X, Y) - H(Y)$, reescribiendo la desigualdad de subaditividad fuerte obtenemos el resultado

$$H(X, Y, Z) - H(Y, Z) \leq H(X, Y) - H(Y). \quad \textbf{QED}$$

3.4 Entropía de von Neumann

La entropía de *Shannon* mide la incertidumbre asociada con una distribución de probabilidad y es un caso particular de la entropía de *von Neumann.*

Definición. Un operador ó matriz $\varrho \in \mathcal{B}(\mathcal{H})$ **es de densidad,** si

$$\varrho \geq 0 \quad \text{y} \quad tr\varrho = 1$$

Observemos que si ϱ se representa espectralmente como $\varrho = \sum_i \lambda_i |\varphi_i\rangle\langle\varphi_i|$, la condición de que ϱ tenga traza 1 significa que $\sum_i \lambda_i = 1$.
Queda como comentario que un operador de densidad también es llamado estado cuántico mezclado, aunque nosotros rara vez usaremos esta denominación.

Definición. La entropía de *von Neumann* de un operador de densidad ϱ es

$$S(\varrho) \equiv -tr\varrho \log \varrho$$

recordemos que hemos convenido en usar logaritmos naturales (ver los comentarios hechos después de la definición de la entropía de *Shannon*).

Si ϱ se representa espectralmente como $\varrho = \sum_i \lambda_i |\varphi_i\rangle\langle\varphi_i|$ la definición anterior se puede expresar como

$$S(\varrho) = -\sum_i \lambda_i \log \lambda_i$$

donde por definición aceptamos que $0 \log 0 = 0$

Observaciones:

a) Por el Teorema 1.1 las matrices de densidad son diagonales respecto a alguna base, es decir, se pueden poner en correspondencia biunívoca con distribuciones de probabilidad, y usando la igualdad anterior vemos que la entropía de *von Neumann* contiene como caso especial a la *entropía* de *Shannon*.

b) Como $\mathcal{H}$ es un espacio de *Hilbert* de dimensión finita, la entropía de *von Neumann* es una función continua con la norma de operadores.

En efecto: Sean $\{A,\ A_1,\ A_2, \ldots\} \subseteq \mathcal{B}(\mathcal{H})$, operadores de densidad positivo semi-definidos tales que $A_n \overset{\|\cdot\|}{\longrightarrow} A$. Si para cada n, $\mu_k(A_n), \mu_k(A)$ son los valores propios enumerados en orden decreciente, de A_n y A respectivamente, usando el *teorema del mini-máx* se puede demostrar (ver el corolario III.2.6 de [1]), que para cada k, $|\mu_k(A_n) - \mu_k(A)| \leq \|A_n - A\|$ y por ende para cada k $\mu_k(A_n) \longrightarrow \mu_k(A)$ cuando $n \longrightarrow \infty$. Ahora usamos la continuidad de la función $-x \log x$ en el intervalo $[0,1]$ para concluir que $\sum_k \mu_k(A_n) \log(\mu_k(A_n)) \longrightarrow \sum_k \mu_k(A) \log(\mu_k(A))$ cuando $n \longrightarrow \infty$, es decir $\lim_n S(A_n) = S(A)$.

Lema 4 *Si $U \in \mathcal{B}(\mathcal{H})$ es un operador unitario y $\varrho \geq 0$ es un operador en $\mathcal{H}$, entonces*

$$U(\log \varrho)U^* = \log(U\varrho U^*).$$

Demostración. Si $\varrho = \sum_i \lambda_i |\varphi_i\rangle\langle\varphi_i|$ es la descomposición espectral de ϱ, entonces $U\varrho U^* = \sum_i \lambda_i U |\varphi_i\rangle\langle\varphi_i| U^* = \sum_i \lambda_i |U\varphi_i\rangle\langle U\varphi_i|$ es la descomposición espectral de $U\varrho U^*$ ya que $\{U\varphi_i\}$ es una base ortonormal de $\mathcal{H}$ porque U es unitario.

Por tanto, $\log U\varrho U^* = \sum_i \log \lambda_i |U\varphi_i\rangle\langle U\varphi_i| = U\left(\sum_i \log \lambda_i |\varphi_i\rangle\langle\varphi_i|\right)U^* = U(\log \varrho)U^*$. **QED**

Proposición 5 *(Invarianza de la entropía de von Neumann bajo conjugación unitaria). Si $U \in \mathcal{B}(\mathcal{H})$ es un operador unitario y $\varrho \geq 0$ es un operador de densidad en $\mathcal{H}$, entonces*

$$S(U\varrho U^*) = S(\varrho).$$

Demostración $trU\varrho U^* \log U\varrho U^* = trU(\varrho \log \varrho)U^* = tr\varrho \log \varrho$. La primera igualdad es por el Lema anterior y la última por la invarianza de la traza bajo conjugación unitaria. **QED**

3.4.1 Entropía relativa de von Neumann

Como en el caso de la entropía de *Shannon* es útil definir la entropía relativa.
Definición. Sean $\rho, \gamma \in \mathcal{B}(\mathcal{H})$ matrices de densidad tales que $\ker(\gamma) \subseteq \ker(\rho)$, se define la entropía relativa de ρ respecto a γ como

$$\begin{aligned} S(\rho \parallel \gamma) &\equiv tr\rho \log \rho - tr\rho \log \gamma \\ &\equiv \sum_i \lambda_i \log \lambda_i - \sum_i \sum_j \lambda_i c_{ij} \log \mu_j, \end{aligned}$$

donde $\rho = \sum_i \lambda_i |u_i\rangle \langle u_i|$ y $\gamma = \sum_j \mu_j |v_j\rangle \langle v_j|$ son descomposiciones espectrales, y para cada i, j $c_{ij} \equiv |\langle u_i, v_j \rangle|^2$.

La primera propiedad de la entropía relativa que discutiremos es su invarianza bajo conjugación unitaria, característica que como vimos en la proposición anterior, comparte con la *entropía de von Neumann.*

Proposición 6 *(Invarianza de la entropía relativa de von Neumann bajo conjugación unitaria). Si $U \in \mathcal{B}(\mathcal{H})$ es un operador unitario y $\rho, \sigma \geq 0$ son operadores de densidad en $\mathcal{H}$, entonces*

$$S(U\rho U^* \parallel U\sigma U^*) = S(\rho \parallel \sigma).$$

Demostración. Se sigue de que $trU\rho U^* \log U\sigma U^* = trU\rho \log \sigma U^* = tr\rho \log \sigma$ y de la definición. **QED**

Para demostrar algunas propiedades de la entropía relativa y de la entropía, necesitamos desarrollar algunos resultados, que desembocarán en la prueba de la *Desigualdad de Klein* **(Corolario 9)**, desigualdad que también será de *suma eficacia* en los próximos capítulos.

Lema 7 *Sea A un operador auto-adjunto en $\mathcal{H}$ y $\varphi \in \mathcal{H}$ tal que $\|\varphi\| = 1$. Si f es una función convexa real con dominio en algún subconjunto convexo de $\mathbb{R}$ que contiene al espectro de A, entonces*

$$f(\langle \varphi, A\varphi \rangle) \leq \langle \varphi, f(A)\varphi \rangle$$

Nota: Cuando f es estrictamente convexa, la igualdad se da si, y sólo si, φ es vector propio de A.

Demostración. Sea $A = \sum_i a_i |\varphi_i\rangle \langle \varphi_i|$ la descomposición espectral de A donde $\{\varphi_i\}$ es una base ortonormal de vectores propios y sea $\varphi = \sum_j c_j \varphi_j$ la expresión de φ en términos de los básicos, entonces $\sum_j |c_j|^2 = 1$, y $f(A) = \sum_i f(a_i) |\varphi_i\rangle \langle \varphi_i|$ luego $\langle \varphi, f(A)\varphi \rangle = \sum_k |c_k|^2 f(a_k) \geq f\left(\sum_k |c_k|^2 a_k\right) = f(\langle \varphi, A\varphi \rangle)$. **QED**

Teorema 8 *(Klein). Sea f una función de valores reales, convexa y diferenciable. Si A,B son operadores auto-adjuntos y sus espectros están contenidos en el dominio de f, entonces*

$$tr\,[\ f\,(A) - f\,(B)] \geq tr\,[(A - B)\, f'\,(B)]\,.$$

Demostración. Sea $\{\psi_i\}$ una base ortonormal de vectores propios de B con valores propios asociados $\{b_i\}$, por el lema anterior y por el Lema A_5 del apéndice se cumple que

$$\begin{aligned}
\langle \psi_i, (A - B)\, f'\,(B)\, \psi_i \rangle &= \langle \psi_i, A f'\,(b_i)\, \psi_i \rangle - \langle \psi_i, b_i f'\,(b_i)\, \psi_i \rangle \\
&= f'\,(b_i)\,[\langle \psi_i, A\psi_i \rangle - b_i] \\
&\leq [f\,(\langle \psi_i, A\psi_i \rangle) - f\,(b_i)] \\
&\leq \langle \psi_i, f\,(A)\, \psi_i \rangle - f\,(b_i) \\
&= \langle \psi_i, f\,(A)\, \psi_i \rangle - \langle \psi_i, f\,(B)\, \psi_i \rangle\,,
\end{aligned}$$

sumando obtenemos el resultado.

La siguiente observación se usará en el próximo corolario.
Si $trA \neq trB$, entonces dada una base ortonormal $\{\psi_i\}$ de vectores propios de B con valores propios asociados $\{b_i\}$, existe ψ_j tal que $\langle \psi_j, A\psi_j \rangle \neq \langle \psi_j, B\psi_j \rangle = b_j$. Por lo que **si f es estrictamente convexa y tiene segunda derivada**, su derivada es estrictamente creciente y por el teorema del valor medio, se cumple que $f'\,(b_j)\,[\langle \psi_j, A\psi_j \rangle - b_j] < f\,(\langle \psi_j, A\psi_j \rangle) - f\,(b_j)$ y la cadena anterior de desigualdades implica que
$0 < f\,(\langle \psi_j, A\psi_j \rangle) - f\,(b_j) - f'\,(b_j)\,[\langle \psi_j, A\psi_j \rangle - b_j] \leq$
$tr\,[f\,(A) - f\,(B) - (A - B)\, f'\,(B)]$, es decir

$$0 < tr\,[f\,(A) - f\,(B) - (A - B)\, f'\,(B)]\,. \quad \textbf{QED}$$

Corolario 9 *(Desigualdad de Klein). Para $A, B > 0$ en $\mathcal{B}\,(\mathcal{H})$ se cumple que*

$$tr\,(A \log A - A \log B) \geq tr\,(A - B)\,, \tag{KLE}$$

con igualdad si, y sólo si, $A = B$.

Demostración. Si $f\,(x) = x \log x$, entonces

$$tr\,(A \log A - B \log B) \geq tr\,[(\log B + I)\,(A - B)]\,,$$

desarrollando ambos lados y usando la ciclicidad de la traza vemos que esta desigualdad equivale a $trA \log A - trB \log B \geq trA \log B - trB \log B + tr\,(A - B)$ de donde se sigue **(KLE)**.

Para discutir la condición de igualdad, por la observación hecha al final de la demostración del Teorema 8 es suficiente considerar el caso cuando $trA = trB$.

Sean $A = \sum_i a_i \left|u_i\right\rangle \left\langle u_i\right|$ y $B = \sum_j b_j \left|v_j\right\rangle \left\langle v_j\right|$ las descomposiciones espectrales de A y B donde $\{u_i\}, \{v_j\}$ son bases ortonormales para $\mathcal{H}$.
Para cada i, j sea $c_{ij} = \left|\left\langle u_i, v_j\right\rangle\right|^2$, luego $\left\langle u_i, A \log B u_i\right\rangle = \left\langle A u_i, \log B u_i\right\rangle = \sum_j c_{ij} a_i \log b_j$ y por tanto

$$\begin{aligned}
0 &= tr\,(A - B) = tr\,(A \log A) - tr\,(A \log B) \\
&= \sum_i a_i \left(\log a_i - \sum_j c_{ij} \log b_j \right) \\
&\geq \sum_i a_i \left[\log a_i - \log \left(\sum_j c_{ij} b_j \right) \right] \\
&= H\,(a_i \parallel r_i)\ ,
\end{aligned}$$

la desigualdad es debido a que $\log(\cdot)$ es una función cóncava (estrictamente) y hemos definido $r_i \equiv \sum_j c_{ij} b_j$. Como sabemos, si la entropía de *Shannon* es cero, entonces $a_i = r_i$ para toda i.

Sustituyendo esto en los términos de arriba tenemos

$$\begin{aligned}
0 &\geq \sum_i a_i \left[\log a_i - \sum_j c_{ij} \log b_j \right] \\
&= \sum_i r_i \left[\log r_i - \sum_j c_{ij} \log b_j \right] \geq 0,
\end{aligned}$$

por la concavidad del logaritmo cada uno de los últimos sumandos es no negativo y entonces se cumple que $\forall i$, $\log r_i = \sum_j c_{ij} \log b_j$. De nuevo por la concavidad estricta del logaritmo existe un, necesariamente, único j tal que $c_{ij} = 1$ y $c_{ij\prime} = 0$ para $j\prime \neq j$ por lo tanto, existe un único j tal que $a_i = r_i \equiv \sum_j c_{ij} b_j = b_j$.

En resumen, A y B tienen el mismo espectro y $\forall i\ \exists j$ tal que $\left\langle u_i, v_{j\prime}\right\rangle = 0$ cuando $j\prime \neq j$. Por tanto $u_i = \sum_{j\prime} \left\langle v_{j\prime}, u_i\right\rangle v_{j\prime} = \left\langle v_j, u_i\right\rangle v_j$ y análogamente $v_j = \sum_{i\prime} \left\langle u_{i\prime}, v_j\right\rangle u_{i\prime} = \left\langle u_i, v_j\right\rangle u_i$, usando estas relaciones y el hecho de que

$$A u_i = a_i u_i \ \text{ y } \ B v_j = b_j v_j = a_i v_j,$$

después de algunos cálculos obtenemos que $BAu_i = a_i B u_i = a_i^2 u_i$ y $ABu_i = A \left\langle v_j, u_i\right\rangle B v_j = a_i^2 u_i$, es decir A y B conmutan, por el Teorema 1.2 existe una base común de vectores propios con respecto a la que son diagonales y como tienen el mismo espectro concluimos que A y B son iguales. **QED**

Teorema 10 *(No negatividad de la entropía relativa) Si $\rho, \gamma \in \mathcal{B}(\mathcal{H})$ son operadores de densidad tales que $\rho, \gamma > 0$, entonces*

$$S\,(\rho \parallel \gamma) \geq 0 \ \textit{con igualdad si, y sólo si, } \rho = \gamma$$

Demostración. Usamos directamente la *desigualdad de Klein*. Por esto suele llamársele desigualdad de *Klein* a la no negatividad de la entropía relativa junto con su condición de igualdad. **QED**

3.5 Propiedades de la entropía de von Neumann

El siguiente resultado reúne algunas de las propiedades básicas de la entropía de *von Neumann*.

Teorema 11 *(Propiedades básicas de la entropía de von Neumann.)*

a) La entropía es no negativa. La entropía de un estado es cero si, y sólo si, el estado es un *estado puro* (i.e. es de la forma $\varrho = |\varphi\rangle\langle\varphi|$ para algún $\varphi \in \mathcal{H}$ de norma 1).

b) En un espacio de *Hilbert* de dimensión d la entropía es a lo más $\log d$. La entropía es igual a $\log d$ si, y sólo si, el sistema "está en el estado uniformemente mezclado $\frac{I}{d}$".

c) Si $\varrho_{12} \in \mathcal{B}(\mathcal{H}_{12})$ es un *estado puro* (i.e. es de la forma $\varrho_{12} = |\varphi\rangle\langle\varphi|$ para algún $\varphi \in \mathcal{H}_{12}$ de norma 1), entonces

$$S(\varrho_1) = S(\varrho_2)$$

Se suele enunciar este resultado diciendo que si un sistema compuesto está en estado cuántico puro, entonces las entropías de los subsistemas que lo componen son iguales.

d) (*Entropía de un producto tensorial*). Para $\rho \in \mathcal{B}(\mathcal{H}_1)$, $\sigma \in \mathcal{B}(\mathcal{H}_2)$ operadores de densidad estrictamente positivos se cumple que

$$S(\rho \otimes \sigma) = S(\rho) + S(\sigma).$$

Desde el punto de vista de la teoría de la información, esta propiedad se interpreta como sigue: si tenemos dos sistemas independientes descritos por $\rho, \mathcal{H}_1$ y $\sigma, \mathcal{H}_2$ respectivamente, entonces la información total del sistema descrita por $\rho\otimes\sigma, \mathcal{H}_1\otimes\mathcal{H}_2$ es igual a la suma de información contenida en sus subsistemas.

e) (*La entropía relativa es aditiva*). Si $\rho_i, \gamma_i > 0$ con $\rho_i, \gamma_i \in \mathcal{B}(\mathcal{H}_i)$ son operadores de densidad, entonces

$$S(\rho_1 \otimes \rho_2 \parallel \gamma_1 \otimes \gamma_2) = S(\rho_1 \parallel \gamma_1) + S(\rho_2 \parallel \gamma_2)$$

f) Si $\{p_i\}$ es una distribución de probabilidad y los estados ρ_i tienen soporte en subespacios ortogonales, entonces

$$S\left(\sum_i p_i\rho_i\right) = H\left(p_i\right) + \sum_i p_i S\left(\rho_i\right),$$

donde el soporte de un operador T se define como $SuppT = [\ker T]^{\perp}$.

g) *Teorema de la entropía conjunta.* Supongamos que $\{p_i\}$ es una distribución de probabilidad, $\{|i\rangle\}$ es un subconjunto ortogonal de $\mathcal{H}_A$ y $\{\rho_i\} \subset \mathcal{B}\left(\mathcal{H}_B\right)$ es un subconjunto de operadores de densidad, entonces

$$S\left(\sum_i p_i\left|i\right\rangle\left\langle i\right| \otimes \rho_i\right) = H\left(p_i\right) + \sum_i p_i S\left(\rho_i\right)$$

Las hipótesis de este resultado se suelen enunciar de la siguiente manera: p_i son probabilidades, $\{|i\rangle\}$ son estados ortogonales del sistema A y ρ_i es un conjunto de operadores de densidad de otro sistema B.

Demostración.

a) La no negatividad de la entropía es inmediata de la definición. Si $\varrho = \left|\varphi\right\rangle\left\langle\varphi\right|$ para algún φ de norma 1 es claro que su entropía es cero.

Recíprocamente, supongamos que el operador de densidad ϱ tiene entropía cero y sea $\varrho = \sum_i \lambda_i \left|u_i\right\rangle\left\langle u_i\right|$ su descomposición espectral. La condición

$$\sum_i \left(-\lambda_i \log\lambda_i\right) = 0,$$

implica que para cada i, $-\lambda_i\log\lambda_i = 0$, por lo tanto $\lambda_{i_0} = 1$ para un único i_0, pues $\sum_i \lambda_i = tr\varrho = 1$, porque ϱ es un operador de densidad. Concluimos que $\varrho = \left|u_{i_0}\right\rangle\left\langle u_{i_0}\right|$, para un elemento u_{i_0} de norma 1.

b) Usamos la no negatividad de la entropía relativa y que $tr\varrho = 1$.
$0 \leq S\left(\varrho \parallel \frac{I}{d}\right) = -S\left(\varrho\right) - tr\varrho\log\frac{I}{d} = -S\left(\varrho\right) - tr\varrho\left(\log\frac{1}{d}I\right) = -S\left(\varrho\right) - \log\frac{1}{d}tr\varrho = -S\left(\varrho\right) - \log\frac{1}{d} = -S\left(\varrho\right) + \log d$, y como sabemos por la condición de igualdad de la entropía relativa, hay igualdad a cero si, y sólo si, $\varrho = \frac{I}{d}$.

c) De la descomposición de *Schmidt* (Teorema 2.7), sabemos que las trazas parciales tienen los mismos valores propios y como la entropía está completamente determinada por los valores propios se tiene el resultado.

d) Sabemos (lema 1.14), que $\log\left(\rho\otimes\sigma\right) = \left(\log\rho\right)\otimes I_2 + I_1\otimes\left(\log\sigma\right)$ y como $tr\rho = tr\sigma = 1$ tenemos que
$tr\left[\rho\otimes\sigma\log\left(\rho\otimes\sigma\right)\right] = tr\rho\left(\log\rho\right) + tr\sigma\left(\log\sigma\right)$, de donde se deduce el resultado.

e) Usamos de nuevo que $\log\left(\gamma_1\otimes\gamma_2\right) = \left(\log\gamma_1\right)\otimes I_2 + I_1\otimes\left(\log\gamma_2\right)$ y como en el inciso anterior se tiene que,

$tr\left(\rho_1 \otimes \rho_2\right) \log\left(\gamma_1 \otimes \gamma_2\right) = tr\left(\rho_1 \log \gamma_1\right) tr\rho_2 + tr\rho_1 tr\left(\rho_2 \log \gamma_2\right)$, es decir

$$tr\left(\rho_1 \otimes \rho_2\right) \log\left(\gamma_1 \otimes \gamma_2\right) = tr\left(\rho_1 \log \gamma_1\right) + tr\left(\rho_2 \log \gamma_2\right)$$

y de aquí, usando la definición de entropía relativa tenemos

$$S\left(\rho_1 \otimes \rho_2 \parallel \gamma_1 \otimes \gamma_2\right) = tr\left(\rho_1 \otimes \rho_2\right) \log\left(\rho_1 \otimes \rho_2\right) - tr\left(\rho_1 \otimes \rho_2\right) \log\left(\gamma_1 \otimes \gamma_2\right)$$

$$= -S\left(\rho_1 \otimes \rho_2\right) - tr\left(\rho_1 \log \gamma_1\right) - tr\left(\rho_2 \log \gamma_2\right)$$

$$= -S\left(\rho_1\right) - S\left(\rho_2\right) - tr\left(\rho_1 \log \gamma_1\right) - tr\left(\rho_2 \log \gamma_2\right)$$

la última igualdad es por el inciso anterior. Usando las definiciones y reagrupando se tiene el resultado.

f) Supongamos que p_i es una distribución de probabilidad y que los estados ρ_i tienen soportes ortogonales, es decir $\left[\ker \rho_i\right]^{\perp} \perp \left[\ker \rho_j\right]^{\perp}$ si $i \neq j$. En el caso de dimensión finita se cumple que $\left[\ker T\right]^{\perp} = ranT$ y así, lo anterior equivale a que $ran\rho_i \perp ran\rho_j$ para $i \neq j$ emplearemos esta forma en la demostración.
Sean λ_i^j y e_i^j los correspondientes elementos propios asociados de ρ_i. Observemos que $p_i\lambda_i^j$ y e_i^j son valores propios y vectores propios asociados de $\sum_i p_i\rho_i$. Para verlo usaremos que $ran\ \rho_i \perp ran\ \rho_j$ cuando $i \neq j$ como sigue:

$\left\langle \rho_i e_k^j, \rho_i e_k^j \right\rangle = \frac{1}{\lambda_k^j} \left\langle \rho_i^2 e_k^j, \lambda_k^j e_k^j \right\rangle = \frac{1}{\lambda_k^j} \left\langle \rho_i^2 e_k^j, \rho_k e_k^j \right\rangle = \frac{1}{\lambda_k^j} \left\langle \rho_i \left(\rho_i e_k^j\right), \rho_k \left(e_k^j\right) \right\rangle =$ 0, si $k \neq i$. Por lo tanto $\left\|\rho_i e_k^j\right\|^2 = 0$ si $k \neq i$ y en consecuencia,

$$\left(\sum_i p_i\rho_i\right) e_k^j = \sum_i p_i \left(\rho_i e_k^j\right) = p_k \rho_k e_k^j = p_k \lambda_k^j e_k^j$$

y así, $S\left(\sum_i p_i\rho_i\right) = -\sum_{i,j} p_i\lambda_i^j \log p_i\lambda_i^j = -\sum_{i,j} p_i\lambda_i^j \log p_i - \sum_{i,j} p_i\lambda_i^j \log \lambda_i^j =$

$-\sum_i p_i \log p_i \sum_j \lambda_i^j - \sum_i p_i \sum_j \lambda_i^j \log \lambda_i^j = H\left(p_i\right) + \sum_i p_i S\left(\rho_i\right).$

g) Usando el inciso precedente, tenemos que

$S\left(\sum_i p_i \left|i\right\rangle \left\langle i\right| \otimes \rho_i\right) = H\left(p_i\right) + \sum_i p_i S\left(\left|i\right\rangle \left\langle i\right| \otimes \rho_i\right)$

$= H\left(p_i\right) + \sum_i p_i \left[S\left(\left|i\right\rangle \left\langle i\right|\right) + S\left(\rho_i\right)\right] = H\left(p_i\right) + \sum_i p_i S\left(\rho_i\right),$

la última igualdad se obtiene usando los incisos d) y a). **QED**

Capítulo 4

Desigualdades de la entropía

4.1 Herramienta matemática

En esta sección desarrollaremos el bagaje matemático necesario para el capítulo.
Definición Supongamos que $f(n)$, $g(n)$ son dos funciones con valores reales y definidas en los enteros no negativos.
Diremos que *f(n)* está en la clase de funciones $O(g(n))$ ó sólo que $f(n)$ es $O(g(n))$, si existen constantes $C > 0$ y N_0 tales que

$$\text{si } n \geq N_0, \text{entonces } |f(n)| \leq Cg(n).$$

Esto es, para n suficiente grande la función *g(n)* es una cota superior de *f(n)* salvo un factor constante.

Lema 1 *(Fórmula del producto de Lie). Para cualesquiera matrices A, B se cumple*

$$\lim_{m \to \infty} \left(e^{\frac{A}{m}} e^{\frac{B}{m}} \right)^m = e^{A+B}$$

Demostración. Para dos matrices X, Y y $m \in \mathbb{N}$ se tiene $X^m - Y^m = \sum_{i=0}^{m-1} X^{m-1-i}(X - Y)Y^i$, entonces

$$\|X^m - Y^m\| \leq mM^{m-1} \|X - Y\|,$$

donde $M = \max\{\|X\|, \|Y\|\}$.

Poniendo $X = e^{\frac{A+B}{m}}$ y $Y = e^{\frac{A}{m}} e^{\frac{B}{m}}$ vemos que están acotadas por $e^{\left(\frac{\|X\|+\|Y\|}{m}\right)}$. Usando la expansión en serie de potencias de la función exponencial se tiene

$X - Y =$

$\left[1+\frac{A+B}{m}+\frac{1}{2}\left(\frac{A+B}{m}\right)^2+\cdots\right]-\left[1+\frac{A}{m}+\frac{1}{2}\left(\frac{A}{m}\right)^2+\cdots\right]\left[1+\frac{B}{m}+\frac{1}{2}\left(\frac{B}{m}\right)^2+\cdots\right]$ $=O\left(\frac{1}{m^2}\right)$, para m grande. Como para todo m, $X^m=e^{A+B}$ por la primera desigualdad obtenemos

$$\left\|e^{A+B}-\left(e^{\frac{A}{m}}e^{\frac{B}{m}}\right)^m\right\|\leq m\ e^{(\|X\|+\|Y\|)}O\left(\frac{1}{m^2}\right)$$

que tiende a cero cuando $m\to\infty$. **QED**

Teorema 2 *(Golden-Thompson). Para operadores auto-adjuntos A y B,*

$$tr\ e^{A+B}\leq tr\ e^Ae^B. \tag{G-T}$$

Nota: hay igualdad en la desigualdad anterior si, y sólo si, A y B conmutan, ver [16].

Demostración. Sea T un operador cualquiera. Observemos que si $\{t_i\}$ son los valores propios de T usando la descomposición canónica de *Jordan* se tiene que $\{t_i^m\}_i$ son los valores propios de T^m y que $\left\{|t_i|^2\right\}$ son los valores propios de T^*T, y entonces $\left|trT^{2m}\right|=\left|\sum_i t_i^{2m}\right|\leq\sum_i\left|t_i^{2m}\right|=\sum_i\left(|t_i|^2\right)^m=tr\left[(T^*T)^m\right]$, es decir

$$\left|trT^{2m}\right|\leq tr\left[(T^*T)^m\right].$$

Usando esta desigualdad y la ciclicidad de la traza tenemos,

$$\begin{aligned}\left|tr\left(e^{\frac{A}{2^{m+1}}}e^{\frac{B}{2^{m+1}}}\right)^{2^{m+1}}\right| &\leq tr\left[\left(e^{\frac{A}{2^{m+1}}}e^{\frac{B}{2^{m+1}}}\right)^*\left(e^{\frac{A}{2^{m+1}}}e^{\frac{B}{2^{m+1}}}\right)\right]^{2^m}\\ &= tr\left[\left(e^{\frac{A}{2^{m+1}}}\right)^2\left(e^{\frac{B}{2^{m+1}}}\right)^2\right]^{2^m},\end{aligned}$$

es decir

$$\left|tr\left(e^{\frac{A}{2^{m+1}}}e^{\frac{B}{2^{m+1}}}\right)^{2^{m+1}}\right|\leq tr\left[e^{\frac{A}{2^m}}e^{\frac{B}{2^m}}\right]^{2^m}.$$

Repitiendo el procedimiento con este último término obtenemos

$$\begin{aligned}tr\left[e^{\frac{A}{2^m}}e^{\frac{B}{2^m}}\right]^{2^m} &\leq tr\left[\left(e^{\frac{A}{2^m}}e^{\frac{B}{2^m}}\right)^*\left(e^{\frac{A}{2^m}}e^{\frac{B}{2^m}}\right)\right]^{2^{m-1}}\\ &\leq tr\left[e^{\frac{A}{2^{m-1}}}e^{\frac{B}{2^{m-1}}}\right]^{2^{m-1}},\end{aligned}$$

lo que muestra que $tr\left[e^{\frac{A}{2^m}}e^{\frac{B}{2^m}}\right]^{2^m}$ es una sucesión decreciente acotada superiormente por tre^Ae^B, término que corresponde a $m=0$. Haciendo tender $m\to\infty$, por la continuidad de la traza y la fórmula del producto de *Lie* obtenemos el resultado. **QED**

Lema 3 *Sean* $A, B \in \mathcal{B}(\mathcal{H})$ *tales que* $A, B > 0$, *entonces*

$$\log B - \log A = \int_0^\infty \left[(xI + A)^{-1} - (xI + B)^{-1} \right] dx$$

Demostración. Por el teorema fundamental del cálculo para funciones con valores en el álgebra $\mathcal{B}(\mathcal{H})$, (ver Apéndice D) tenemos que

$$\begin{aligned} \int_0^\eta (xI + B)^{-1} dx &= \log (\eta I + B) - \log B = \log \eta \left(I + \eta^{-1} B \right) - \log B \\ &= \log \eta I + \log \left(I + \eta^{-1} B \right) - \log B \end{aligned}$$

Similarmente
$\int_0^\eta (xI + A)^{-1} dx = \log \eta I + \log \left(I + \eta^{-1} A \right) - \log A$, y por tanto

$\int_0^\eta (xI + A)^{-1} dx - \int_0^\eta (xI + B)^{-1} dx = - \log \left(I + \eta^{-1} B \right) + \log B + \log \left(I + \eta^{-1} A \right) - \log A$ y dado que $\log (\cdot)$ es función continua se sigue el resultado. **QED**

Corolario 4 *Para todo* $x \in [0, \infty)$ *y para cualesquiera operadores* $S, T > 0$,

$$\log (S + xT) - \log S = \int_0^\infty \frac{1}{(S + uI)} xT \frac{1}{(S + xT + uI)} du$$

Demostración. Sean $B = S + xT$ y $A = S$. Ahora usamos que $A^{-1} - B^{-1} = A^{-1} [B - A] B^{-1} = A^{-1} xTB^{-1}$. **QED**

Lema 5 *Si* $A \in \mathcal{B}(\mathcal{H})$ *es un operador positivo definido y* $B \in \mathcal{B}(\mathcal{H})$ *es un operador arbitrario, entonces*

$$I = \int_0^\infty (xI + A)^{-1} A (xI + A)^{-1} dx.$$

Como consecuencia

$$\textbf{\textit{a)}} \quad A = \int_0^\infty A (xI + A)^{-1} A (xI + A)^{-1} dx$$

y

$$\textbf{\textit{b)}} \quad B = \int_0^\infty B (xI + A)^{-1} A (xI + A)^{-1} dx.$$

Demostración. La primera integral se sigue de la continuidad de la inversión matricial y de las siguientes igualdades,

$$\begin{aligned}\int_0^\eta (xI+A)^{-1} A (xI+A)^{-1} dx &= \int_0^\eta A (xI+A)^{-2} dx \\ &= A\left[-(xI+A)^{-1}\right]_0^\eta \\ &= A\left[-(\eta I+A)^{-1}+A^{-1}\right] \\ &= A\left[-\eta^{-1}\left(I+\eta^{-1}A\right)^{-1}+A^{-1}\right]. \textbf{ QED}\end{aligned}$$

El siguiente teorema fue establecido por *E. Lieb* en [8] y después *Epstein* [4] encontró una prueba más breve que fue simplificada por *M.B. Ruskai* [16] que es la que recomendamos.

Teorema 6 *Sea K una matriz fija. Si K es auto-adjunta, entonces la función $A \longrightarrow F(A) = tre^{K+\log A}$ es cóncava en $A>0$.*

Demostración. Ver el artículo de *Mary Beth Ruskai* [16], para una prueba simplificada, breve y elegante encontrada originalmente por *Epstein* en [4].

Teorema 7 ***(Lieb)*** *Para $R, S, T > 0$,*

$$tr\ e^{\log R - \log S + \log T} \leq tr \int_0^\infty R \frac{1}{(S+uI)} T \frac{1}{(S+uI)} du.$$

Demostración. Por el teorema anterior, $A \overset{F}{\mapsto} tr\ e^{K+\log A}$ es cóncava para $A>0$ y K fija. Sea x un número real tal que $x>0$. Usando que $\log(xA) = \log xI + \log A$ (Lema 1.13), tenemos $e^{K+\log(xA)} = e^{K+\log xI + \log A} = xe^{K+\log A}$, es decir, la función F también es homogénea.

Sea $K = \log R - \log S$, por el Lema A_6 (Apéndice A) con $a = S$ y $b = T$, se tiene que

$$\begin{aligned} xtr\ e^{(\log R - \log S)+\log T} &\leq tre^{(\log R-\log S)+\log(S+xT)} - tre^{\log R} \\ &\leq tre^{\log R} e^{\log(S+xT)-\log S} - trR, \end{aligned}$$

la última desigualdad es por **(G-T)**. En resumen

$$tre^{\log R - \log S + \log T} \leq \frac{trRe^{\log(S+xT)-\log S} - trR}{x}.$$

Por otra parte, aplicando el Corolario 4 se tiene que

$e^{\log(S+xT)-\log S} = e^{\int_0^\infty \frac{1}{(S+uI)} xT \frac{1}{(S+xT+uI)} du} = I + x \int_0^\infty \frac{1}{(S+uI)} T \frac{1}{(S+xT+uI)} du + o(x)$,

y tomando la traza
$trRe^{\log(S+xT)-\log S} = trR + xtrR \int_0^\infty \frac{1}{(S+uI)} T \frac{1}{(S+xT+uI)} du + o(x)$, por ende

$$tr\ e^{\log R - \log S + \log T} \leq trR \int_0^\infty \frac{1}{(S+uI)} T \frac{1}{(S+xT+uI)} du + \frac{o(x)}{x}$$

de donde se sigue el resultado al hacer tender $x \to 0^+$. **QED**

4.2 Subaditividad fuerte

En adelante usaremos la convención hecha en la introducción de que, salvo que se establezca otra cosa, trabajaremos **operadores de densidad estrictamente positivos** (i.e. operadores de traza 1 y positivos estrictamente). Ocasionalmente para enfatizar alguna propiedad escribiremos explícitamente esta condición.

Observación: Usando la igualdad (Proposición 2.4),

$$tr\varrho_{12} \frac{1}{(\varrho_2 + uI)} \varrho_{23} \frac{1}{(\varrho_2 + uI)} = tr\varrho_2 \frac{1}{(\varrho_2 + uI)} \varrho_2 \frac{1}{(\varrho_2 + uI)},$$

y la continuidad de la traza deducimos que

$$\begin{aligned}
tr \int_0^\infty \varrho_{12} \frac{1}{(\varrho_2 + uI)} \varrho_{23} \frac{1}{(\varrho_2 + uI)} du &= \int_0^\infty tr \left[\varrho_{12} \frac{1}{(\varrho_2 + uI)} \varrho_{23} \frac{1}{(\varrho_2 + uI)} \right] du \\
&= \int_0^\infty tr \left[\varrho_2 \frac{1}{(\varrho_2 + uI)} \varrho_2 \frac{1}{(\varrho_2 + uI)} \right] du \\
&= tr \int_0^\infty \varrho_2 \frac{1}{(\varrho_2 + uI)} \varrho_2 \frac{1}{(\varrho_2 + uI)} du,
\end{aligned}$$

es decir

$$\begin{aligned}
tr \int_0^\infty \varrho_{12} \frac{1}{(\varrho_2 + uI)} \varrho_{23} \frac{1}{(\varrho_2 + uI)} du &= tr \int_0^\infty \varrho_2 \frac{1}{(\varrho_2 + uI)} \varrho_2 \frac{1}{(\varrho_2 + uI)} du \\
&= tr\varrho_2,
\end{aligned}$$

la última igualdad se justifica por a) del Lema 5**.** Usaremos este resultado en la demostración del teorema siguiente.

Teorema 8 *(Subaditividad Fuerte). Si $\varrho_{123} \in \mathcal{B}(\mathcal{H}_{123})$ es un operador de densidad estrictamente positivo, entonces*

$$S(\varrho_{123}) + S(\varrho_2) \leq S(\varrho_{12}) + S(\varrho_{23}) \tag{SSA$_1$}$$

y la igualdad se alcanza si, y sólo si, $\log \varrho_{123} - \log \varrho_{12} = \log \varrho_{23} - \log \varrho_2$.

Recordemos que, *e.g.* $\log \varrho_{12}$ significa $(\log \varrho_{12}) \otimes I_3$.

Demostración. Por la propiedad TP2 de la traza parcial (Proposición 2.1) se cumplen las siguientes igualdades:
$tr\varrho_{12} \log \varrho_{12} = tr\varrho_{123} (\log \varrho_{12} \otimes I_3)$, $tr\varrho_2 \log \varrho_2 = tr\varrho_{123} (I_1 \otimes \log \varrho_2 \otimes I_3)$ y $tr\varrho_{23} \log \varrho_{23} = tr\varrho_{123} (I_1 \otimes \log \varrho_{23})$.

Sean

$$A = \varrho_{123} \text{ y } B = e^{\log \varrho_{12} - \log \varrho_2 + \log \varrho_{23}},$$

es decir B se ha elegido de tal forma que satisface $\log B = \log \varrho_{12} - \log \varrho_2 + \log \varrho_{23}$, entonces

$$-S(\varrho_{123}) + S(\varrho_{12}) - S(\varrho_2) + S(\varrho_{23}) = tr\varrho_{123} (\log \varrho_{123} - \log B).$$

La *Desigualdad de Klein* (Corolario 3.9) y el teorema de *Lieb* (Teorema 7), respectivamente implican

$$\begin{aligned}
tr\varrho_{123} (\log \varrho_{123} - \log B) &\geq tr\left(\varrho_{123} - e^{\log \varrho_{12} - \log \varrho_2 + \log \varrho_{23}}\right) \\
&\geq tr\left(\varrho_{123} - \int_0^\infty \varrho_{12} \frac{1}{(\varrho_2 + uI)} \varrho_{23} \frac{1}{(\varrho_2 + uI)} du\right) \\
&= tr\varrho_{123} - tr \int_0^\infty \varrho_2 \frac{1}{(\varrho_2 + uI)} \varrho_2 \frac{1}{(\varrho_2 + uI)} du \\
&= (tr\varrho_{123} - tr\varrho_2) = 1 - 1 = 0,
\end{aligned}$$

donde la primera de las igualdades es por la **observación** previa a este teorema, y la segunda es porque el operador traza parcial preserva la traza (Proposición 2.2, b)) por ende se cumple (SSA$_1$)

Para la condición de igualdad, notemos que hemos mostrado que

$$\begin{aligned}
-S(\varrho_{123}) + S(\varrho_{12}) - S(\varrho_2) + S(\varrho_{23}) &= tr\varrho_{123} (\log \varrho_{123} - \log B) \\
&\geq tr\left(\varrho_{123} - e^{\log \varrho_{12} - \log \varrho_2 + \log \varrho_{23}}\right) \\
&\geq 0,
\end{aligned}$$

de tal manera que si en **(SSA$_1$)** hay igualdad, entonces
$0 = tr\varrho_{123} (\log \varrho_{123} - \log B) \geq tr\left(\varrho_{123} - e^{\log \varrho_{12} - \log \varrho_2 + \log \varrho_{23}}\right) = 0$, y por la condición para igualdad correspondiente en la *Desigualdad de Klein* (Corolario 3.9), tenemos que

$$\varrho_{123} = e^{\log \varrho_{12} - \log \varrho_2 + \log \varrho_{23}}.$$

Recíprocamente, si $\log \varrho_{123} - \log \varrho_{12} = \log \varrho_{23} - \log \varrho_2$, entonces

$$0 = tr\varrho_{123} (\log \varrho_{123} - \log B) = -S(\varrho_{123}) + S(\varrho_{12}) - S(\varrho_2) + S(\varrho_{23}). \quad \textbf{QED}$$

4.3 Desigualdades de la entropía relativa

Teorema 9 *(Monotonía bajo la traza parcial). Si* $\rho_{12}, \gamma_{12} \in \mathcal{B}(\mathcal{H}_{12})$ *son operadores de densidad estrictamente positivos, entonces*

$$S(\rho_2 \parallel \gamma_2) \le S(\rho_{12} \parallel \gamma_{12}) \qquad \text{(MPT)}$$

Con igualdad si, y sólo si, $\log \rho_2 - \log \gamma_2 = \log \rho_{12} - \log \gamma_{12}$.

Recordemos que esta igualdad debe interpretarse como $I_1 \otimes (\log \rho_2 - \log \gamma_2) = \log \rho_{12} - \log \gamma_{12}$.

Demostración. Por la propiedad (TP2) de la traza parcial (Prop. 2.1), $tr\rho_2 \log \rho_2 = tr\rho_{12}(I_1 \otimes \log \rho_2)$ y $tr\rho_2 \log \gamma_2 = tr\rho_{12}(I_1 \otimes \log \gamma_2)$, entonces

$$S(\rho_2 \parallel \gamma_2) = tr[\rho_{12}(I_1 \otimes \log \rho_2) - \rho_{12}(I_1 \otimes \log \gamma_2)]$$

y por ende
$S(\rho_{12} \parallel \gamma_{12}) - S(\rho_2 \parallel \gamma_2) =$
$tr[\rho_{12} \log \rho_{12} - \rho_{12}(\log \gamma_{12} + I_1 \otimes \log \rho_2 - I_1 \otimes \log \gamma_2)]$.

Sean

$$A = \rho_{12} \ \text{ y } \ B = e^{\log \gamma_{12} + I_1 \otimes \log \rho_2 - I_1 \otimes \log \gamma_2}$$

$(\Leftrightarrow \log B = \log \gamma_{12} + I_1 \otimes \log \rho_2 - I_1 \otimes \log \gamma_2)$, entonces la *Desigualdad de Klein* (Cor. 3.9) y el teorema de *Lieb* (Teo.7), implican

$$\begin{aligned}
S(\rho_{12} \parallel \gamma_{12}) - S(\rho_2 \parallel \gamma_2) &\ge tr\left(\rho_{12} - e^{\log \gamma_{12} + \log \rho_2 - \log \gamma_2}\right) \\
&\ge tr\left(\rho_{12} - \int_0^\infty \gamma_{12} \frac{1}{(\gamma_2 + uI)} \rho_2 \frac{1}{(\gamma_2 + uI)} du\right) \\
&= tr\rho_{12} - tr \int_0^\infty \gamma_2 \frac{1}{(\gamma_2 + uI)} \rho_2 \frac{1}{(\gamma_2 + uI)} du \\
&= (tr\rho_{12} - tr\rho_2) = 1 - 1 = 0,
\end{aligned}$$

la primera igualdad se demuestra de manera semejante a la **observación** hecha previamente al enunciado del Teorema 8 y la segunda es porque el operador traza parcial preserva la traza y porque al usar la ciclicidad de la traza tenemos

$tr \int_0^\infty \gamma_2 \frac{1}{\gamma_2 + uI} \rho_2 \frac{1}{\gamma_2 + uI} du = \int_0^\infty tr\gamma_2 \frac{1}{\gamma_2 + uI} \rho_2 \frac{1}{\gamma_2 + uI} du = \int_0^\infty tr\rho_2 \frac{1}{\gamma_2 + uI} \gamma_2 \frac{1}{\gamma_2 + uI} du$

$= tr \int_0^\infty \rho_2 \frac{1}{\gamma_2 + uI} \gamma_2 \frac{1}{\gamma_2 + uI} du = tr\rho_2$, donde la última integral resulta de aplicar el inciso b) del Lema 5.

Para la condición de igualdad, observe que hemos mostrado la validez de

$$\begin{aligned}
S(\rho_{12} \parallel \gamma_{12}) - S(\rho_2 \parallel \gamma_2) &= tr(A \log A - A \log B) \\
&\ge tr(A - B) \\
&\ge 0
\end{aligned}$$

de tal manera que si en **(MPT)** hay igualdad, entonces

$$\begin{aligned} 0 &= tr\left(A\log A - A\log B\right) \\ &= tr\left(A - B\right) \end{aligned}$$

y por la condición correspondiente en la *Desigualdad de Klein* (Cor. 3.9) concluimos que $A = B$ ó lo que es lo mismo $\rho_{12} = e^{\log\gamma_{12}+\log\rho_2-\log\gamma_2}$.
Recíprocamente si $\log\rho_{12} - \log\gamma_{12} = I_1\otimes\log\rho_2 - I_1\otimes\log\gamma_2$, entonces $\log\rho_{12} = \log\gamma_{12} + I_1\otimes\log\rho_2 - I_1\otimes\log\gamma_2$, o sea $\log A = \log B$ y en consecuencia

$$\begin{aligned} 0 &= trA\left(\log A - \log B\right) \\ &= S\left(\rho_{12}\parallel\gamma_{12}\right) - S\left(\rho_2\parallel\gamma_2\right). \quad \textbf{QED} \end{aligned}$$

Teorema 10 *(Monotonía de la entropía relativa bajo operadores completamente positivos) Sea $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ un operador* CP *que preserva la traza. Si ρ y γ son operadores de densidad estrictamente positivos en $\mathcal{H}_A$, entonces*

$$S\left(\Phi(\rho)\parallel\Phi(\gamma)\right) \le S\left(\rho\parallel\gamma\right), \tag{MONO}$$

con igualdad si, y sólo si, $\widehat{\Phi}\left[\log\Phi(\rho) - \log\Phi(\gamma)\right] = \log\rho - \log\gamma$, donde $\widehat{\Phi}$ denota el adjunto de Φ respecto al producto interno de Hilbert -Schmidt.

Demostración. Supongamos que Φ tiene la representación de *Krauss* $\Phi(\varrho) = \sum_{i=1}^{n} E_i\varrho E_i^*$. Por la representación de *Lindblad* (Teorema 2.10, c)), existe una isometría parcial $V : \mathcal{H}_A \longrightarrow \mathcal{H}_B\otimes\mathbb{C}^n$ tal que

$$\Phi(\varrho) = tr_2 V\varrho V^* \quad \varrho\in\mathcal{B}(\mathcal{H}_A);$$

y por el teorema previo (Monotonía bajo la traza parcial), tenemos que

$$\begin{aligned} S\left(\Phi(\rho)\parallel\Phi(\gamma)\right) &= S\left(tr_2 V\rho V^*\parallel tr_2 V\gamma V^*\right) \\ &\le S\left(V\rho V^*\parallel V\gamma V^*\right) \\ &= S\left(\rho\parallel\gamma\right), \end{aligned}$$

la última igualdad se debe a que $trV\rho V^*\log V\gamma V^* = tr\rho\log\gamma$, ver [16] para la isometría parcial V. Esto muestra que se cumple **(MONO)**.

Demostremos que la condición de igualdad es necesaria. Usando de nuevo que para el operador V se cumple que $S(\rho\parallel\gamma) = S(V\rho V^*\parallel V\gamma V^*)$ se obtiene

$$S\left(\Phi(\rho)\parallel\Phi(\gamma)\right) = S\left(\rho\parallel\gamma\right) \Leftrightarrow S\left(tr_2 V\rho V^*\parallel tr_2 V\gamma V^*\right) = S\left(V\rho V^*\parallel V\gamma V^*\right)$$

y por la condición de igualdad en el Teorema 4.9 (*Monotonía bajo la traza parcial*), lo anterior implica que si $I_2\in\mathcal{B}(\mathbb{C}^n)$ es el operador identidad, entonces

$$\begin{aligned} \left(\log tr_2 V\rho V^* - \log tr_2 V\gamma V^*\right)\otimes I_2 &= \log V\rho V^* - \log V\gamma V^* \qquad (4.1) \\ &= V\left(\log\rho - \log\gamma\right)V^*, \end{aligned}$$

y como $V^*V = I_{\mathcal{H}_A}$, lo que se puede demostrar a partir de la construcción de V en el Teorema 2.10, concluimos que si se alcanza la igualdad en **(MONO)**, entonces

$$V^* \left[\left(\log \Phi(\rho) - \log \Phi(\gamma) \right) \otimes I_2 \right] V = \log \rho - \log \gamma; \tag{4.2}$$

y usando la igualdad $\widehat{\Phi}(\varrho) = V^* (\varrho \otimes I_2) V, \ \ \forall \varrho \in \mathcal{B}(\mathcal{H}_B)$, (Teorema 2.10, d) concluimos que la condición es necesaria.

Recíprocamente, como $\widehat{\Phi}$ es el adjunto de Φ respecto del producto interno de *Hilbert-Schmidt* se cumple $tr\sigma^* \widehat{\Phi}(\gamma) = tr \left[\Phi(\sigma) \right]^* \gamma$, para todos $\sigma \in \mathcal{B}(\mathcal{H}_A)$ y $\gamma \in \mathcal{B}(\mathcal{H}_B)$; usamos esta observación después de multiplicar por ρ la igualdad (4.2) en ambos miembros y tomar la traza para obtener que $tr\rho \widehat{\Phi} \left[\log \Phi(\rho) - \log \Phi(\gamma) \right] = tr \Phi(\rho) \left[\log \Phi(\rho) - \log \Phi(\gamma) \right]$ y de aquí

$$tr \Phi(\rho) \left[\log \Phi(\rho) - \log \Phi(\gamma) \right] = tr \left[\rho \left(\log \rho - \log \gamma \right) \right]. \ \ \textbf{QED}$$

Corolario 11 *Una condición necesaria para que se alcance la igualdad en MONO es que*

$$n \left[\log \Phi(\rho) - \log \Phi(\gamma) \right] = \Phi\left(\log \rho - \log \gamma \right).$$

donde n es el número de sumandos en la representación de Krauss de Φ.

Demostración Se sigue de la igualdad (4.1) que aparece líneas arriba y del inciso c) de la Proposición 2.2. **QED**

Antes de continuar argumentaremos lo afirmado en la introducción acerca de que la SSA_1 se puede ver como una propiedad de la entropía relativa y de la monotonía de la traza parcial.

Proposición 12 *La monotonía de la entropía relativa bajo la traza parcial implica la subaditividad fuerte.*

Demostración. Sea $\varrho_{123} \in \mathcal{B}(\mathcal{H}_{123})$ con $tr\varrho_{123} = 1$ y $\varrho_{123} > 0$. Usando el Lema 2.3 y la hipótesis se cumple que

$$S\left(\varrho_{12} \parallel I_1 \otimes \varrho_2 \right) \leq S\left(\varrho_{123} \parallel I_1 \otimes \varrho_{23} \right).$$

Por la definición de $S(\cdot \parallel \cdot)$ y la propiedad (TP2) de la traza parcial tenemos,

$$S\left(\varrho_{12} \parallel I_1 \otimes \varrho_2 \right) = tr\varrho_{12} \log \varrho_{12} - tr\varrho_{12} \left(I_1 \otimes \log \varrho_2 \right) = tr\varrho_{12} \log \varrho_{12} - tr\varrho_2 \log \varrho_2$$

y $S\left(\varrho_{123} \parallel I_1 \otimes \varrho_{23} \right) = tr\varrho_{123} \log \varrho_{123} - tr\varrho_{123} \left(I_1 \otimes \log \varrho_{23} \right) = tr\varrho_{123} \log \varrho_{123} - tr\varrho_{23} \log \varrho_{23}$, así la anterior desigualdad equivale a

$$-S(\varrho_{12}) + S(\varrho_2) \leq -S(\varrho_{123}) + S(\varrho_{23})$$

y ésta a su vez, después de reagruparla vemos que equivale a ($\mathbf{SSA}_1$). **QED**

Teorema 13 *(Convexidad conjunta de la entropía relativa). Sean* $\lambda_1, \ldots, \lambda_n$ *números reales no negativos tales que* $\sum_k \lambda_k = 1$*. Si* $\rho = \sum_k \lambda_k \rho_k$ *y* $\gamma = \sum_k \lambda_k \gamma_k$ *donde para cada* k*,* $\rho_k\ ,\gamma_k \in \mathcal{B}(\mathcal{H})$ *son operadores de densidad estrictamente positivos, entonces*

$$S(\rho \parallel \gamma) \leq \sum_k \lambda_k S(\rho_k \parallel \gamma_k). \tag{JCX}$$

Con igualdad si, y sólo si, $\log\rho - \log\gamma = \log\rho_k - \log\gamma_k$ *para toda* k*.*

Demostración. Para cada k, poniendo $A_k = \rho_k$ y eligiendo B_k tal que $\log B_k = \log\gamma_k + \log\rho - \log\gamma$ en la *Desigualdad de Klein* tenemos

$$tr\rho_k \log\rho_k - tr\rho_k\left(\log\gamma_k + \log\rho - \log\gamma\right) \geq tr\left(\rho_k - e^{\log\rho - \log\gamma + \log\gamma_k}\right) \tag{C1}$$

Multiplicamos esta desigualdad por λ_k y después sumamos para obtener

$$\begin{aligned}\sum_k \lambda_k S(\rho_k \parallel \gamma_k) - S(\rho \parallel \gamma) &\geq tr\left(\rho - \sum_k \lambda_k e^{\log\rho - \log\gamma + \log\gamma_k}\right) \\ &\geq tr\left(\rho - e^{\log\rho - \log\gamma + \log\sum_k \lambda_k\gamma_k}\right) \\ &= tr\left(\rho - e^{\log\rho}\right) = 0,\end{aligned}$$

la segunda desigualdad es por la concavidad (Teorema 6), de $A \mapsto tr\ e^{K+\log A}$ con $K = \log\rho - \log\gamma$ y $A > 0$.

Discutamos la condición de igualdad: si para algún k hubiera desigualdad estricta en (C1), entonces por el desarrollo anterior tendríamos $\sum_k \lambda_k S(\rho_k \parallel \gamma_k) - S(\rho \parallel \gamma) > 0$. Por lo tanto, si $\sum_k \lambda_k S(\rho_k \parallel \gamma_k) = S(\rho \parallel \gamma)$, necesariamente hay igualdad en (C1) para toda k y la condición de igualdad en la *Desigualdad de Klein* obliga a que $\log\rho - \log\gamma = \log\rho_k - \log\gamma_k$ para toda k.

Recíprocamente, si $\log\rho - \log\gamma = \log\rho_k - \log\gamma_k$ para toda k, entonces después de multiplicar por $\lambda_k\rho_k$ ambos miembros de esta igualdad, tomar la traza y sumar obtenemos $S(\rho \parallel \gamma) = \sum_k \lambda_k S(\rho_k \parallel \gamma_k)$. **QED**

4.4 Relaciones entre las desigualdades

Teorema 14 *Las siguientes proposiciones son equivalentes:*

1. **Monotonía de la entropía relativa bajo operadores completamente positivos** CP **que preservan la traza**.
Sea $\Phi : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$ un operador completamente positivo que preserva la traza. Si ρ y γ son operadores de densidad estrictamente positivos en $\mathcal{H}_A$, entonces

$$S(\Phi(\rho) \parallel \Phi(\gamma)) \leq S(\rho \parallel \gamma), \qquad \text{(MONO)}$$

2. Monotonía de la entropía relativa bajo la traza parcial.
Si $\rho_{12}, \gamma_{12} \in \mathcal{B}(\mathcal{H}_{12})$ son operadores de densidad estrictamente positivos, entonces

$$S(\rho_2 \parallel \gamma_2) \leq S(\rho_{12} \parallel \gamma_{12}). \qquad \text{(MPT)}$$

3. Subaditividad fuerte de la entropía de von Neumann.
Si $\varrho_{123} \in \mathcal{B}(\mathcal{H}_{123})$ es un operador de densidad estrictamente positivo, entonces se cumplen las desigualdades $\mathbf{SSA}_1$ y $\mathbf{SSA}_2$, además son equivalentes entre sí,

$$S(\varrho_{123}) + S(\varrho_2) \leq S(\varrho_{12}) + S(\varrho_{23}) \qquad (\text{SSA}_1)$$

$$S(\varrho_1) + S(\varrho_2) \leq S(\varrho_{13}) + S(\varrho_{23}). \qquad (\text{SSA}_2)$$

4. Convexidad conjunta de la entropía relativa.
Sean $\lambda_1, \ldots, \lambda_n$ números reales *no negativos* tales que $\sum_k \lambda_k = 1$. Si $\rho = \sum_k \lambda_k \rho_k$ y $\gamma = \sum_k \lambda_k \gamma_k$ donde para cada k, ρ_k ,$\gamma_k \in \mathcal{B}(\mathcal{H})$ son operadores de densidad estrictamente positivos, entonces

$$S(\rho \parallel \gamma) \leq \sum_k \lambda_k S(\rho_k \parallel \gamma_k) \qquad \text{(JCX)}$$

5. Concavidad de la entropía condicional.
Sean $\lambda_1, \ldots, \lambda_n$ números reales *no negativos* tales que $\sum_i \lambda_i = 1$. Si para cada i, $\rho_{12}^i \in \mathcal{B}(\mathcal{H}_{12})$, es un operador de densidad estrictamente positivo y $\rho_2^i \equiv tr_1 \rho_{12}^i$ entonces

$$\sum_i \lambda_i \left[S\left(\rho_{12}^i\right) - S\left(\rho_2^i\right)\right] \leq S\left(\sum_i \lambda_i \rho_{12}^i\right) - S\left(\sum_i \lambda_i \rho_2^i\right). \qquad \text{(CON)}$$

Es decir, la función $\rho_{12} \longmapsto S(\rho_{12}) - S(\rho_2)$ es cóncava.

La demostración se hará de acuerdo al siguiente arreglo

$$\begin{array}{ccccc} JCX & \Rightarrow & SSA_2 & \Rightarrow & SSA_1 \\ \Uparrow & & & & \Downarrow \\ MONO & \Leftarrow & MPT & \Leftarrow & CON \end{array}$$

DEMOSTRACIÓN.

JCX⇒SSA$_2$
Sea $\sigma_1 \equiv \frac{1}{d} I_1 \in \mathcal{B}(\mathcal{H}_1)$ donde $d = \dim \mathcal{H}_1$.
Definimos $h : \mathcal{B}(\mathcal{H}_{123}) \longrightarrow \mathbb{R}$ con

$$h(\rho_{123}) = S(\rho_{13} \parallel \sigma_1 \otimes \rho_3).$$

Usando la hipótesis y la linealidad de la traza parcial veremos que h es una función convexa. Para ello, sean $\{\lambda_i\}$ reales tales que $0 \leq \lambda_i \leq 1$ con $\sum_i \lambda_i = 1$ y sean $\rho_{123}^i \in \mathcal{B}(\mathcal{H}_{123})$; evaluando tenemos que

$$h\left(\sum_i \lambda_i \rho_{123}^i\right) = S\left(\sum_i \lambda_i \rho_{123}^i \parallel \sigma_1 \otimes \left(\sum_i \lambda_i \rho_3^i\right)\right) = S\left(\sum_i \lambda_i \rho_{123}^i \parallel \sum_i \lambda_i \sigma_1 \otimes \rho_3^i\right)$$

$$\leq \sum_i \lambda_i S\left(\rho_{123}^i \parallel \sigma_1 \otimes \rho_3^i\right) = \sum_i \lambda_i h\left(\rho_{123}^i\right).$$

Ahora observemos que $S(\rho_{13} \parallel \sigma_1 \otimes \rho_3) = tr\rho_{13} \log \rho_{13} - tr\rho_{13} \log(\sigma_1 \otimes \rho_3) = -S(\rho_{13}) - tr\rho_{13}(\log \sigma_1) \otimes I_3 - tr\rho_{13}(I_1 \otimes \log \rho_3) = -S(\rho_{13}) + S(\sigma_1) + S(\rho_3) = -S(\rho_{13}) + \log d + S(\rho_3)$, es decir

$$-S(\rho_{13}) + S(\rho_3) = S(\rho_{13} \parallel \sigma_1 \otimes \rho_3) - \log d,$$

por lo anterior la función g con dominio en las *matrices de densidad* $\rho_{123} \geq 0$, con regla de correspondencia $g(\rho_{123}) \equiv -S(\rho_{13}) + S(\rho_3)$ es una función convexa en $\rho_{123} > 0$. Extenderemos esta propiedad al caso $\rho_{123} \geq 0$, para lo cual es suficiente ver que los operadores estrictamente positivos son densos, con la norma de operadores, en el conjunto $D_+ \equiv \{\varrho \in \mathcal{B}(\mathcal{H}_{123}) \ / \ \varrho \geq 0 \text{ y } tr\varrho = 1\}$ y que la función g es continua con la norma de operadores en dicho conjunto.
La continuidad de g ocurre porque las funciones $A \to S(A)$, $A \in D_+$ y la traza parcial son funciones continuas con la norma de operadores, lo primero se cumple por la observación b) hecha después de la definición de la entropía de *von Neumann*, y lo segundo es porque la traza parcial es un operador CP y estos, no es difícil comprobarlo, son continuos con la norma de operadores.

Para ver la densidad del conjunto de matrices estrictamente positivas en D_+, sea $\rho_{123} = \sum_j \lambda_j |\psi_j\rangle \langle\psi_j|$, con $\{\psi_j\}$ una base ortonormal de $\mathcal{H}_{123}$. Si definimos para cada $n \in \mathbb{N}$, $\alpha_n = \left(1 + \frac{k_0}{n}\right)$, donde $k_0 = \dim \ker \rho_{123}$ y

$$\mu_j = \begin{cases} \lambda_j, & \text{si } \lambda_j > 0 \\ \frac{1}{n}, & \text{si } \lambda_j = 0 \end{cases},$$

entonces $\rho_{123}^n \equiv \sum_j \alpha_n^{-1}\mu_j \left|\psi_j\right\rangle \left\langle\psi_j\right|$, es una sucesión de operadores de densidad estrictamente positivos que convergen en la norma de operadores a ρ_{123}.

Así, podemos concluir que la función definida en los operadores de densidad $\rho_{123} \geq 0$, mediante

$$f(\rho_{123}) \equiv -S(\rho_{13}) + S(\rho_3) - S(\rho_{12}) + S(\rho_2)$$

es convexa; además si $\psi \in \mathcal{H}_{123}$ es de norma 1, entonces el operador $\rho_{123} = |\psi\rangle\langle\psi|$ satisface $f(\rho_{123}) = 0$ pues por el Teorema 3.11 c) se cumple que $S(\rho_3) = S(\rho_{12})$ y $S(\rho_2) = S(\rho_{13})$.

Para terminar, por el teorema espectral todo $\rho_{123} \in \mathcal{B}(\mathcal{H}_{123})$ se puede expresar en la forma $\rho_{123} = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ con vectores ortonormales $\psi_i \in \mathcal{H}_{123}$ y por la convexidad de f es válida la desigualdad $f(\rho_{123}) \leq \sum_i \lambda_i f(|\psi_i\rangle\langle\psi_i|) = 0$, es decir, para toda matriz de densidad $\rho_{123} \in \mathcal{B}(\mathcal{H}_{123})$ se verifica que

$$0 \geq f(\rho_{123}) \equiv -S(\rho_{13}) + S(\rho_3) - S(\rho_{12}) + S(\rho_2). \quad \textbf{QED}$$

SSA$_2$ ⇒SSA$_1$
Dado $\varrho_{123} \in \mathcal{B}(\mathcal{H}_{123})$ con $tr\varrho_{123} = 1$, por el Lema 2.6 obtenemos un estado puro ϱ_{1234}, tal que $\varrho_{123} = tr_4\varrho_{1234}$ y entonces

$$\begin{aligned} S(\varrho_{123}) + S(\varrho_2) &= S(\varrho_4) + S(\varrho_2) \\ &\leq S(\varrho_{14}) + S(\varrho_{12}) \\ &= S(\varrho_{23}) + S(\varrho_{12}), \end{aligned}$$

las igualdades son por el Teorema 3.11 c) y la desigualdad es por la hipótesis. **QED**

SSA$_1$ ⇒CON.
Para cada $i = 1, 2, ..., n$ sean $\rho_{12}^i \in \mathcal{B}(\mathcal{H}_{12})$ y $\{\lambda_i\}$ un subconjunto de números reales tales que $\lambda_i \geq 0$ y $\sum_i \lambda_i = 1$.

Introducimos un espacio auxiliar $\mathcal{H}_3$ de dimensión n. Sea $\{|i\rangle\}$ una base de $\mathcal{H}_3$ y definimos

$$\rho_{123} = \sum_i \lambda_i \rho_{12}^i \otimes |i\rangle\langle i|,$$

luego por la Proposición 2.2 c) se cumple que $\rho_{12} = tr_3\rho_{123} = \sum_i \lambda_i \rho_{12}^i$ y por el Lema 2.5 a) y b) respectivamente, $\rho_2 = tr_{13}\rho_{123} = \sum_i \lambda_i \rho_2^i$, $\rho_{23} = tr_1\rho_{123} = \sum_i \lambda_i \rho_2^i \otimes |i\rangle\langle i|$, ahora usamos el Teorema 3.11 g) para obtener $S(\rho_{123}) = H(\lambda_i) + \sum_i \lambda_i S(\rho_{12}^i)$, $S(\rho_{23}) = H(\lambda_i) + \sum_i \lambda_i S(\rho_2^i)$ y al usar la hipótesis, tenemos

$$S(\varrho_{123}) + S(\varrho_2) \leq S(\varrho_{12}) + S(\varrho_{23}) \qquad (\text{SSA}_1)$$

que equivale a las siguientes desigualdades

$H(\lambda_i) + \sum_i \lambda_i S(\rho_{12}^i) + S(\sum_i \lambda_i \rho_2^i) \leq S(\sum_i \lambda_i \rho_{12}^i) + H(\lambda_i) + \sum_i \lambda_i S(\rho_2^i)$,

y $\sum_i \lambda_i \left[S(\rho_{12}^i) - S(\rho_2^i)\right] \leq S(\sum_i \lambda_i \rho_{12}^i) - S(\sum_i \lambda_i \rho_2^i)$. **QED**

Observación: Para $\rho, \gamma \in \mathcal{B}(\mathcal{H}_2)$, tales que $\rho, \gamma > 0$ y para cualquier real no negativo x se cumple,

$$\frac{d}{dx} S(\gamma + x\rho) = -tr\rho \log(\gamma + x\rho) - tr\rho.$$

En efecto:

$\frac{d}{dx} S(\gamma + x\rho) = \frac{d}{dx}\left[-tr(\gamma + x\rho)\log(\gamma + x\rho)\right]$

$= -tr\frac{d}{dx}\left[(\gamma + x\rho)\log(\gamma + x\rho)\right]$

$= -tr\left[\rho\log(\gamma + x\rho) + (\gamma + x\rho)(\gamma + x\rho)^{-1}\rho\right]$

$= -tr\left[\rho\log(\gamma + x\rho) + \rho\right] = -tr\rho\log(\gamma + x\rho) - tr\rho.$

CON⇒MPT.

Sean $\rho_{12}, \gamma_{12} \in \mathcal{B}(\mathcal{H}_2)$, demostraremos que

$$S(\rho_1 \parallel \gamma_1) \leq S(\rho_{12} \parallel \gamma_{12}). \tag{MPT}$$

La hipótesis implica que la función

$$f(\varrho_{12}) = S(\varrho_1) - S(\varrho_{12})$$

es **convexa**, ahora veremos que es homogénea: sea $x \in \mathbb{R}, x > 0$ calculando vemos que $S(x\rho_1) = -trx\rho_1 \log x\rho_1 = -trx\rho_1(\log xI_1 + \log\rho_1) = -x\log xtr\rho_1 - xtr\rho_1\log\rho_1$ y similarmente $S(x\rho_{12}) = -x\log xtr\rho_{12} - xtr\rho_{12}\log\rho_{12}$, usando ambas relaciones se sigue que $S(x\varrho_1) - S(x\varrho_{12}) = x[S(\varrho_1) - S(\varrho_{12})]$. Luego el Lema A_6 implica que

$$\lim_{x\to 0}\frac{f(\gamma_{12} + x\rho_{12}) - f(\gamma_{12})}{x} \leq f(\rho_{12}),$$

si definimos $g(x) = f(\gamma_{12} + x\rho_{12})$, entonces la desigualdad anterior significa

$$g'(0) \leq f(\rho_{12}). \tag{D$_1$}$$

Usando la observación previa al inicio de esta demostración y recordando que $tr\rho_1 = tr\rho_{12}$, tenemos $\frac{d}{dx}g(x)\mid_{x=0} = \frac{d}{dx}\mid_{x=0}(S(\gamma_1 + x\rho_1) - S(\gamma_{12} + x\rho_{12})) = -tr\rho_1\log\gamma_1 + tr\rho_{12}\log\gamma_{12}$, por eso **($\mathbf{D_1}$)** equivale a las siguientes desigualdades

$$\begin{aligned} -tr\rho_1\log\gamma_1 + tr\rho_{12}\log\gamma_{12} &\leq S(\rho_1) - S(\rho_{12}), \\ -S(\rho_1) - tr\rho_1\log\gamma_1 &\leq -S(\rho_{12}) - tr\rho_{12}\log\gamma_{12}, \\ S(\rho_1 \parallel \gamma_1) &\leq S(\rho_{12} \parallel \gamma_{12}). \ \textbf{QED} \end{aligned}$$

Observemos que si hubiésemos definido inicialmente $f(\varrho_{12}) = S(\varrho_2) - S(\varrho_{12})$ tendríamos $S(\rho_2 \parallel \gamma_2) \leq S(\rho_{12} \parallel \gamma_{12})$.

MPT⇒MONO.

De hecho esta implicación fue la que demostramos en la prueba del **Teorema 10.**

MONO⇒JCX.

En la proposición 2 de **este** Teorema, elegimos ρ_{12} (y similarmente γ_{12}) como una matriz con bloques en la diagonal iguales a $\lambda_k \rho_k$ (resp. $\lambda_k \gamma_k$), entonces

$$\begin{aligned}
\rho_{12} \log \rho_{12} &= \begin{bmatrix} \lambda_1\rho_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n\rho_n \end{bmatrix} \begin{bmatrix} \log\lambda_1 I + \log\rho_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \log\lambda_n I + \log\rho_n \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1\rho_1 \log\lambda_1 I + \lambda_1\rho_1 \log\rho_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n\rho_n \log\lambda_n I + \lambda_n\rho_n \log\rho_n \end{bmatrix}
\end{aligned}$$

donde hemos usado que $\log(\lambda\varrho) = (\log\lambda) I + \log\varrho$ para todo $\lambda > 0$ y $\varrho > 0$ (Lema 1.13), de manera similar encontramos que

$$\rho_{12} \log \gamma_{12} = \begin{bmatrix} \lambda_1\rho_1 \log\lambda_1 I + \lambda_1\rho_1 \log\gamma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n\rho_n \log\lambda_n I + \lambda_n\rho_n \log\gamma_n \end{bmatrix}$$

por tanto

$$\rho_{12} \log \rho_{12} - \rho_{12} \log \gamma_{12} = \begin{bmatrix} \lambda_1\rho_1 (\log\rho_1 - \log\gamma_1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n\rho_n (\log\rho_n - \log\gamma_n) \end{bmatrix},$$

al tomar la traza obtenemos $S(\rho_{12} \parallel \gamma_{12}) = tr(\rho_{12} \log \rho_{12} - \rho_{12} \log \gamma_{12}) = \sum_i \lambda_i S(\rho_i \parallel \gamma_i) \geq S(\rho_2 \parallel \gamma_2)$ donde la última desigualdad es por la hipótesis, ya que la traza parcial es un operador CP que preserva la traza (Proposición 2.8).

Para terminar, es suficiente demostrar que

$$\begin{aligned}
\rho_2 &\equiv tr_1 \rho_{12} = \sum_k \lambda_k \rho_k \\
&\text{y} \\
\gamma_2 &\equiv tr_1 \gamma_{12} = \sum_k \lambda_k \gamma_k
\end{aligned}$$

lo que haremos usando la Proposición 2.1 como sigue: sea $Y \in \mathcal{B}(\mathcal{H}_2)$, luego

$$\begin{aligned}
tr\rho_{12}(I_1 \otimes Y) &= tr\left[\begin{array}{ccc} \lambda_1\rho_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n\rho_n \end{array}\right]\left[\begin{array}{ccc} Y & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & Y \end{array}\right] \\
&= tr\left[\begin{array}{ccc} \lambda_1\rho_1 Y & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n\rho_n Y \end{array}\right] = \sum_k tr(\lambda_k\rho_k Y) \\
&= tr\left[\left(\sum_k \lambda_k\rho_k\right)Y\right],
\end{aligned}$$

por la unicidad de la traza parcial concluimos que

$$\sum_k \lambda_k\rho_k = \rho_2,$$

y análogamente

$$\sum_k \lambda_k\gamma_k = \gamma_2;$$

con esto terminamos el ciclo. **QED**

Capítulo 5

Aplicaciones, conclusiones y perspectivas

5.1 Subaditividad de la entropía

Como aplicación de la *subaditividad fuerte* tenemos la siguiente propiedad conocida como la *subaditividad de la entropía;* daremos dos demostraciones de ésta, en la primera usaremos como ya quedó dicho, la subaditividad fuerte **(SSA)** y en la segunda volverá a aparecer la **Desigualdad de Klein** con su condición de igualdad.

Teorema 1 *(Subaditividad). Si $\varrho_{12} \in \mathcal{B}(\mathcal{H}_{12})$ es un operador de densidad estrictamente positivo, entonces*

$$S(\varrho_{12}) \leq S(\varrho_1) + S(\varrho_2),$$

con igualdad si, y sólo si, $\varrho_{12} = \varrho_1 \otimes \varrho_2$.

Demostración 1 (no incluye condición de igualdad). Consideremos el operador $\gamma_{123} = \varrho_{12} \otimes 1 \in \mathcal{B}(\mathcal{H}_{12} \otimes \mathbb{C})$, por la **(SSA)** tenemos que $S(\gamma_{123}) + S(\gamma_3) \leq S(\gamma_{13}) + S(\gamma_{23})$, es decir

$$S(\varrho_{12} \otimes 1) + S(1) \leq S(\varrho_1 \otimes 1) + S(\varrho_2 \otimes 1).$$

Por el inciso d) del teorema 3.11 se cumplen las igualdades $S(\varrho_{12} \otimes 1) = S(\varrho_{12}) + S(1)$, $S(\varrho_1 \otimes 1) = S(\varrho_1) + S(1)$, y $S(\varrho_2 \otimes 1) = S(\varrho_2) + S(1)$, al sustituirlas en la desigualdad anterior se obtiene el resultado. **QED**

Demostración 2 (incluye condición de igualdad). Usando la definición de entropía, y por la propiedad TP2, (Proposición 2.1) de la traza parcial se cumple que

$S(\varrho_1) + S(\varrho_2) - S(\varrho_{12}) = -tr\varrho_1 \log \varrho_1 - tr\varrho_2 \log \varrho_2 + tr\varrho_{12} \log \varrho_{12} =$

$-tr\varrho_{12} (\log \varrho_1 \otimes I_2) - tr\varrho_{12} (I_1 \otimes \log \varrho_2) + tr\varrho_{12} \log \varrho_{12} =$

$tr\varrho_{12} [\log \varrho_{12} - (\log \varrho_1 \otimes I_2) - (I_1 \otimes \log \varrho_2)] = tr\varrho_{12} [\log \varrho_{12} - \log (\varrho_1 \otimes \varrho_2)]\,,$

donde hemos usado que $\log(\sigma \otimes \rho) = (\log \sigma_1) \otimes I_2 - I_1 \otimes (\log \gamma)$, (Lema 1.14) y por la *Desigualdad de Klein* (Corolario 3.9); concluimos que

$$\begin{aligned} S(\varrho_1) + S(\varrho_2) - S(\varrho_{12}) &= tr\varrho_{12} [\log \varrho_{12} - \log (\varrho_1 \otimes \varrho_2)] \\ &\geq tr [\varrho_{12} - \varrho_1 \otimes \varrho_2] = 0 \end{aligned}$$

lo que prueba la desigualdad.

Para la condición de igualdad si $S(\varrho_1) + S(\varrho_2) - S(\varrho_{12}) = 0$, entonces usando las desigualdades de arriba y por la condición correspondiente en la *Desigualdad de Klein*, concluimos que $\varrho_{12} = \varrho_1 \otimes \varrho_2$
Recíprocamente, cuando se cumple esta última igualdad tenemos que

$$0 = tr\varrho_{12} [\log \varrho_{12} - \log (\varrho_1 \otimes \varrho_2)] = S(\varrho_1) + S(\varrho_2) - S(\varrho_{12})\,. \quad \textbf{QED}$$

5.2 Concavidad de la entropía

Teorema 2 *Dada una distribución de probabilidad* $\{p_1, ..., p_n\}$ *(números reales no negativos tales que* $\sum_i p_i = 1$*) y operadores de densidad* $\{\rho_1, ..., \rho_n\} \subseteq \mathcal{B}(\mathcal{H}_A)$, *se cumple que*

$$S\left(\sum_{i=1}^{n} p_i \rho_i\right) \geq \sum_{i=1}^{n} p_i S(\rho_i)\,,$$

con igualdad si, y sólo si, todos los ρ_i *son iguales.*

Demostración. Usamos la *subaditividad de la entropía* para lo cual, introducimos un espacio auxiliar $\mathcal{H}_B$ de dimensión n (el número de sumandos) con una base ortonormal $\{|i\rangle\}$. Consideremos el estado

$$\rho_{AB} \equiv \sum_i p_i \rho_i \otimes |i\rangle \langle i|$$

en $\mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$, aplicando el inciso g) del teorema 3.11 obtenemos que

$$S(\rho_{AB}) = S\left(\sum_i p_i \rho_i \otimes |i\rangle \langle i|\right) = H(p_i) + \sum_i p_i S(\rho_i)\,.$$

Notemos que $\rho_A \equiv tr_B \rho_{AB} = tr_B \left(\sum_i p_i \rho_i \otimes |i\rangle \langle i| \right) = \sum_i p_i tr_B \left(\rho_i \otimes |i\rangle \langle i| \right) = \sum_i p_i \rho_i$ y por ende

$$S(\rho_A) = S\left(\sum_i p_i \rho_i \right),$$

análogamente $\rho_B \equiv tr_A \rho_{AB} = tr_A \left(\sum_i p_i \rho_i \otimes |i\rangle \langle i| \right) = \sum_i tr_A \left(p_i \rho_i \otimes |i\rangle \langle i| \right) = \sum_i tr\left(p_i \rho_i \right) |i\rangle \langle i| = \sum_i p_i |i\rangle \langle i|$ y

$$S(\rho_B) = S\left(\sum_i p_i |i\rangle \langle i| \right) = H(p_i),$$

para terminar sustituimos las expresiones anteriores en la desigualdad de *subaditividad* $S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$, y después de simplificar se obtiene la desigualdad deseada.

Antes de discutir la condición de igualdad necesitamos el siguiente

Lema:
Cuando $f_1 \otimes g_1$ y $f_2 \otimes g_2$ son elementos distintos de cero en un producto tensorial de espacios de *Hilbert*, entonces $f_1 \otimes g_1 = f_2 \otimes g_2$ si, y sólo si, existe un escalar c distinto de cero tal que $f_1 = cf_2$ y $g_1 = c^{-1} g_2$.

Demostración. Demostraremos sólo la condición necesaria porque la otra se comprueba fácilmente. Por hipótesis, $0 \neq f_1 \otimes g_1 = f_2 \otimes g_2$, en particular f_1, g_1, f_2 y g_2 son elementos **distintos de cero**; luego $0 < \|f_1\|^2 \|g_1\|^2 = \langle f_1, f_1 \rangle \langle g_1, g_1 \rangle = \langle f_1 \otimes g_1, f_1 \otimes g_1 \rangle = \langle f_1 \otimes g_1, f_2 \otimes g_2 \rangle = \langle f_1, f_2 \rangle \langle g_1, g_2 \rangle$, es decir

$$\|f_1\|^2 \|g_1\|^2 = \langle f_1, f_2 \rangle \langle g_1, g_2 \rangle, \tag{5.1}$$

y por consiguiente, el segundo término de la igualdad es estrictamente mayor que cero.
También se cumple que

$$\begin{aligned} \langle f_1, f_2 \rangle \langle g_1, g_2 \rangle &= \langle f_1 \otimes g_1, f_2 \otimes g_2 \rangle \\ &= \|f_1 \otimes g_1\| \|f_2 \otimes g_2\| \\ &= \|f_1\| \|g_1\| \|f_2\| \|g_2\| \end{aligned} \tag{5.2}$$

la segunda igualdad de este bloque es por la condición de igualdad en la *Desigualdad de Schwarz.* Esto último implica que existen escalares c,k distintos de cero tales que $f_1 = cf_2$ y $g_1 = kg_2$ porque de lo contrario usando nuevamente la *Desigualdad de Schwarz,* se tendría que $|\langle f_1, f_2 \rangle| < \|f_1\| \|f_2\|$ ó $|\langle g_1, g_2 \rangle| < \|g_1\| \|g_2\|$ y al multiplicarlas se cumpliría que $|\langle f_1, f_2 \rangle| |\langle g_1, g_2 \rangle| < \|f_1\| \|f_2\| \|g_1\| \|g_2\|$ contradiciendo las igualdades en el bloque **(5.2).**

Ahora en (**5.1**) sustituimos $f_2 = c^{-1} f_1$ y $g_2 = k^{-1} g_1$ para obtener

$$\begin{aligned} \|f_1\|^2 \|g_1\|^2 &= \left\langle f_1, c^{-1} f_1 \right\rangle \left\langle g_1, k^{-1} g_1 \right\rangle \\ &= c^{-1} k^{-1} \|f_1\|^2 \|g_1\|^2 \end{aligned}$$

como f_1, g_1 son distintos de cero podemos cancelar el término $\|f_1\|^2 \|g_1\|^2$ y concluir que $1 = c^{-1}k^{-1}$.

Continuemos con la discusión de la condición de igualdad del Teorema 2. Sólo probaremos la necesidad de la condición porque la suficiencia es inmediata. Conservando la notación usada en la primera parte de la prueba, si suponemos válida la igualdad $\sum_{i=1}^{n} p_i S(\rho_i) = S\left(\sum_{i=1}^{n} p_i \rho_i\right)$, después de sumar $H(p_i)$ en ambos lados obtenemos $S(\rho_{AB}) = S(\rho_A) + S(\rho_B)$, y por el Teorema 1 esto sólo es posible si $\rho_{AB} = \rho_A \otimes \rho_B$, usaremos esta igualdad después de algunos cálculos con las expresiones de ρ_{AB}, ρ_A y ρ_B como sigue:

Si $a \in \mathcal{H}_A$ y $|j\rangle$ es un elemento de la base $\{|i\rangle\}$, entonces

$\rho_{AB}(a \otimes |j\rangle) = \left(\sum_i p_i \rho_i \otimes |i\rangle\langle i|\right)(a \otimes |j\rangle) = \sum_i p_i \rho_i a \otimes (|i\rangle\langle i| \, |j\rangle) = \rho_j a \otimes p_j |j\rangle$,

y $(\rho_A \otimes \rho_B)(a \otimes |j\rangle) = \rho_A a \otimes \rho_B |j\rangle = \left(\sum_r p_r \rho_r a\right) \otimes \left(\sum_s p_s |s\rangle\langle s| \, |j\rangle\right) =$

$\left(\sum_r p_r \rho_r a\right) \otimes (p_j |j\rangle)$ y por lo tanto

$$\rho_j a \otimes p_j |j\rangle = \left(\sum_r p_r \rho_r a\right) \otimes (p_j |j\rangle);$$

por el lema para algún escalar no cero se cumple que $\rho_j a = c\left(\sum_r p_r \rho_r a\right)$ y $p_j |j\rangle = c^{-1} p_j |j\rangle$ y de aquí inferimos que $\rho_j a = \sum_r p_r \rho_r a$ para toda j es decir, $\rho_j a = \rho_k a$ para todo par *j,k*. **QED**

Como veremos en el **Corolario 6**, los siguientes resultados le darán más profundidad a la concavidad de la entropía y nos serán de utilidad posteriormente.

Lema 3 *(Las mediciones proyectivas incrementan la entropía) Si* $\{P_i\} \subseteq \mathcal{B}(\mathcal{H})$ *es un conjunto completo de proyecciones ortogonales (i.e.* $\sum_i P_i = I$ *) y* $\rho \in \mathcal{B}(\mathcal{H})$ *es un operador de densidad, entonces* $\rho' \equiv \sum_i P_i \rho P_i$ *tiene entropía al menos tan grande como la de* ρ,

$$S(\rho') \geq S(\rho)$$

con igualdad si, y sólo si, $\rho' = \rho$.

Antes de proceder con la demostración observemos que ρ' es un operador de densidad.

Demostración Por hipótesis $P_j P_i = \delta_{ij} P_i$ y esto implica que $\left(\sum_i P_i \rho P_i\right) P_j = P_j \rho P_j = P_j \left(\sum_i P_i \rho P_i\right)$, es decir que para toda j, ρ' y P_j conmutan y por ende también conmutan P_j y $\log \rho'$.

Ahora aplicamos que $\sum_i P_i = I$, $P_i^2 = P_i$, la ciclicidad de la traza y la observación anterior para realizar los siguientes cálculos:

$$\begin{aligned} tr\rho\log\rho' &= tr\left(\sum_i P_i\rho\log\rho'\right) = \sum_i tr\left(P_i\rho\log\rho'\right) \\ &= \sum_i tr\left(P_i\rho\log\rho' P_i\right) = \sum_i tr\left(P_i\rho P_i\log\rho'\right) \\ &= tr\left(\sum_i P_i\rho P_i\log\rho'\right) = tr\rho'\log\rho' \end{aligned}$$

y de aquí $-tr\rho\log\rho' = S(\rho')$; esta igualdad y la no negatividad de la entropía relativa (Teorema 3.10), implican que

$$0 \leq S(\rho \parallel \rho') = -S(\rho) - tr\rho\log\rho' = -S(\rho) + S(\rho') . \quad \textbf{QED}$$

Teorema 4 *Dada una distribución de probabilidad* $\{p_1, ..., p_n\}$ *(números reales no negativos tales que* $\sum_i p_i = 1$*) y vectores normalizados* $\{\psi_1, ..., \psi_n\} \subseteq \mathcal{H}_A$*, se cumple que*

$$S\left(\sum_{i=1}^{n} p_i \left|\psi_i\right\rangle\left\langle\psi_i\right|\right) \leq H(p_i)$$

con igualdad si, y sólo si, $\{\psi_1, ..., \psi_n\}$ *son ortogonales.*

Demostración Introducimos un espacio auxiliar $\mathcal{H}_B$ de dimensión n (el número de sumandos) con una base ortonormal $\{|i\rangle\}$. Consideremos el vector

$$\varphi \equiv \sum_i \sqrt{p_i}\psi_i \otimes |i\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B,$$

es fácil ver que $\langle\varphi,\varphi\rangle = 1$, y por ende $|\varphi\rangle\langle\varphi| = \sum_{i,j}\sqrt{p_i}\sqrt{p_j}\,|\psi_i\rangle\langle\psi_j| \otimes |i\rangle\langle j|$ es un estado puro.

Usando las propiedades de la traza parcial se puede mostrar que

$$\rho_A = tr_B\,|\varphi\rangle\langle\varphi| = \sum_i p_i\,|\psi_i\rangle\langle\psi_i| \qquad (*)$$

y

$$\rho_B = tr_A\,|\varphi\rangle\langle\varphi| = \sum_{i,j}\sqrt{p_i}\sqrt{p_j}\,\langle\psi_i,\psi_j\rangle\,|i\rangle\langle j| \qquad (**)$$

Ahora aplicamos el inciso c) del teorema 3.11 para obtener que

$$S(\rho_B) = S(\rho_A)$$

Definiendo el operador $\Psi : \mathcal{B}(\mathcal{H}_B) \longrightarrow \mathcal{B}(\mathcal{H}_B)$ tal que $\Psi(\varrho) = \sum_i P_i \varrho P_i$ donde $P_i = |i\rangle \langle i|$, luego de algunas operaciones podemos hallar que

$$\Psi(\rho_B) = \sum_i p_i |i\rangle \langle i| \qquad (***)$$

y por el lema anterior concluimos que $S(\rho_B) \leq S(\Psi \rho_B)$ es decir, $S(\rho_A) \leq S(\Psi \rho_B)$ pero por (*) y (***) esto demuestra la desigualdad enunciada en el teorema.

Por el Lema 3 se da la igualdad si, y sólo si, $\rho_B = \Psi(\rho_B)$ y como veremos a continuación esta condición equivale a que los vectores $\{\psi_i\}$ son ortonormales; en efecto:

$\Leftarrow$] Si $\{\psi_i\}$ son ortonormales,

$$\begin{aligned} \rho_B &= \sum_{i,j} \sqrt{p_i}\sqrt{p_j} \langle \psi_i, \psi_j \rangle |i\rangle \langle j|, \quad \text{por } (**) \\ &= \sum_{i,j} \sqrt{p_i}\sqrt{p_j} \delta_{ij} |i\rangle \langle j| \\ &= \sum_i p_i |i\rangle \langle i| = \Psi(\rho_B), \quad \text{por } (***). \end{aligned}$$

$\Rightarrow$] Si $\rho_B = \Psi(\rho_B)$, usando (**) tenemos que para $|m\rangle \in \{|i\rangle\}$ se cumple lo siguiente:

$\rho_B |m\rangle = \sum_{i,j} \sqrt{p_i}\sqrt{p_j} \langle \psi_i, \psi_j \rangle |i\rangle \langle j| |m\rangle = \sum_i \sqrt{p_i}\sqrt{p_m} \langle \psi_i, \psi_m \rangle |i\rangle$, mientras que usando (***),

$\Psi \rho_B |m\rangle = \sum_i p_i |i\rangle \langle i| |m\rangle = p_m |m\rangle$ e igualando ambas expresiones concluimos que para todo $|m\rangle \in \{|i\rangle\}$

$$\begin{aligned} 0 &= \sum_i \sqrt{p_i}\sqrt{p_m} \langle \psi_i, \psi_m \rangle |i\rangle - p_m |m\rangle \\ &= \sum_{i \neq m} \sqrt{p_i}\sqrt{p_m} \langle \psi_i, \psi_m \rangle |i\rangle + p_m [\langle \psi_m, \psi_m \rangle - 1] |m\rangle \end{aligned}$$

y por la independencia lineal del conjunto $\{|i\rangle\}$ se concluye $\langle \psi_i, \psi_m \rangle = \delta_{im}$. **QED**

Como consecuencia tenemos el resultado siguiente.

Teorema 5 *Dada una distribución de probabilidad $\{p_1, ..., p_n\}$ (números reales no negativos tales que $\sum_i p_i = 1$) y operadores de densidad $\{\rho_1, ..., \rho_n\} \subseteq \mathcal{B}(\mathcal{H})$, se cumple que*

$$S\left(\sum_i p_i \rho_i\right) \leq \sum_i p_i S(\rho_i) + H(p_i),$$

con igualdad si, y sólo si, $\{\rho_1, ..., \rho_n\}$ tienen soportes ortogonales.

Intuitivamente, esta desigualdad nos dice que nuestra incertidumbre sobre el estado promedio $\sum_i p_i\rho_i$ nunca es mayor que nuestra incertidumbre promedio acerca de los estados ρ_i, más la contribución adicional de $H(p_i)$.

Demostración Sean $\rho=\sum_i p_i\rho_i$, y para cada i, $\rho_i=\sum_j p_j^i \left|\varphi_j^i\right\rangle\left\langle\varphi_j^i\right|$ una descomposición ortonormal, entonces para cada i se cumple que $\sum_j p_j^i=1$ y $\rho=\sum_{i,j} p_i p_j^i \left|\varphi_j^i\right\rangle\left\langle\varphi_j^i\right|$; como $\sum_i\sum_j p_i p_j^i=\sum_i p_i\sum_j p_j^i=1\cdot 1$ podemos usar el teorema anterior para tener

$$S(\rho)\leq H\left(p_i p_j^i\right), \tag{*}$$

pero al usar la definición de entropía de *Shannon* y desarrollar tenemos

$$H\left(p_i p_j^i\right)=-\sum_{i,j} p_i p_j^i \log\left(p_i p_j^i\right)=-\sum_i p_i\log p_i\sum_j p_j^i-\sum_i p_i\sum_j p_j^i\log p_j^i=$$

$$-\sum_i p_i\log p_i+\sum_i p_i\left(-\sum_j p_j^i\log p_j^i\right)=H(p_i)+\sum_i p_i S(\rho_i).$$

También por el teorema anterior la igualdad se alcanza en **(*)** si, y sólo si, los elementos $\left\{\varphi_j^i\right\}$ son ortogonales y esta condición equivale a que los operadores $\rho_i=\sum_j p_j^i\left|\varphi_j^i\right\rangle\left\langle\varphi_j^i\right|$ tienen soportes ortogonales como veremos a continuación:

Recordemos, que el soporte de un operador T es el conjunto $[\ker T]^{\perp}$ y que es igual al conjunto $ranT$ si éste es de dimensión finita; entonces los operadores $\{\rho_i\}$ tienen soportes ortogonales si, y sólo, si $ran\rho_i\perp ran\rho_j$ para $i\neq j$.

$\Rightarrow]$ Sean $x,y\in\mathcal{H}$, luego $\langle\rho_i x,\rho_j y\rangle=\left\langle\sum_r p_r^i\left|\varphi_r^i\right\rangle\left\langle\varphi_r^i\right|x,\sum_s p_s^j\left|\varphi_s^j\right\rangle\left\langle\varphi_s^j\right|y\right\rangle=$ $\sum_r p_r^i\overline{\langle\varphi_r^i,x\rangle}\sum_s p_s^j\left\langle\varphi_s^j,y\right\rangle\left\langle\varphi_r^i,\varphi_s^j\right\rangle=0$

$\Leftarrow]$ Observemos que $\rho_i\varphi_r^i=p_r^i\varphi_r^i$ y como estamos suponiendo que tienen soportes ortogonales, para todos r,s, y para $i\neq j$ tenemos que

$$\begin{aligned}0&=\left\langle\rho_i\varphi_r^i,\rho_j\varphi_s^j\right\rangle=\left\langle p_r^i\varphi_r^i,p_s^j\varphi_s^j\right\rangle\\&=p_r^i p_s^j\left\langle\varphi_r^i,\varphi_s^j\right\rangle.\quad\textbf{QED}\end{aligned}$$

Corolario 6 *Dada una distribución de probabilidad* $\{p_1,...,p_n\}$ *y operadores de densidad* $\{\rho_1,...,\rho_n\}\subseteq\mathcal{B}(\mathcal{H})$, *se cumple que*

$$\sum_i p_i S(\rho_i)\leq S\left(\sum_i p_i\rho_i\right)\leq\sum_i p_i S(\rho_i)+H(p_i)$$

Demostración Sólo estamos resumiendo el teorema anterior y la concavidad de la entropía. **QED**

Ahora veamos algunos ejemplos propios de la Teoría de la información.

5.3 Desigualdad del procesamiento de datos

Iniciamos con una observación: por el Teorema 3.3, 5) la información común o compartida $H(X:Y) \equiv H(X) + H(Y) - H(X,Y)$, es *igual a cero* si, y sólo si, las variables aleatorias X, Y son independientes lo que refueza la idea de que $H(X:Y)$, mide que tanto X, Y conocen una acerca de la otra.

Cuando tenemos información disponible, antes de llegar a nosotros, ha sido sometida a "ruido". La desigualdad del procesamiento de datos establece que la información de salida de una fuente puede únicamente decrecer con el tiempo.

Para trabajar la noción del proceso de la información se usa el concepto de cadena de *Markov*, que fue definido al inicio de la sección 3.3; recordemos que una sucesión de variables aleatorias $X \longrightarrow Y \longrightarrow Z$ es una cadena de *Markov* si $P(Z \mid Y, X) = P(Z \mid Y)$.

Antes de enunciar el siguiente teorema, recordemos que se define la *entropía condicional de Shannon* de las variables aleatorias X_1, X_2 como $H(X_1 \mid X_2) = H(X_1, X_2) - H(X_2)$.

Teorema 7 *(Desigualdad del procesamiento de datos) Si $X \longrightarrow Y \longrightarrow Z$ es una cadena de Markov, entonces*

$$H(X) \geq H(X:Y) \geq H(X:Z),$$

y en la primera desigualdad hay igualdad si, y sólo si, dado Y es posible reconstruir X.

Una interpretación intuitiva de este resultado es que Z no puede conocer más sobre X que lo que Y conoce, lo que es menos que la información contenida en X.

Demostración La primera desigualdad y la condición de igualdad son consecuencias directas del Teorema 3.3, 2).
Para la otra desigualdad, como $X \longrightarrow Y \longrightarrow Z$ es una cadena de *Markov*, es inmediato comprobar que $Z \longrightarrow Y \longrightarrow X$ es también una cadena de *Markov*, luego usando esto es fácil verificar, a partir de la definición de la entropía de *Shannon,* que $H(X \mid Y) = H(X \mid Y, Z)$.

Pero por el mismo Teorema 3.3, 7), tenemos que $H(X \mid Y, Z) \leq H(X \mid Z)$ es decir $H(X \mid Y) \leq H(X \mid Z)$ y al usar la definición de $H(\cdot \mid \cdot)$ tenemos

$$H(X,Y) - H(Y) \leq H(X,Z) - H(Z),$$

y después de sumar $H(X)$ en ambos lados de la desigualdad anterior y de reagrupar, vemos que

$$H(Z) + H(X) - H(X,Z) \leq H(Y) + H(X) - H(X,Y),$$

y por definición concluimos que $H(X:Z) \leq H(X:Y)$. **QED**

Definición.
Una familia **finita** $\{E_b\} \subseteq \mathcal{B}(\mathcal{H})$ es una *medición con valores en los operadores positivos* ó **POVM** (por sus siglas en inglés: positive operator valued measurement) si para toda b,

$$E_b \geq 0 \quad \text{y} \quad \sum_b E_b = I.$$

La última igualdad se llama la *relación de completez.*

Dada una **POVM** fija $\mathcal{M}=\{E_b\} \subseteq \mathcal{B}(\mathcal{H})$ y una matriz de densidad γ, i.e. $\gamma \geq 0$ y $tr\gamma = 1$, queda definida una distribución de probabilidad clásica $p(b) = tr(\gamma E_b)$, cuya entropía de *Shannon* denotamos por

$$H(tr\gamma E_b).$$

Veamos que en efecto $\{tr\gamma E_b\}_b$ es una distribución de probabilidad.
Como γ es un operador no negativo sabemos, por la Proposición 1.5 que en su descomposición espectral $\gamma = \sum_n \lambda_n |u_n\rangle\langle u_n|$ se cumple que $\lambda_n \geq 0$ para cada n, luego
$tr\gamma E_b = tr\left(\sum_n \lambda_n |u_n\rangle\langle u_n| E_b\right) = \sum_n \lambda_n tr |u_n\rangle\langle E_b u_n| = \sum_n \lambda_n \langle E_b u_n, u_n\rangle \geq 0$
pues $E_b \geq 0$, es decir

$$tr\gamma E_b \geq 0 \quad \forall b.$$

Y por la relación de completez, $\sum_b tr(\gamma E_b) = tr\left(\sum_b \gamma E_b\right) = tr\left(\gamma \sum_b E_b\right) = tr\gamma = 1.$

5.4 La cota de Holevo

Lema 8 *Si $\{\lambda_1, ..., \lambda_n\}$ es una distribución de probabilidad y $\{\rho_1, ..., \rho_n\}$ es un conjunto de matrices de densidad estrictamente positivas en $\mathcal{H}$, entonces la matriz de densidad promedio $\rho = \sum_j \lambda_j \rho_j$ tiene la siguiente propiedad*

$$\sum_j \lambda_j S(\rho_j \parallel \rho) = S(\rho) - \sum_j \lambda_j S(\rho_j).$$

Observemos que debido a la concavidad de la entropía de *von Neumann* la cantidad del lado derecho de la igualdad es no negativa. Algunas veces se le llama la cantidad de *Holevo* y se denota con el símbolo χ, es decir

$$\chi \equiv S(\rho) - \sum_j \lambda_j S(\rho_j).$$

Demostración.

$\sum_j \lambda_j S(\rho_j \parallel \rho) = \sum_j \lambda_j tr\rho_j \log \rho_j - \sum_j \lambda_j tr\rho_j \log \rho$

$= -\sum_j \lambda_j S(\rho_j) - tr\rho \log \rho = -\sum_j \lambda_j S(\rho_j) + S(\rho)$. **QED**

Antes de proseguir tenemos las siguientes

Definiciones:

a) Dado un conjunto de matrices $\{\rho_1, ..., \rho_n\} \subseteq \mathcal{B}(\mathcal{H})$ y una distribución de probabilidad $\{\lambda_1, ..., \lambda_n\}$, su **ensemble** es el conjunto $\mathcal{E} = \{\lambda_j \rho_j\}_j$.

b) Sea $\{E_b\} \subseteq \mathcal{B}(\mathcal{H})$ una POVM. Si $\mathcal{E} = \{\lambda_j \rho_j\}_j$ es un ensemble de matrices donde cada $\rho_j \geq 0$ es matriz de densidad y si $\rho = \sum_j \lambda_j \rho_j$ es la matriz de densidad promedio, la **información accesible** en el ensemble se define como

$$I(\mathcal{E}, \mathcal{M}) \equiv H(tr\rho E_b) - \sum_j \lambda_j H(tr\rho_j E_b)$$

Una motivación física para el concepto de la *información accesible* es la siguiente:

Tenemos dos jugadores llamados Alice y Bob. Supongamos que Alice tiene información clásica, es decir una variable aleatoria A con valores $\{a\}$ y probabilidades λ_a. También cuenta con un conjunto finito fijo de estados (operadores de densidad) $\{\rho_a\}$ y elige un estado acorde al valor $a \in \{a\}$ que desea trasmitir, y se lo da a Bob quien tiene una POVM $\{E_b\}$ para medir el estado que recibe y obtener un resultado b que es un valor de una variable aleatoria B con valores $\{b\}$. El propósito del juego es que Bob determine el valor de A basado en los resultados B de su medición.

Hablemos con un poco más de precisión sobre la medición que hace *Bob*; él puede aplicar cada elemento E_b de la POVM a cualquier estado γ que Alice le mande y una de las reglas del juego es que la probabilidad de obtener el resultado b está dada por

$$P(b) = tr(\gamma E_b).$$

En particular

$$P(b) = tr(\rho E_b), \text{ donde } \rho = \sum_a \lambda_a \rho_a.$$

Así, cuando Bob hace su medición en el estado ρ_a, la probabilidad condicional de que resulte b es

$$P(b \mid a) = tr\rho_a E_b$$

y ésta induce una distribución conjunta sobre el mensaje de entrada a de Alice y el mensaje b decodificado por *Bob*,

$$\begin{aligned} P(a,b) &= \lambda_a P(b \mid a) \\ &= \lambda_a tr\rho_a E_b. \end{aligned}$$

Ya que existe una distribución de probabilidad conjunta entre los mensajes de entrada y de salida (las variables aleatorias A y B respectivamente) la información transferida puede ser analizada por la teoría de la información clásica: la información que Bob puede conocer acerca del mensaje de Alice, esto es la *información accesible* a Bob, está dada por la información común o compartida $H(A:B)$, entre A y los resultados de su medición B; como comentamos en el preámbulo del Teorema 7, ésta es una buena medida de cuanta información gana Bob sobre A, a través de su medición.
Debido a la desigualdad del procesamiento de datos (Teorema 7), sabemos que Bob puede inferir A de B si, y sólo si, $H(A:B) = H(A)$ y que en general $H(A:B) \leq H(A)$; así la meta de Bob es elegir la medición que maximice

$$H(A:B) \equiv H(A) + H(B) - H(A,B),$$

aproximando este valor tan cerca como le sea posible a $H(A)$.

Para determinar $H(A:B)$, necesitamos los cálculos siguientes:

$H(A,B) = \sum_{a,b} P(a,b) \log P(a,b) = \sum_{a,b} (\lambda_a tr\rho_a E_b) \log (\lambda_a tr\rho_a E_b) =$

$\sum_{a,b} (\lambda_a tr\rho_a E_b) \log \lambda_a + \sum_{a,b} (\lambda_a tr\rho_a E_b) \log (tr\rho_a E_b) =$

$\sum_a \lambda_a \log \lambda_a \sum_b tr\rho_a E_b + \sum_a \lambda_a \sum_b (tr\rho_a E_b) \log (tr\rho_a E_b) =$

$H(A)\, tr\rho_a + \sum_a \lambda_a H(tr\rho_a E_b) = H(A) + \sum_a \lambda_a H(tr\rho_a E_b)$, esto es

$$H(A,B) = H(A) + \sum_a \lambda_a H(tr\rho_a E_b),$$

de allí que

$$\begin{aligned} H(A:B) &= H(A) + H(B) - H(A,B) \\ &= H(B) - \sum_a \lambda_a H(tr\rho_a E_b), \end{aligned}$$

esto es

$$H(A:B) = H(tr(\rho E_b)) - \sum_a \lambda_a H(tr\rho_a E_b)$$

que es el término definido como la información accesible $I(\mathcal{E}, \mathcal{M})$.

Teorema 9 *(La cota de Holevo). Sean $\{\lambda_1, ..., \lambda_n\}$ una distribución de probabilidad y $\{\rho_1, ..., \rho_n\}$ un conjunto de matrices de densidad estrictamente positivas en $\mathcal{H}$. La cota de Holevo asegura que si, $\rho = \sum_j \lambda_j \rho_j$ es la matriz de densidad promedio, entonces la información accesible en el* ensemble *satisface*

$$I(\mathcal{E}, \mathcal{M}) \leq \chi,$$

*para cualquier **POVM**, $\mathcal{M} = \{E_b\}$.*

Queda como comentario que hay igualdad en la cota de *Holevo* si, y sólo si, $\rho_1, \ldots, \rho_n$ conmutan entre sí, ver [16].

Demostración. Sean $\mathcal{M}=\{E_b\} \subseteq \mathcal{B}(\mathcal{H})$ una **POVM** y $\beta = \{b\}$ una base ortonormal de $\mathcal{H}$.
Definimos el operador $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, como

$$\Phi(\varrho) = \sum_{b\in\beta} |b\rangle\langle b|\, tr(\varrho E_b), \quad \varrho \in \mathcal{B}(\mathcal{H}).$$

Veamos que Φ preserva la traza. En efecto, usando la misma base $\beta = \{b\}$ para calcular la traza tenemos:

$tr\Phi(\varrho) = \sum_{b'} \langle b', \sum_b |b\rangle\langle b|\, tr(\varrho E_b)\, b'\rangle = \sum_{b'} \sum_b \langle b', |b\rangle\langle b|\, tr(\varrho E_b)\, b'\rangle =$

$\sum_{b'} \sum_b \langle |b\rangle\langle b|\, b', tr(\varrho E_b)\, b'\rangle = \sum_{b'} \sum_b \langle b', b\rangle \langle b, tr(\varrho E_b)\, b'\rangle =$

$\sum_{b'} \langle b', tr(\varrho E_{b'})\, b'\rangle = \sum_{b'} tr(\varrho E_{b'}) \langle b', b'\rangle = \sum_{b'} tr(\varrho E_{b'}) =$

$tr\varrho \left(\sum_{b'} E_{b'}\right) = tr\varrho,$

en la última igualdad hemos usado la *relación de completez* $\sum_{b'\in\beta} E_{b'} = I$.

Ahora afirmamos que Φ es un operador completamente positivo; para comprobarlo veremos que tiene la siguiente representación de *Krauss*:

$$\Phi(\varrho) = \sum_{k\in\beta} \sum_{b\in\beta} |b\rangle\langle k|\, \sqrt{E_b}\varrho\sqrt{E_b}\, |k\rangle\langle b|, \quad \varrho \in \mathcal{B}(\mathcal{H}).$$

En efecto, sean b', b'' elementos de la base $\{b\}$, calculando tenemos,

$\langle b', \sum_b |b\rangle\langle b|\, tr(\varrho E_b)\, b''\rangle = \sum_b \langle b', |b\rangle\langle b|\, tr(\varrho E_b)\, b''\rangle$

$= \sum_b \langle |b\rangle\langle b|\, b', tr(\varrho E_b)\, b''\rangle = \sum_b \overline{\langle b, b'\rangle} \langle b, tr(\varrho E_b)\, b''\rangle$

$= \overline{\langle b', b'\rangle} \langle b', tr(\varrho E_{b'})\, b''\rangle = \langle b', tr(\varrho E_{b'})\, b''\rangle = \delta_{b'b''} tr(\varrho E_{b'}),$

esto es

$$\left\langle b', \sum_b |b\rangle\langle b|\, tr(\varrho E_b)\, b''\right\rangle = \delta_{b'b''} tr(\varrho E_{b'}). \tag{5.3}$$

Y por otra parte

$\left\langle b', \sum_{k,b} |b\rangle\langle k|\, \sqrt{E_b}\varrho\sqrt{E_b}\, |k\rangle\langle b|\, b''\right\rangle = \sum_k \left[\sum_b \left\langle b', \delta_{bb''}\, |b\rangle\langle k|\, \sqrt{E_b}\varrho\sqrt{E_b}k\right\rangle\right] =$

$\sum_k \left[\left\langle b', |b''\rangle\langle k|\, \sqrt{E_{b''}}\varrho\sqrt{E_{b''}}k\right\rangle\right] = \sum_k \left[\left\langle |k\rangle\langle b''|\, b', \sqrt{E_{b''}}\varrho\sqrt{E_{b''}}k\right\rangle\right] =$

$\sum_k \left[\overline{\langle b'', b' \rangle} \left\langle k, \sqrt{E_{b''}} \varrho \sqrt{E_{b''}} k \right\rangle \right] = \langle b', b'' \rangle \sum_k \left\langle k, \sqrt{E_{b''}} \varrho \sqrt{E_{b''}} k \right\rangle =$

$\delta_{b'b''} tr \left(\sqrt{E_{b''}} \varrho \sqrt{E_{b''}} \right) = \delta_{b'b''} tr \left(\varrho E_{b''} \right),$

la última igualdad es por la ciclicidad de la traza, o sea

$$\left\langle b', \sum_{k,b} |b\rangle\langle k| \sqrt{E_b} \varrho \sqrt{E_b} |k\rangle\langle b| \, b'' \right\rangle = \delta_{b'b''} tr \left(\varrho E_{b''} \right), \tag{5.4}$$

comparando (5.3) con (5.4) se comprueba lo afirmado.

Para continuar, aplicamos la monotonía de la entropía relativa bajo operadores completamente positivos (Teorema 4.10, proposición 1), y obtenemos que para cada j

$$S \left(\Phi \left(\rho_j \right) \parallel \Phi \left(\rho \right) \right) \leq S \left(\rho_j \parallel \rho \right)$$

ó equivalentemente, usando la definición de entropía relativa en el lado izquierdo

$$tr \left(\Phi \left(\rho_j \right) \log \Phi \left(\rho_j \right) \right) - tr \left(\Phi \left(\rho_j \right) \log \Phi \left(\rho \right) \right) \leq S \left(\rho_j \parallel \rho \right).$$

Pero por la definición de Φ tenemos que

$\log \Phi \left(\rho_j \right) = \sum_b |b\rangle\langle b| \log tr \left(\rho_j E_b \right), \log \Phi \left(\rho \right) = \sum_b |b\rangle\langle b| \log tr \left(\rho E_b \right)$, y por ende

$\Phi \left(\rho_j \right) \log \Phi \left(\rho_j \right) = \sum_b |b\rangle\langle b| \, tr \left(\rho_j E_b \right) \log tr \left(\rho_j E_b \right)$ y

$\Phi \left(\rho_j \right) \log \Phi \left(\rho \right) = \sum_b |b\rangle\langle b| \, tr \left(\rho_j E_b \right) \log tr \left(\rho E_b \right),$

por esto podemos reescribir la desigualdad anterior como

$$\sum_b tr \left(\rho_j E_b \right) \log tr \left(\rho_j E_b \right) - \sum_b tr \left(\rho_j E_b \right) \log tr \left(\rho E_b \right) \leq S \left(\rho_j \parallel \rho \right)$$

luego multiplicamos por λ_j y sumamos para obtener

$\sum_j \lambda_j \left[\sum_b tr \left(\rho_j E_b \right) \log \left(tr \rho_j E_b \right) \right] - \sum_b tr \left(\rho E_b \right) \log tr \left(\rho E_b \right) \leq$

$\sum_j \lambda_j S \left(\rho_j \parallel \rho \right) = \chi$, la igualdad es por el Lema 8. **QED**

La cota de Holevo es un resultado clave para demostrar varios resultados en la teoría de la información cuántica, ver [11]. Nosotros sólo lo aplicamos en el juego de Alice y Bob como sigue:
Por el Teorema 5 se infiere que

$$S \left(\sum_a \lambda_a \rho_a \right) - \sum_a \lambda_a S \left(\rho_a \right) \leq H \left(A \right),$$

con igualdad si, y sólo si, los estados ρ_a tienen soportes ortogonales. Así, cuando los operadores ρ_a de Alice no tienen soportes ortogonales la desigualdad anterior y el Teorema 9 (la cota de *Holevo*), implican que $H(A:B)$ es estrictamente menor que $H(A)$ y consecuentemente, es imposible para Bob determinar con certeza absoluta, el estado que Alice le está enviando basado en los resultados de su medición.

Terminamos con un ejemplo concreto del juego de Bob y Alice en el cual usaremos logaritmos de base 2 para calcular la entropía: Sean $\mathcal{H}$ un espacio de *Hilbert* de dimensión 2 con una base ortonormal $\{|0\rangle, |1\rangle\}$ y $\psi = \cos\theta\,|0\rangle + sen\ \theta\,|1\rangle$, donde θ es algún parámetro real.

Alice tiene los estados $\{|0\rangle\langle 0|, |\psi\rangle\langle\psi|\} \subseteq \mathcal{B}(\mathcal{H})$ y los valores 0 y 1 que elige lanzando una moneda honesta. Si la moneda cae cara, Alice prepara el estado $|0\rangle\langle 0|$ y cuando la moneda cae águila elije el estado

$$|\psi\rangle\langle\psi| = \cos^2\theta\,|0\rangle\langle 0| + \cos\theta sen\ \theta\,|0\rangle\langle 1| + \cos\theta sen\ \theta\,|1\rangle\langle 0| + sen^2\ \theta\,|1\rangle\langle 1|.$$

Usando la expresión anterior de $|\psi\rangle\langle\psi|$ es fácil ver, que respecto a la base $\{|0\rangle, |1\rangle\}$, el estado promedio $\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|\psi\rangle\langle\psi|$ tiene la expresión matricial

$$\rho = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \cos^2\theta & \cos\theta sen\ \theta \\ \cos\theta sen\ \theta & sen^2\ \theta \end{bmatrix},$$

y que sus valores propios son $\frac{1\pm\cos\theta}{2}$.

En este caso, la cota de *Holevo* es

$$\begin{aligned} \chi &= S(\rho) - \frac{1}{2}S(|0\rangle\langle 0|) - \frac{1}{2}S(|\psi\rangle\langle\psi|) \\ &= S(\rho) \\ &= H_{bin}\left(\frac{1+\cos\theta}{2}\right), \end{aligned}$$

pues $S(|0\rangle\langle 0|) = S(|\psi\rangle\langle\psi|) = 0$ por tener respectivamente los valores propios 1 y 0. Es decir, χ es la entropía binaria

$$H_{bin}\left(\tfrac{1+\cos\theta}{2}\right) =$$

$$-\tfrac{1}{\ln 2}\left(\tfrac{1+\cos\theta}{2}\right)\ln\left(\tfrac{1+\cos\theta}{2}\right) - \tfrac{1}{\ln 2}\left(\tfrac{1-\cos\theta}{2}\right)\ln\left(\tfrac{1-\cos\theta}{2}\right),$$

por lo cual la cota de *Holevo* es máxima cuando $\theta = \pi/2$, alcanzando el valor máximo de 1, que corresponde al caso cuando los estados preparados por Alice son ortogonales en cuyo caso Bob puede determinar con certeza los estados preparados por Alice. Para otros valores de θ, la cota de *Holevo* es menor estrictamente que $H(A) = 1$ y es imposible para Bob determinar con certeza cuál es el estado preparado por Alice.

5.5 Conclusiones

La teoría que presentamos es muy interesante y rica. Integra conceptos, ideas y métodos clásicos con nociones modernas. Como ejemplos de los primeros y de los segundos, respectivamente tenemos:

a) Convexidad, derivabililidad, integral de *Riemannn* y la prueba de *Epstein* del Teorema 4.6 pues, aunque no lo probamos aquí, no podemos dejar de mencionar que se demuestra usando métodos del análisis complejo.

b) El operador traza parcial, los operadores completamente positivos, la representación generalizada de *Lindblad,* así como la presentación del *proceso de purificación* para obtener desigualdades (la demostración del Teorema 4.14 es un ejemplo).

Dentro de los logros de este trabajo merece la pena resaltar que:

a) Desarrollamos las propiedades básicas del operador traza parcial para el caso de dimensión finita que se pueden generalizar al caso en que el espacio de *Hilbert* es de dimensión infinita pero separable.

b) Demostramos la desigualdad de *Klein* junto con una condición de igualdad.

c) Demostramos que el operador traza parcial es un operador CP que preserva la traza.

d) Demostramos el caso general, que para los operadores $\Phi : \mathcal{B}(\mathcal{H}_1) \longrightarrow \mathcal{B}(\mathcal{H}_2)$, ser CP y preservar la traza equivale a tener una representación de *Lindblad.*

e) Diseñamos una nueva ruta para demostrar la equivalencia entre las desigualdades SSA, MONO, MPT, CVX y CON que aparecen en el Teorema 4.14

f) Encontramos una nueva condición necesaria para que se cumpla la igualdad en MONO, (Corolario 4.11).

g) Presentamos una prueba simple de la cota de *Holevo.*

5.6 Perspectivas

Nuestro trabajo ha sido tan sólo un asomo a la interesante y de intenso desarrollo teoría de la información cuántica, por lo mismo las tareas por realizar son muchas, sin embargo, proponemos tan sólo algunas que consideramos importantes:

a) Extender los conceptos presentados e investigar cuáles resultados siguen siendo válidos en el caso en que los espacios de *Hilbert* subyacentes son de dimensión no finita pero separables.

b) Adaptar o buscar nuevas demostraciones de las desigualdades para el caso menos restrictivo en que los operadores de densidad sean positivos semi-definidos ($\rho \geq 0$).

c) Estudiar e investigar las propiedades de continuidad de la entropía y entropía relativa de *von Neumann*.

d) Explorar la conexión entre las desigualdades para las normas de la clase de traza L_p con el estudio de la entropía y la capacidad en información cuántica. Según *Mary Beth Ruskai* [16], dicha conexión promete ser una vía promisoria de estudio.

Apéndice A

La desigualdad de Jensen

Definición A$_1$ (*Conjunto convexo*) Un subconjunto $\mathcal{C}$ de un espacio vectorial real se llama *convexo* si para todas $a, b \in \mathcal{C}$ y $0 \leq \lambda \leq 1,\ \lambda \in \mathbb{R}$ se cumple que $\lambda a + (1 - \lambda)\, b \in \mathcal{C}$.

Se puede demostrar que la condición anterior equivale a pedir que para cualquier subconjunto finito de números reales $\{\lambda_i\}$ tales que $0 \leq \lambda_i \leq 1$ y $\sum_i \lambda_i = 1$ y cualquier subconjunto $\{a_i\}$ de elementos de $\mathcal{C}$ se cumple que $\sum_i \lambda_i a_i \in \mathcal{C}$, por ejemplo usando Inducción Matemática :
Si $a_1, ..., a_n, a_{n+1} \in \mathcal{C}$ y $0 \leq \lambda_i \leq 1$ con $\sum_{i=1}^{n+1} \lambda_i = 1$, entonces se deduce que $\lambda_1 a_1 + (1 - \lambda_1) \left[\sum_{i=2}^{n+1} \frac{\lambda_i}{1-\lambda_1} a_i\right] = \sum_{i=1}^{n+1} \lambda_i a_i \in \mathcal{C}$ suponiendo que $\sum_{i=2}^{n+1} \frac{\lambda_i}{1-\lambda_1} a_i \in \mathcal{C}$ y que $\lambda_1 < 1$.

Definición A$_2$ (*Función Convexa*). Sea $\mathcal{C}$ un conjunto convexo, una función $f : \mathcal{C} \to \mathbb{R}$ es *convexa*, si para todas $a, b \in \mathcal{C}$ y $0 \leq \lambda \leq 1$,

$$f\left(\lambda a + (1 - \lambda)\, b\right) \leq \lambda f(a) + (1 - \lambda)\, f(b)\,.$$

Si además se da la igualdad sólo cuando $\lambda = 0$ (ó $\lambda = 1$) ó $a = b$ se dice que f es estrictamente convexa.

Se puede demostrar que esta definición equivale a pedir que para cualquier subconjunto finito de números reales $\{\lambda_i\}$ tales que $0 \leq \lambda_i \leq 1$ y $\sum_i \lambda_i = 1$ y cualquier subconjunto $\{a_i\}$ de elementos de $\mathcal{C}$ se cumple que

$$f\left(\sum_i \lambda_i a_i\right) \leq \sum_i \lambda_i f(a_i)$$

y si además se da la igualdad sólo cuando $\lambda_i = 1$ para algún i ó cuando todos los a_i son iguales, entonces f es estrictamente convexa.

Definición A$_3$ *(Función cóncava)* Una función es cóncava si $-f$ es convexa.

A continuación discutimos el conocido criterio para decidir si una función dos veces derivable es convexa con base en el signo de la segunda derivada.

Proposición A$_4$ Sea $\mathcal{I} = (a, b)$ un intervalo abierto de $\mathbb{R}$ y sea $f : \mathcal{I} \to \mathbb{R}$ una función que tiene segunda derivada tal que

$$f'' > 0, \quad \text{en } \mathcal{I} \tag{A.1}$$

entonces f es estrictamente convexa en $\mathcal{I}$.

Demostración. Sean reales λ_i tales que $0 < \lambda_i < 1$, con $\sum_i \lambda_i = 1$ y $\{a_i\} \subset \mathcal{I}$, entonces $\overline{x} \equiv \sum_i \lambda_i a_i \in \mathcal{I}$.
Por el **Teorema de Taylor** para cada i, existe algún ξ_i entre $\overline{x}$ y a_i tal que

$$f(a_i) = f(\overline{x}) + (a_i - \overline{x}) f'(\overline{x}) + \frac{1}{2}(a_i - \overline{x})^2 f''(\xi_i),$$

multiplicando por λ_i y sumando obtenemos,

$$\sum_i \lambda_i f(a_i) = f(\overline{x}) + \sum_i \frac{1}{2}\lambda_i (a_i - \overline{x})^2 f''(\xi_i),$$

y por (A.1)

$$\sum_i \lambda_i f(a_i) - f(\overline{x}) = \sum_i \frac{1}{2}\lambda_i (a_i - \overline{x})^2 f''(\xi_i) \quad \geq 0$$

ó equivalentemente,

$$\begin{aligned} \sum_i \lambda_i f(a_i) &\geq f(\overline{x}) \\ &= f\left(\sum_i \lambda_i a_i\right); \end{aligned}$$

además si hay igualdad, entonces $\sum_i \frac{1}{2}\lambda_i (a_i - \overline{x})^2 f''(\xi_i) = 0$ y como por hipótesis $0 < \lambda_i < 1$, concluimos que $(a_i - \overline{x})^2 = 0$ para cada i y en consecuencia todos los a_i son iguales a $\overline{x}$. **QED**

Lema A_5 Sea $\mathcal{I}$ un intervalo abierto de $\mathbb{R}$ y sea $f : \mathcal{I} \longrightarrow \mathbb{R}$ una función diferenciable y convexa, entonces para todos $a, b \in \mathcal{I}$.

$$(a-b) f'(b) \leq f(a) - f(b)$$

Demostración.
Sea $h \in (0,1]$, por la convexidad tenemos que $f((1-h)b + ha) \leq (1-h) f(b) + hf(a)$ lo que implica

$$\frac{f((a-b)h+b) - f(b)}{h} \leq f(a) - f(b).$$

Definiendo la función $g : [0,1] \to \mathbb{R}$ con $g(x) = f((a-b)x+b)$, tenemos $g'(x) = f'((a-b)x+b)(a-b)$ y por ende

$$\begin{aligned} f'(b)(a-b) &= g'(0) \\ &= \lim_{h\to 0} \frac{g(h) - g(0)}{h} \\ &= \lim_{h\to 0} \frac{f((a-b)h+b) - f(b)}{h} \\ &\leq f(a) - f(b). \quad \textbf{QED} \end{aligned}$$

Definición A_6 (*Cono convexo*). Un subconjunto $\mathcal{C}$ de un espacio vectorial real se llama *cono convexo* si es un subconjunto convexo y $xa \in \mathcal{C}$, $\forall x > 0,\ x \in \mathbb{R}$

Observemos que un cono convexo tiene la propiedad de que $\forall\ a, b \in \mathcal{C}$ y $\forall\ x > 0$, $x \in \mathbb{R}$ se cumple que $a + xb \in \mathcal{C}$ porque podemos escribir $a + xb = (1+x)\left[(1+x)^{-1} a + x(1+x)^{-1} b\right]$.

Lema A_6 Sea $\mathcal{C}$ un cono convexo y $f : \mathcal{C} \longrightarrow \mathbb{R}$ una función cóncava y homogénea (i.e. $f(xa) = xf(a)$ para $x > 0$). Entonces para $x \in \mathbb{R}$ tal que $x > 0$,

$$\frac{f(a+xb) - f(a)}{x} \geq f(b)$$

y $\lim_{x\to 0^+} \frac{f(a+xb)-f(a)}{x} \geq f(b)$ cuando f es diferenciable a la derecha en a en el sentido de que $\lim\limits_{x\to 0^+} \frac{f(a+xb)-f(a)}{x}$ existe para todo par $a, b \in \mathcal{C}$.

Demostración. De la concavidad y homogeneidad se sigue que $\frac{1}{2} f(a+xb) = f\left(\frac{a}{2} + \frac{xb}{2}\right) \geq \frac{1}{2} f(a) + \frac{1}{2} f(xb)$ que equivale a $\frac{f(a+xb)-f(a)}{x} \geq f(b)$.
La segunda parte se sigue directamente de la desigualdad mostrada. **QED**

Apéndice B

La derivada de funciones con valores en $\mathcal{B}(\mathcal{H})$

A lo largo de este apéndice simbolizamos con E al espacio de *Banach* $\mathcal{B}(\mathcal{H})$ donde $\mathcal{H}$ es un espacio de *Hilbert* complejo.

Definición B$_1$ Sea $[a, b]$ un intervalo de números reales.

$f : [a, b] \to E$ es una función derivable en $t_0 \in (a, b)$, si existe un vector $x_{t_0} \in E$ tal que

$$\lim_{t \to t_0} \left\| \frac{f(t) - f(t_0)}{t - t_0} - x_{t_0} \right\| = 0,$$

cuando t_0 es alguno de los extremos a ó b pedimos que los límites laterales existan. Es fácil verificar que si existe x_{t_0}, es único y que en este caso la función es continua en $t = t_0$ y se escribe

$$x_{t_0} = f'(t_0) = \left. \frac{df}{dt} \right|_{t=t_0}.$$

Proposición B$_2$ Si $f : [a, b] \to E$ es una función derivable en t y $\Lambda \in \mathcal{B}(E, \mathbb{C})$, entonces

$$(\Lambda \circ f)'(t) = \Lambda \circ f'(t).$$

En este caso la función $\Lambda \circ f : [a, b] \to \mathbb{C}$ es derivable en t en el sentido usual.

Demostración. El resultado se sigue de la siguiente igualdad que se cumple por la linealidad de Λ,

$$\frac{\Lambda f(t) - \Lambda f(t_0)}{t - t_0} = \Lambda \frac{f(t) - f(t_0)}{t - t_0}. \quad \textbf{QED}$$

Si $F:[a,b]\to\mathcal{B}(\mathcal{H})$ es una función y $x,y\in\mathcal{H}$ definimos la función

$$F_{x,y}:[a,b]\to\mathbb{C},\ \ F_{x,y}(t)=\langle x,F(t)\,y\rangle\,,\quad t\in[a,b]$$

La siguiente proposición se refiere a la función recientemente definida.

Proposición B$_3$ Si $F:[a,b]\to\mathcal{B}(\mathcal{H})$ es una función derivable en $t\in[a,b]$ y $x,y\in\mathcal{H}$, entonces la función $F_{x,y}$ es derivable en t y

$$F'_{x,y}(t)=\langle x,F'(t)\,y\rangle$$

Demostración.
Observemos que para $s\neq t$

$$\frac{F_{x,y}(s)-F_{x,y}(t)}{s-t}-\langle x,F'(t)\,y\rangle=\frac{1}{s-t}\langle x,[F(s)-F(t)-(s-t)\,F'(t)]\,y\rangle\,,$$

de allí que

$$\left|\tfrac{F_{x,y}(s)-F_{x,y}(t)}{s-t}-\langle x,F'(t)\,y\rangle\right|\leq\tfrac{1}{|s-t|}\left\|F(s)-F(t)-(s-t)\,F'(t)\right\|\|x\|\,\|y\|$$

$=\left\|\frac{F(s)-F(t)}{s-t}-F'(t)\right\|\|x\|\,\|y\|$ de donde se sigue el resultado. **QED**

Apéndice C

Integral de Riemann

La integral de *Riemann* para funciones $F : [a,b] \to \mathcal{B}(\mathcal{H})$, donde $[a,b]$ es un intervalo de números reales se define de la manera usual.
Una partición $\mathcal{P}$ del intervalo $[a,b]$ de $\mathbb{R}$ es un conjunto de números reales de la forma $\mathcal{P} = \{a = t_0 < t_1 < t_2 <, ..., < t_{n-1} < t_n = b\}$; la norma de $\mathcal{P}$ se define como $\mu(\mathcal{P}) = \max\{t_j - t_{j-1} : 1 \leq j \leq n\}$. Una elección asociada a la partición $\mathcal{P}$ es un conjunto $\mathcal{A} = \{\xi_1, ..., \xi_n\}$ donde para cada j, $\xi_j \in [t_{j-1}, t_j]$.

Sea $F : [a,b] \to \mathcal{B}(\mathcal{H})$, si $\mathcal{A}$ es una elección asociada a la partición $\mathcal{P}$ se define la suma de *Riemann* de F como

$$\sum(F; \mathcal{P}, \mathcal{A}) = \sum_{j=1}^{n} (t_j - t_{j-1}) F(\xi_j)$$

Proposición C$_1$ Si $F : [a,b] \to \mathcal{B}(\mathcal{H})$ es una función continua, entonces existe un único operador J con la propiedad de que para cada $\epsilon > 0$, existe $\delta > 0$ tal que si $\mathcal{P}$ es una partición de $[a,b]$ con $\mu(\mathcal{P}) < \delta$, entonces $\left\|\sum(F; \mathcal{P}, \mathcal{A}) - J\right\| < \epsilon$ para cualquier elección $\mathcal{A}$ asociada a $\mathcal{P}$.

El operador J se llama la integral de F sobre $[a,b]$ y se denota por

$$J = \int_a^b F(t)\ dt$$

Demostración. Por ser similar al caso real y complejo sólo se indican los pasos esenciales. Como F es uniformemente continua en $[a,b]$, para cada $\epsilon > 0$, existe δ tal que si $\mu(\mathcal{P}), \mu(\mathcal{P}') < \delta$, entonces

$$\left\|\sum(F; \mathcal{P}, \mathcal{A}) - \sum(F; \mathcal{P}', \mathcal{A}')\right\| < \epsilon,$$

para cualesquier elecciones $\mathcal{A},\mathcal{A}'$asociadas a las particiones de $\mathcal{P}$ y $\mathcal{P}'$ respectivamente.

Después se toma una sucesión de particiones $(\mathcal{P}_n)_n$ de $[a,b]$ con $\mu(\mathcal{P}_n)\to 0$ cuando $n\to\infty$ y cualquier sucesión de aumentaciones $(\mathcal{A}_n)_n$, por último se verifica que la sucesión $\{\sum(F;\mathcal{P}_n,\mathcal{A}_n)\}_n$ converge en la norma de $\mathcal{B}(\mathcal{H})$ a un único límite J. **QED**

La integral tiene las siguientes propiedades.

Proposición C$_2$ Si $F,F_2:[a,b]\to\mathcal{B}(\mathcal{H})$ son funciones continuas y $\lambda\in\mathbb{C}$, entonces

a) *La integral es lineal*

$$\int_a^b [\lambda F(t)+F_2(t)]\,dt=\lambda\int_a^b F(t)\,dt+\int_a^b F_2(t)\,dt$$

b)

$$\left\|\int_a^b F(t)\,dt\right\|\leq\int_a^b \|F(t)\|\,dt$$

c) La función $G:[a,b]\to\mathcal{B}(\mathcal{H})$ tal que $G(t)=\int_a^t F(s)\,ds$ es derivable y $G'(t)=F(t)$ para todo $t\in[a,b]$.

d) Si $x,y\in\mathcal{H}$, entonces

$$\left\langle x,\left(\int_a^b F(t)\,dt\right)y\right\rangle=\int_a^b \langle x,F(t)\,y\rangle\,dt$$

Demostración.

a) Se omite.

b) Sea $J=\int_a^b F(t)\,dt$. Si $\epsilon>0$, existe $\delta>0$ tal que si $\mathcal{P}$ es una partición de $[a,b]$ con $\mu(\mathcal{P})<\delta$, entonces $\|\sum(F;\mathcal{P},\mathcal{A})-J\|<\epsilon$.

Ahora escogemos una partición $\mathcal{Q}$ con $\mu(\mathcal{Q})<\delta$ y una elección $\mathcal{A}$ asociada a $\mathcal{Q}$, luego

$$\|J\|\leq\left\|\sum(F;\mathcal{Q},\mathcal{A})\right\|+\left\|J-\sum(F;\mathcal{Q},\mathcal{A})\right\|. \tag{C.1}$$

Dado que $\left\|\sum\left(F;\mathcal{Q},\mathcal{A}\right)\right\| \leq \sum_{j=1}^{n}\left(t_j - t_{j-1}\right)\left\|F\left(\xi_j\right)\right\| \leq \int_a^b \left\|F\left(t\right)\right\| dt$, la desigualdad (C.1) se puede expresar como

$$\left\|J\right\| \leq \int_a^b \left\|F\left(t\right)\right\| dt + \epsilon$$

y de aquí se infiere el resultado.

c) Sea $t_o \in [a,b]$. Por la continuidad de F en t_0 dada $\epsilon > 0$ existe $\delta > 0$ tal que $\left\|F\left(s\right) - F\left(t_0\right)\right\| < \epsilon$, cuando $\left|s - t_0\right| < \delta$. Usando el inciso precedente, y si $0 < \left|t - t_0\right| < \delta$, tenemos las siguientes estimaciones:

$$\begin{aligned}
\left\|\left(\frac{G\left(t\right) - G\left(t_0\right)}{t - t_o} - F\left(t_0\right)\right)\right\| &= \frac{1}{\left|t - t_o\right|}\left\|\int_{t_o}^{t}\left[F\left(s\right) - F\left(t_0\right)\right] ds\right\| \\
&\leq \frac{1}{\left|t - t_o\right|}\int_{t_o}^{t}\left\|F\left(s\right) - F\left(t_0\right)\right\| ds \\
&\leq \frac{1}{\left|t - t_o\right|}\int_{t_o}^{t}\epsilon ds = \epsilon,
\end{aligned}$$

de donde se sigue el resultado.

Hagamos una pausa para recordar que si $f : [a,b] \to \mathbb{C}$ es una función con valores complejos tal que sus partes real e imaginaria $\Re\left(f\right), \Im\left(f\right) : [a,b] \to \mathbb{R}$ son funciones *Riemann* integrables se define

$$\int_a^b f\left(t\right) dt = \int_a^b \Re\left(f\right)\left(t\right) dt + i\int_a^b \Im\left(f\right)\left(t\right) dt.$$

Esta integral tiene propiedades semejantes a las de la integral de *Riemann* de funciones reales de variable real, una de ellas que nos interesa remarcar es:

El **Teorema fundamental del cálculo clásico**. Si $f : [a,b] \to \mathbb{C}$ es una función integrable y $h : [a,b] \to \mathbb{C}$ es una función derivable con $h' = f$ en $[a,b]$, entonces

$$\int_a^b f\left(t\right) dt = h\left(b\right) - h\left(a\right).$$

Se demuestra usando el caso real en las funciones parte real e imaginaria de f respectivamente.

d) Continuando, definimos la función $g : [a,b] \to \mathbb{C}$ con

$$g(t) = \left\langle x, \left(\int_a^t F(s)\, ds \right) y \right\rangle,$$

luego por la Proposición $\mathbf{B}_3$ y el inciso anterior tenemos $g'(t) = \langle x, F(t)\, y \rangle$, y por el Teorema fundamental del cálculo clásico concluimos que

$$\begin{aligned} \int_a^b \langle x, F(t)\, y \rangle\, dt &= g(b) - g(a) \\ &= \left\langle x, \left(\int_a^b F(s)\, ds \right) y \right\rangle \end{aligned}$$

Proposición $\mathbf{C}_3$ (*Teorema fundamental del cálculo*). Si $F : [a,b] \to \mathcal{B}(\mathcal{H})$ es una función continua y $G : [a,b] \to \mathcal{B}(\mathcal{H})$ es derivable con $G' = F$ en $[a,b]$, entonces

$$\int_a^b F(t)\, dt = G(b) - G(a)$$

Demostración. Para evidenciar la igualdad es suficiente verificar que para cualquier par $x, y \in \mathcal{H}$ se satisface que

$$\left\langle x, \left(\int_a^b F(t)\, dt \right) y \right\rangle = \langle x, [G(b) - G(a)]\, y \rangle;$$

con esto en mente, definimos la función $g : [a,b] \to \mathbb{C}$ mediante $g(t) = \langle x, G(t)\, y \rangle$; por la Proposición B_3 y por hipótesis g satisface que $g'(t) = \langle x, G'(t)\, y \rangle = \langle x, F(t)\, y \rangle$, después usando el inciso d) de la proposición C_2 y el Teorema fundamental del cálculo clásico para el caso complejo vemos que

$$\begin{aligned} \left\langle x, \left(\int_a^b F(t)\, dt \right) y \right\rangle &= \int_a^b \langle x, F(t)\, y \rangle\, dt \\ &= g(t)\big]_a^b = g(b) - g(a) \\ &= \langle x, G(b)\, y \rangle - \langle x, G(a)\, y \rangle. \quad \textbf{QED} \end{aligned}$$

Bibliografía

[1] R. Bhatia, *Matrix Analysis* (Springer-Verlag, 1997).

[2] O. Bratteli y D. Robinson, *Operator Algebras and Quantum Statistical Mechanics* Vol II (Springer-Verlag, 1981; segunda edición 1997).

[3] J. B. Conway, *A Course in Functional Analysis* (Springer-Verlag, 1990; cuarta impresión corregida 1997).

[4] H. Epstein, "Remarks on Two Theorems of E. Lieb" *Commun. Math. Phys.* **31**, 317-325 (1973).

[5] A. S. Holevo, "Information Theoretical aspects of Quantum Measurement" *Prob. Inf. Transmission* USSR **9**, 31-42 (1973).

[6] K. Krauss, *States, Effects and Operations: Fundamental Notions of Quantum Theory* (Springer-Verlag, 1983).

[7] O. Landford y D. Robinson, "Mean Entropy of States in Quantum Statistical Mechanics" *J. Math. Phys.* **9**, 1120-1125 (1968).

[8] E. Lieb, "Convex Trace Functions and the Wigner-Yanase-Dyson Conjecture" *Adv. Math.* **11**, 267-288 (1973).

[9] E. Lieb y M. B. Ruskai, "Proof of the Strong Subadditivity of Quantum Mechanical Entropy" *J. Math. Phys.* **14**, 1938-1941 (1973).

[10] G. Linblad, "Completely Positive Maps and Entropy Inequalities" *Commun. Math. Phys.* **40**, 147-151 (1975).

[11] M. A. Nielsen y I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

[12] H. A. Rincon, Álgebra Lineal, (Coordinación de Servicios Editoriales, Facultad de Ciencias, UNAM, 2001).

[13] R. T. Rockafellar, *Convex Analysis* (Princeton, 1972).

[14] W. Rudin, *Functional Analysis* (McGraw-Hill, 1973).

[15] D. Ruelle, *Statistical Mechanics- Rigorous Results*, (W. A. Benjamin, New York, 1969).

[16] M. B. Ruskai, "Inequalities for Quantum Entropy" *J. Math. Phys.* **43**, No 9, 4358-4375 (2002).

[17] C. Shannon, "A Mathematical Theory of Communication" *Bell Systems Tech Journal* (1948).

[18] V. Vedral, "The Role of Relative Entropy in Quantum Information Theory" *Rev. Mod. Phys.* **74** 197-233 (2002).

[19] J. von Neumann, "Thermodynamik Quantenmechanischer Gesamheiten" *Gött. Nach.* **1**, 273-291 (1927).

[20] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Berlin, 1932); Traducción al inglés por R. T. Beyer *Mathematical Foundations of Quantum Mechanics* (Princeton, 1955).

[21] A. Werhl, "General Properties of Entropy" *Rev. Mod. Phys.* **50**, 221-260 (1978).