**Espacios $L_p$ no-conmutativos y aplicaciones.**

Tesis que presenta:
Gildardo Barrientos Sánchez
para obtener el grado de
Maestro en Ciencias (Matemáticas)

Posgrado en Matemáticas
Departamento de Matemáticas
Universidad Autónoma Metropolitana
Iztapalapa

——————————————-

Asesor: Dr. Roberto Quezada Batalla.

26 de enero de 2010

# Índice general

# Reconocimientos

Agradezco principalmente a mi asesor, el Dr. Roberto Quezada, por la enorme paciencia, su apoyo durante mis estudios de maestría y sus invaluables enseñanzas. También un agradecimiento enorme a los sinodales, Dr. Luis Rincón y Dr. Fernando Galaz F., por la lectura del manuscrito en especial al Dr. Fernando por las observaciones hechas las cuales contribuyeron significativaente al trabajo.

Agradezco a la UAM-I las facilidades para llevar a cabo mis estudios de Maestría y para la realización de esta tesis. Agrego que mis estudios de Maestría fueron apoyados por el CONACYT mediante una beca-crédito y el proyecto de investigación número 49510-F.

Finalmente, gracias a ustedes: Tomasín (papá), Luchita (mamá), Edgardo B.S., José Luis Q. H., Cristhyan P. S. y Juan Pablo V. M., por estar cuando más se les necesita.

# Introducción

Los espacios $L_p$ no-conmutativos, también conocidos como ideales de Schatten, son una herramienta fundamental en Análisis Matemático y sus aplicaciones. Sobresalen sus usos recientes en Probabilidad Cuántica y Teoría Cuántica de la Información, véanse por ejemplo las referencias [4], [5], [7] y [17]. Sin embargo, hasta donde sabemos, no existe una monografía autocontenida que incluya una construcción elemental, directa y completa de estos espacios. La construcción que se encuentra en el libro de R. Bathia [1], que es la que desarrollamos en este trabajo, se encuentra dispersa en varios ejercicios y diferentes capítulos junto con otros resultados importantes sobre matrices, lo cual dificulta su lectura. El propósito principal de este trabajo es llenar este hueco, por lo menos en el caso de espacios de Hilbert de dimensión finita, es decir sobre matrices complejas $n \times n$. En el caso de dimensión infinita el lector puede consultar [9]. Para ilustrar su utilidad en Teoría Cuántica de la Información, desarrollamos en detalle la aplicación que realizan E. Lieb y E. Carlen en [4] y [5], donde obtienen algunas desigualdades tipo Minkowski que permiten dar una demostración corta de la subaditividad de la entropía de von Neumann, una desigualdad fundamental en Teoría Cuántica de la Información. Una de las consecuencias más importantes de la subaditividad fuerte es la célebre cota de Holevo para la información accesible, al respecto véanse por ejemplo las referencias [11], [16] y [13].

Este trabajo también es una breve introducción a la teoría de matrices estocásticas y transformaciones completamente positivas en dimensión finita. Nuestra aportación principal es la ordenación del material y la demostración detallada de todos los resultados que en [1], [4] y [5] aparecen sólo enunciados sin demostración o como ejercicios para el lector.

En el Capítulo 1 incluimos el concepto de mayorización y matrices estocásticas. El Capítulo 2 contiene una demostración de las propiedades de la norma en $L_p(\mathcal{H})$ donde $\mathcal{H}$ es un espacio de Hilbert de dimensión finita.

El material de estos dos capítulos es una introducción elemental y directa a los espacios $L_p$ no-conmutativos en dimensión finita. En el capítulo 3 desarrollamos en detalle la demostración de Carlen y Lieb de la subaditividad fuerte de la entropía de von Neumann usando una desigualdad tipo Minkowski. Incluimos un apéndice que contiene una introducción elemental de las transformaciones completamente positivas, la definición y propiedades elementales de la entropía de von Neumann y algunos resultados que se usan en los capítulos anteriores.

# Capítulo 1

# Mayorización y matrices biestocásticas

## 1.1. Mayorización

Sea $x = (x_1, ..., x_n) \in \mathbb{R}^n$, denotaremos por $x^\downarrow$ y $x^\uparrow$ a los vectores obtenidos a partir de $x$ reordenando las coordenadas en forma decreciente y en forma creciente respectivamente, es decir:

*(i)* $x^\downarrow = (x_1^\downarrow, ..., x_n^\downarrow)$ donde $x_1^\downarrow \geq x_2^\downarrow \geq ... \geq x_n^\downarrow$

*(ii)* $x^\uparrow = (x_1^\uparrow, ..., x_n^\uparrow)$ donde $x_1^\uparrow \leq x_2^\uparrow \leq ... \leq x_n^\uparrow$.

Obsérvese que

$$x_j^\uparrow = x_{n-j+1}^\downarrow \qquad j = 1, ..., n.$$

**Definición 1.1.1.** *Si $x,y \in \mathbb{R}^n$ decimos que $x$ es **mayorizado** por $y$ si se cumplen las condiciones:*

$$\sum_{j=1}^{k} x_j^\downarrow \leq \sum_{j=1}^{k} y_j^\downarrow, \quad \textit{para cada } 1 \leq k < n \ \textit{ y} \tag{1.1}$$

$$\sum_{j=1}^{n} x_j^\downarrow = \sum_{j=1}^{n} y_j^\downarrow. \tag{1.2}$$

*y en este caso escribiremos $x \prec y$. A la última igualdad le llamaremos condición de la traza, función que definiremos más adelante.*

**Ejemplo 1.1.2.** *Si* $x_j \geq 0$ *y* $\sum_{j=1}^{n} x_j = 1$, *entonces*

$$\left(\frac{1}{n}, ..., \frac{1}{n}\right) \prec (x_1, ...x_n) \prec (1, 0, ..., 0).$$

*Demostración.* La mayorización $x \prec (1, 0, ..., 0)$ es clara, pues la desigualdad

$$\sum_{j=1}^{k} x_j^{\downarrow} \leq 1$$

se cumple para cada $1 \leq k \leq n$ y también se tiene

$$\sum_{j=1}^{k} x_j = \sum_{j=1}^{k} x_j^{\downarrow} = 1.$$

Para demostrar la otra mayorización procederemos por inducción sobre $n$. Si $n = 1$ tenemos sólo un sumando y por tanto

$$\frac{1}{n} = 1 = x_1.$$

Por hipótesis de inducción, dados $x_1, ..., x_n$ tales que $\sum_{j=1}^{n} x_j = 1$ entonces para cada $1 \leq k \leq n$ se tiene que

$$\frac{k}{n} = \sum_{j=1}^{k} \left(\frac{1}{n}\right)^{\downarrow} \leq \sum_{j=1}^{k} x_j^{\downarrow}.$$

Sean $x_1, ...x_n, x_{n+1}$ tales que

$$\sum_{j=1}^{n+1} x_j = 1, \tag{1.3}$$

de esto tenemos que

$$x_{n+1} = 1 - \sum_{j=1}^{n} x_j. \tag{1.4}$$

Nótese que

$$x_{n+1}^{\downarrow} \leq \frac{1}{n+1},$$

pues si $x^{\downarrow}_{n+1} > \frac{1}{n+1}$ entonces como $x^{\downarrow}_1 \geq \cdots x^{\downarrow}_n \geq x^{\downarrow}_{n+1} > \frac{1}{n+1}$ tendríamos que

$$\sum_{j=1}^{n+1} x^{\downarrow}_j > \sum_{j=1}^{n+1} \frac{1}{n+1} = 1,$$

lo cual contradice (1.3). Sean $\widetilde{x}_j{}^{\downarrow} = \frac{x^{\downarrow}_j}{\sum_{j=1}^n x^{\downarrow}_j}$, $1 \leq j \leq n$. Entonces $\sum_{j=1}^n \tilde{x}^{\downarrow}_j = 1$ y por la hipótesis de inducción para cada $1 \leq k \leq n$, tenemos que $\frac{k}{n} \leq \sum_{j=1}^k \widetilde{x}_j{}^{\downarrow} = \sum_{j=1}^k \frac{x^{\downarrow}_j}{\sum_{j=1}^n x^{\downarrow}_j}$, equivalentemente $\frac{k}{n}\sum_{j=1}^n x^{\downarrow}_j \leq \sum_{j=1}^k x^{\downarrow}_j$ para cada $1 \leq k \leq n$. Entonces, usando (1.4), obtenemos

$$\frac{k}{n+1} = \frac{k}{n}\Big(1 - \frac{1}{n+1}\Big) \leq \frac{k}{n}\big(1 - x^{\downarrow}_{n+1}\big) = \frac{k}{n}\sum_{j=1}^n x^{\downarrow}_j \leq \sum_{j=1}^k x^{\downarrow}_j,$$

para cada $1 \leq k \leq n+1$. Esto termina la demostración. □

La noción de mayorización ocurre naturalmente en varios contextos, por ejemplo, en Física la relación $x \prec y$ indica que $x$ es un estado más caótico que $y$, pensando a $x$ como un proceso estocástico, esto es, cada $x_i$ es la probabilidad de encontrarse en el estado $i$.

Otro ejemplo ocurre en Economía, si $x_1, ..., x_n$ , $y_1, ..., y_n$ denotan los ingresos de $n-$individuos, entonces $x \prec y$ significa que hay una distribución más equitativa en el estado $x$ que en el estado $y$. El ejemplo anterior ilustra este hecho.

Por otra parte, recordemos que $x^{\uparrow}_j = x^{\downarrow}_{n-j+1}$, entonces

$$\sum_{j=1}^n x_j = \sum_{j=1}^n x^{\uparrow}_j = \sum_{j=1}^k x^{\uparrow}_j + \sum_{j=k+1}^n x^{\downarrow}_{n-j+1} = \sum_{j=1}^k x^{\uparrow}_j + \sum_{j=1}^{n-k} x^{\downarrow}_j.$$

De aquí se sigue la relación

$$\sum_{j=1}^k x^{\uparrow}_j = \sum_{j=1}^n x_j - \sum_{j=1}^{n-k} x^{\downarrow}_j, \quad 1 \leq k \leq n, \tag{1.5}$$

que usaremos en la demostración de la siguiente

**Proposición 1.1.3.** *$x \prec y$ si y sólo si*

$$\sum_{j=1}^{k} x_j^{\uparrow} \geq \sum_{j=1}^{k} y_j^{\uparrow} \quad 1 \leq k \leq n, \qquad \sum_{j=1}^{n} x_j^{\uparrow} = \sum_{j=1}^{n} y_j^{\uparrow}.$$

*Demostración.* Veamos que la condición es suficiente. Nótese que

$$\sum_{j=1}^{n} x_j^{\uparrow} = \sum_{j=1}^{n} x_j^{\downarrow} = \sum_{j=1}^{n} x_j.$$

Esta relación indica que la condición de la traza se cumple.

Usando (1.5) y la condición de la traza tenemos que para cada $1 \leq k < n$,

$$\sum_{j=1}^{k} x_j^{\downarrow} = \sum_{j=1}^{n} x_j - \sum_{j=1}^{n-k} x_j^{\uparrow} \leq \sum_{j=1}^{n} y_j - \sum_{j=1}^{n-k} y_j^{\uparrow} = \sum_{j=1}^{k} y_j^{\downarrow}.$$

Esto demuestra que $x \prec y$.

El recíproco se demuestra de manera análoga. □

Denotaremos por $\mathbf{e}$ al vector $(1, 1, ..., 1)$. Para cualquier subconjunto $I$ de $\{1, 2, ..., n\}$, $\mathbf{e}_I$ denotará el vector correspondiente a la función indicadora de $I$ y $|I|$ su cardinalidad.

Si $x \in \mathbb{R}^n$ definimos la traza de $x$ como

$$\mathrm{Tr}\,(x) := \sum_{j=i}^{n} x_j = \langle x, \mathbf{e} \rangle,$$

donde $\langle \cdot, \cdot \rangle$ es el producto interior en $\mathbb{R}^n$.

**Proposición 1.1.4.** *Sea $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$, para cada $1 \leq k \leq n$, se tiene que*

$$\sum_{j=1}^{k} x_j^{\downarrow} = \max_{|I|=k} \langle x, \mathbf{e}_I \rangle.$$

*Demostración.* Procederemos por inducción sobre $n$. Si $n = 1$ claramente tenemos

$$x_1^{\downarrow} = \max_{|I|=1} \langle x, \mathbf{e}_I \rangle.$$

Supongamos que se cumple para $n$, esto es,

$$\sum_{j=1}^{k} x_j^{\downarrow} = \max_{|I|=k} \langle x, \mathbf{e}_I \rangle, \ \ \forall \ 1 \leq k \leq n. \tag{1.6}$$

Demostraremos que (1.6) vale para cada $1 \leq k \leq n+1$. Sea $1 \leq k \leq n+1$, usando la hipótesis de inducción para $I \subset \{1, \cdots, k-1\}$ y $L \subset \{k, \cdots, n\}$ tenemos que

$$\begin{aligned}\sum_{j=1}^{k} x_j^{\downarrow} = \sum_{j=1}^{k-1} x_j^{\downarrow} + x_k^{\downarrow} &= \max_{|I|=k-1} \langle x, \mathbf{e}_I \rangle + x_k^{\downarrow} = \max_{|I|=k-1} \langle x, \mathbf{e}_I \rangle + \max_{|L|=1} \langle x, \mathbf{e}_L \rangle \\ &= \max_{|J|=k} \langle x, \mathbf{e}_J \rangle, \quad J \subset \{1, \cdots, n+1\}. \qquad \square\end{aligned}$$

Ahora tenemos la siguiente

**Proposición 1.1.5.** *$x \prec y$ si y sólo si para cada subconjunto $I$ de $\{1, 2, ..., n\}$ existe $J \subset \{1, 2, \ldots, n\}$ tal que $|I| = |J|$,*

$$\langle x, \mathbf{e}_I \rangle \leq \langle y, \mathbf{e}_J \rangle \qquad y \qquad \mathrm{Tr}\,(x) = \mathrm{Tr}\,(y).$$

*Demostración.* Supongamos que para cada $I \subset N = \{1, 2, ..., n\}$ existe $J \subset N$ tal que $|J| = |I|$,

$$\langle x, \mathbf{e}_I \rangle \leq \langle y, \mathbf{e}_J \rangle \qquad y \qquad \mathrm{Tr}\,(x) = \mathrm{Tr}\,(y).$$

Tomando máximo tenemos que

$$\langle x, \mathbf{e}_I \rangle \leq \max_{|J|=k} \langle y, \mathbf{e}_J \rangle,$$

es decir,

$$\langle x, \mathbf{e}_I \rangle \leq \sum_{j=1}^{k} y_j^{\downarrow}$$

y también se cumple

$$\max_{|I|=k} \langle x, \mathbf{e}_I \rangle \leq \sum_{j=1}^{k} y_j^{\downarrow}.$$

Entonces

$$\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow}.$$

Esto demuestra que $x \prec y$. El recíproco se demuestra de manera análoga. $\square$

**Definición 1.1.6.** *Diremos que un vector $x \in \mathbb{R}^n$ es* ***submayorizado débilmente*** *por un vector $y$ si*

$$\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow} \qquad k = 1, 2, ..., n,$$

*en este caso escribiremos $x \prec_\omega y$.*

*Diremos que un vector $x \in \mathbb{R}^n$ es* ***supermayorizado débilmente*** *por un vector $y$ si*

$$\sum_{j=1}^{k} x_j^{\uparrow} \geq \sum_{j=1}^{k} y_j^{\uparrow} \qquad k = 1, 2, ..., n,$$

*en este caso escribiremos $x \prec^\omega y$.*

**Proposición 1.1.7.** *Algunas propiedades de la mayorización.*

*(i)* $x \prec y$ *si y sólo si* $x \prec_\omega y$, $x \prec^\omega y$.

*(ii)* *Si $\alpha$ es un número no negativo , entonces* $x \prec_\omega y \Rightarrow \alpha x \prec_\omega \alpha y$; $x \prec^\omega y \Rightarrow \alpha x \prec^\omega \alpha y$.

*(iii)* $x \prec_\omega y$ *si y sólo si* $-x \prec^\omega -y$.

*(iv)* *Para cualquier* $\alpha \in \mathbb{R}$, $x \prec y$ *entonces* $\alpha x \prec \alpha y$.

*Demostración.*

*(i)* Si $x \prec y$, la relación $x \prec_\omega y$ se cumple por la definición 1.1.1, sin la condición de la traza. La relación $x \prec^\omega y$ se cumple por la Proposición 1.1.3.

Ahora supóngase que las relaciones $x \prec^\omega y$ y $x \prec_\omega y$ se cumplen. Bastará demostrar que la condición sobre la traza se cumple. Tomando $k = n$, por la submayorización, tenemos que

$$\operatorname{Tr}(x) = \sum_{j=1}^{n} x_j^{\downarrow} \leq \sum_{j=1}^{n} y_j^{\downarrow} = \operatorname{Tr}(y)$$

por otra parte; de la supermayorización, se tiene que

$$\operatorname{Tr}(x) = \sum_{j=1}^{n} x_j^{\uparrow} \geq \sum_{j=1}^{n} y_j^{\uparrow} = \operatorname{Tr}(y).$$

Esto demuestra que $x \prec y$.

*(ii)* Si $x \prec y$ y $\alpha \geq 0$ tenemos que

$$\sum_{j=1}^{n} (\alpha x_j)^{\downarrow} = \alpha\big(\sum_{j=1}^{n} x_j^{\downarrow}\big) \leq \alpha\big(\sum_{j=1}^{n} y_j^{\downarrow}\big) = \sum_{j=1}^{n} (\alpha y_j)^{\downarrow},$$

de donde se obtiene que $\alpha x \prec_{\omega} \alpha y$. La demostración de la otra implicación es similar.

*(iii)* Recordemos que para $x^{\downarrow}$ tenemos

$$x_1^{\downarrow} \geq x_2^{\downarrow} \geq ... \geq x_n^{\downarrow}$$

equivalentemente,

$$-x_1^{\downarrow} \leq -x_2^{\downarrow} \leq \cdots \leq -x_n^{\downarrow}.$$

Por otra parte para $(-x)^{\uparrow}$ tenemos que

$$(-x)_1^{\uparrow} \leq (-x)_2^{\uparrow} \leq \cdots \leq (-x)_n^{\uparrow},$$

entonces $(-x)_j^{\uparrow} = -x_j^{\downarrow}$ para cada $1 \leq j \leq n$ y por lo tanto $(-x)^{\uparrow} = -x^{\downarrow}$. La desigualdad

$$\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow}$$

se cumple si y sólo si

$$-\sum_{j=1}^{k} x_j^{\downarrow} \geq -\sum_{j=1}^{k} y_j^{\downarrow},$$

equivalentemente $\sum_{j=1}^{k}(-x)_j^{\uparrow} \geq \sum_{j=1}^{k}(-y)_j^{\uparrow}$, o bien $-x \prec^{\omega} -y$.

*(iv)* La condición de las trazas es evidente. Si $\alpha \geq 0$ la afirmación se sigue de $(ii)$ y de la identidad $\mathrm{Tr}\,(\alpha x) = \alpha \mathrm{Tr}\,(x)$, que se cumple para cada $\alpha \in \mathbb{R}$. En el caso $\alpha < 0$ usamos $(ii)$ y $(iii)$ para obtener $\alpha x \prec_{\omega} \alpha y$ y $\alpha x \prec^{\omega} \alpha y$. Entonces el resultado es consecuencia de la equivalencia en $(i)$. □

Cada una de las relaciones $\prec, \prec_{\omega}, \prec^{\omega}$ es reflexiva y transitiva, pero ninguna define un orden parcial. Por ejemplo, si $x \prec y$ y $y \prec x$, sólo podemos decir que $Py = x$, donde $P$ es la matriz asociada a una permutación $\sigma \in S_n$, este último es el grupo de permutaciones de un conjunto de $n$ elementos. Tal matriz está definida por

$$P_{\sigma} = \delta_{\sigma(i)j} = (p_{ij}) = \begin{cases} 1 & \text{si } j = \sigma(i) \\ 0 & \text{si } j \neq \sigma(i) \end{cases}$$

Denotaremos también $P \in S_n$. Tenemos una relación de equivalencia si definimos $x \sim y$ cuando $x = P_\sigma y$, para alguna permutacion $\sigma \in S_n$. La reflexividad se obtiene tomando $P = I$; además, si $x \sim y$ y $x = Py$ para una matriz de permutación, como $P^{-1}$ es también una matriz de permutación podemos escribir $x = P^{-1}y$ y se obtiene la simetría $y \sim x$. Por otra parte, como el producto de permutaciones es una permutación, si se cumple que $x \sim y$ y $y \sim z$, escribiendo $x = Py = P(Qz)$ se tiene que $x \sim z$, lo cual demuestra la transitividad. Si denotamos por $\mathbb{R}^n_{sym} = \mathbb{R}^n / \sim$, entonces claramente $\prec$ define un orden parcial en $\mathbb{R}^n_{sym}$. Esta relación también es un orden parcial sobre el conjunto $\{x \in \mathbb{R}^n : x_1 \geq \cdots \geq x_n\}$. La misma afirmación se cumple para las relaciones $\prec_\omega$ y $\prec^\omega$. Para $x, y \in \mathbb{R}^n$ definimos

$$\begin{aligned} x \vee y &= (x_1 \vee y_1, x_2 \vee y_2, ..., x_n \vee y_n) \\ x \wedge y &= (x_1 \wedge y_1, x_2 \wedge y_2, ..., x_n \wedge y_n) \end{aligned}$$

donde $x_j \vee y_j = \text{máx}(x_j, y_j)$ y $x_j \wedge y_j = \text{mín}(x_j, y_j)$. Por otra parte, sean

$$x^+ = x \vee 0, \qquad |x| = x \wedge (-x), \qquad x^- = -x \vee 0.$$

Nótese que $x^+$ es el vector obtenido al reemplazar las coordenadas negativas de $x$ por ceros, $|x| = (|x_1|, \ldots, |x_n|)$, es decir es el resultado de tomar el valor absoluto en cada coordenada del vector y $x^-$ toma los valores negativos del vector $x$.

Con las definiciones anteriores tenemos el siguiente resultado, que caracteriza la mayorización sin involucrar el reordenamiento de las coordenadas.

**Teorema 1.1.8.** *Sean $x, y \in \mathbb{R}^n$, entonces se cumplen*

*(i) $x \prec_\omega y$ si y sólo si para cualquier $t \in \mathbb{R}$*

$$\sum_{j=1}^{n} (x_j - t)^+ \leq \sum_{j=1}^{n} (y_j - t)^+.$$

*(ii) $x \prec^\omega y$ si y sólo si para cualquier $t \in \mathbb{R}$*

$$\sum_{j=1}^{n} (t - x_j)^+ \leq \sum_{j=1}^{n} (t - y_j)^+.$$

*(iii)* $x \prec y$ *si y sólo si para cualquier* $t \in \mathbb{R}$

$$\sum_{j=1}^{n} |x_j - t| \leq \sum_{j=1}^{n} |y_j - t|.$$

*Demostración.*

*(i)* Supóngase $x \prec_\omega y$ y tómese $t > x_1^\downarrow$, entonces tenemos que $(x_j - t)^+ = 0$ $\forall j$ y claramente se cumple que

$$\sum_{j=1}^{n}(x_j - t)^+ \leq \sum_{j=1}^{n}(y_j - t)^+.$$

Supongamos ahora que $x_{k+1}^\downarrow \leq t \leq x_k^\downarrow$ para algún $1 \leq k \leq n$ y por conveniencia escribamos $x_{n+1}^\downarrow = -\infty$, entonces

$$\begin{aligned}\sum_{j=1}^{n} (x_j - t)^+ &= \sum_{j=1}^{k} (x_j^\downarrow - t) = \sum_{j=1}^{k} x_j^\downarrow - tk \\ &\leq \sum_{j=1}^{k} y_j^\downarrow - tk \leq \sum_{j=1}^{k} (y_j^\downarrow - t)^+ \\ &\leq \sum_{j=1}^{n} (y_j - t)^+.\end{aligned}$$

Recíprocamente; supóngase que

$$\sum_{j=1}^{n} (x_j - t)^+ \leq \sum_{j=1}^{n} (y_j - t)^+ \qquad \forall t \in \mathbb{R},$$

Si $t = y_k^\downarrow$ entonces $y_1^\downarrow \geq y_2^\downarrow \geq ... \geq y_{k-1}^\downarrow \geq t$, y tenemos que

$$\sum_{j=1}^{n} (y_j - t)^+ = \sum_{j=1}^{k} (y_j^\downarrow - t) = \sum_{j=1}^{k} y_j^\downarrow - kt.$$

Por otro lado

$$\begin{aligned}\sum_{j=1}^{k} x_j^\downarrow - kt &= \sum_{j=1}^{k}(x_j^\downarrow - t) \leq \sum_{j=1}^{n}(x_j^\downarrow - t)^+ = \sum_{j=1}^{n}(x_j - t)^+ \\ &\leq \sum_{j=1}^{n}(y_j - t)^+ = \sum_{j=1}^{n} y_j^\downarrow - kt.\end{aligned}$$

Por lo tanto se debe tener que $x \prec_\omega y$. Esto demuestra $(i)$.

*(ii)* Por el inciso $(iii)$ de la Proposición 1.1.7 sabemos que $x \prec^\omega y \Leftrightarrow -x \prec_\omega -y$. El resultado se sigue al tomar $-t \in \mathbb{R}$ y aplicar $(i)$.

*(iii)* Basta recordar que $|x| = (x^+ + x^-)$, aplicar el inciso $(i)$ de la Proposición 1.1.7 y los incisos $(i)$ y $(ii)$ de esta proposición. $\square$

**Corolario 1.1.9.** *Si $x \prec y$ en $\mathbb{R}^n$ y $u \prec w$ en $\mathbb{R}^m$, entonces $(x,u) \prec (y,w)$ en $\mathbb{R}^{m+n}$. Además $x \prec y$ si y sólo si $(x,u) \prec (y,u)$, para cualquier $u$.*

*Demostración.* Supongamos que $x \prec y$ en $\mathbb{R}^n$ y $u \prec w$ en $\mathbb{R}^m$. Sean $\tilde{x} = (x,u)$ y $\tilde{y} = (y,w)$. Para $t \in \mathbb{R}$

$$\begin{aligned}\sum_{j=1}^{n+m} |\tilde{x}_j - t| &= \sum_{j=1}^{n} |x_j - t| + \sum_{i=1}^{m} |u_i - t| \\ &\le \sum_{j=1}^{n} |y_j - t| + \sum_{i=1}^{m} |w_i - t| = \sum_{j=1}^{n+m} |\tilde{y}_j - t|,\end{aligned}$$

por $(iii)$ de la proposición anterior tenemos que $(x,u) \prec (y,w)$.

Ahora, si $x \prec y$ por lo mostrado anteriormente se sigue que $(x,u) \prec (y,u)$ para toda $u$. Supongamos que $(x,u) \prec (y,u)$. Por $(iii)$ de la proposición anterior

$$\sum_{j=1}^{n} |x_j - t| + \sum_{j=n+1}^{n+m} |u_j - t| \le \sum_{j=1}^{n} |y_j - t| + \sum_{j=n+1}^{n+m} |u_j - t|.$$

Luego

$$\sum_{j=1}^{n} |x_j - t| \le \sum_{j=1}^{n} |y_j - t|,$$

se sigue así la mayorización $x \prec y$. $\square$

## 1.2. Matrices biestocásticas

En la presente sección definiremos el concepto de matriz biestocástica y demostraremos algunos resultados que relacionan a estas matrices con la mayorización. Al conjunto de matrices cuadradas de tamaño $n \times n$ lo denotaremos por $\mathbf{M}_n$.

**Definición 1.2.1.** *Una matriz $A \in \mathbf{M}_n$, es **biestocástica** si cumple*

$$a_{ij} \geq 0 \qquad \forall i, j = 1, 2, \ldots, n \tag{1.7}$$

$$\sum_{i=1}^{n} a_{ij} = 1, \qquad \forall j = 1, 2, \ldots, n \tag{1.8}$$

$$\sum_{j=1}^{n} a_{ij} = 1, \qquad \forall i = 1, 2, \ldots, n. \tag{1.9}$$

*Y decimos que $A \in \mathbf{M}_n$ es **bisubestocástica** si $a_{ij} \geq 0$ para toda $i, j$, $\sum_i a_{ij} \leq 1$ para toda $j$ y $\sum_j a_{ij} \leq 1$ para toda $i$.*

**Definición 1.2.2.** *Para $A \in \mathbf{M}_n$. Si $\sigma \in S_n$, el conjunto*

$$\{a_{1\sigma(1)}, a_{2\sigma(2)} \ldots, a_{n\sigma(n)}\},$$

*se le llama **diagonal** de $A$.*

**Definición 1.2.3.** *Una transformación lineal $A : \mathbb{C}^n \to \mathbb{C}^n$ **preserva positividad** si para vectores con coordenadas no negativas la imagen correspondiente es un vector con coordenadas no negativas.*

**Definición 1.2.4.** *Si $A \in \mathbf{M}_n$ decimos que **preserva traza** si $\operatorname{Tr}(Ax) = \operatorname{Tr}(x)$ para cualquier $x \in \mathbb{C}^n$.*

**Definición 1.2.5.** *Para $A \in \mathbf{M}_n$, decimos que es **unital** (o que preserva unidad) si $A\boldsymbol{e} = \boldsymbol{e}$; donde $\boldsymbol{e} = (1, 1, ..., 1)$.*

**Proposición 1.2.6.** *Una matriz $A \in \mathbf{M}_n$, es biestocástica si y sólo si preserva positividad, preserva traza y es unital.*

*Demostración.* Supongamos que $A \in \mathbf{M}_n$ es matriz biestocástica.

$(i)$ Sea $x \in \mathbb{C}^n$ tal que $x_j \geq 0$, para toda $j = 1, \ldots, n$. Tenemos

$$Ax = \left( \sum_{j=1}^{n} a_{1j} x_j, ..., \sum_{j=1}^{n} a_{nj} x_j \right) = (y_1, ..., y_n).$$

Como $x_j \geq 0$ y $a_{ij} \geq 0$, para toda $i, j$ se sigue que $y_i = \sum_{j=1}^{n} a_{ij} x_j \geq 0$ para toda $i = 1, ..., n$ por tanto, $A$ preserva positividad.

$(ii)$ Calculamos:

$$\operatorname{Tr}(Ax) = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}x_j\right) = \sum_{j=1}^{n} x_j \underbrace{\left(\sum_{i=1}^{n} a_{ij}\right)}_{=1} = \sum_{j=1}^{n} x_j = \operatorname{Tr}(x),$$

tenemos así que se cumple $\operatorname{Tr}(Ax) = \operatorname{Tr}(x)$, es decir $A$ preserva traza.

$(iii)$ Finalmente,

$$A\mathbf{e} = A(1,\ldots,1) = \left(\sum_{j=1}^{n} a_{1j},\ldots,\sum_{j=1}^{n} a_{nj}\right) = (1,1,...,1) = \mathbf{e}.$$

Es decir, A es unital.

Recíprocamente:

$(i)$ Por la Definición 1.2.3, tomando $e_j = (0,\cdots,\underbrace{1}_{j},\cdots 0)$ tenemos que el vector $Ae_j = (a_{1j}, a_{2j},\cdots,a_{nj})$ tiene coordenadas no negativas para cada $j$. Entonces $a_{ij} \geq 0$ para cada $1 \leq i,j \leq n$.

$(ii)$ Por la Definición 1.2.4, para cada vector $x \in \mathbb{C}^n$ se tiene la identidad

$$\operatorname{Tr}(Ax) = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}x_j\right) = \sum_{j=1}^{n} x_j = \operatorname{Tr}(x),$$

tomando $x = e_j = (0,\cdots,\underbrace{1}_{j},\cdots 0)$ se tiene $\sum_{i=1}^{n} a_{ij} = 1$, para toda $j = 1,2,\ldots,n$.

$(iii)$ Por la Definición 1.2.5 tenemos que

$$\left(\sum_{j=1}^{n} a_{1j},\ldots,\sum_{j=1}^{n} a_{nj}\right) = (1,1,...,1)$$

de donde se obtiene

$$\sum_{j=1}^{n} a_{ij} = 1$$

para toda $i = 1,\ldots,n$. Esto demuestra que la matriz es biestocástica. $\square$

**Proposición 1.2.7.** *(i) Las matrices biestocásticas $n \times n$ forman un conjunto convexo y cerrado bajo la multiplicación de matrices y la toma de adjuntos.*

*(ii) Toda matriz asociada a una permutación $\sigma \in S_n$ es biestocástica y es un punto extremo del conjunto de matrices biestocásticas.*

*Demostración.* $(i)$ Sean $A, B$ matrices biestocásticas y $t \in [0,1]$ , mostraremos que la matriz

$$Q = (1-t)A + tB$$

es biestocástica. Para $x \in \mathbb{C}^n$ tal que $x_j \geq 0$ para toda $j = 1, ..., n$ se tiene

$$Qx = (1-t)Ax + tBx.$$

Como $A, B$ preservan positividad, se sigue que $Q$ también preserva la positividad.

Usando las propiedades de la traza se obtiene que

$$\operatorname{Tr}(Qx) \;=\; (1-t)\operatorname{Tr} Ax + t \operatorname{Tr} Bx = (1-t)\operatorname{Tr} x + t \operatorname{Tr} x = \operatorname{Tr} x.$$

Por tanto, $Q$ preserva la traza.

Aplicando $Q$ a $\mathbf{e} = (1, 1, ..., 1)$ se obtiene que

$$Q\mathbf{e} = (1-t)A\mathbf{e} + tB\mathbf{e} = \mathbf{e}.$$

Por la Proposicion 1.2.6 obtenemos que $Q$ es una matriz biestocástica.

Veamos ahora que el conjunto de matrices biestocásticas es cerrado bajo el producto y toma de adjuntos.

Sean $A, B$ matrices biestocásticas y sea $P = AB$, tenemos que

$$\operatorname{Tr} Px = \operatorname{Tr} ABx = \operatorname{Tr} Bx = \operatorname{Tr} x,$$

pues $A, B$ preservan traza.

Sea $x \in \mathbb{C}^n$ con $x_j \geq 0,\ 1 \leq j \leq n$. Como $Bx$ preserva positividad, para $y = Bx$ se tiene $y_j \geq 0$ para toda $j = 1, \ldots, n$, y como $A$ también preserva positividad tenemos que $(Px)_j = (Ay)_j \geq 0$ para toda $j = 1, \ldots, n$ por lo tanto, $P$ preserva positividad.

Como $A, B$ son unitales se sigue que

$$P\mathbf{e} = AB\mathbf{e} = A\mathbf{e} = \mathbf{e}.$$

Por la Proposición 1.2.6, esto demuestra que $P$ es biestocástica. Además para cada matriz biestocástica $A$, se tiene que

$$A^* = \overline{A^t} = (\overline{a}_{ji}) = (a_{ji}),$$

de aquí se sigue que $A^*$ es biestocástica.

$(ii)$ Supongamos que $P_\sigma = (p_{ij})$ es una matriz asociada a la permutación $\sigma$, tenemos

$$\begin{array}{rl} p_{ij} \geq 0, & \\ \sum_{j=1}^n p_{ij} = 1 & \text{para toda } i, \\ \sum_{i=1}^n p_{ij} = 1 & \text{para toda } j. \end{array}$$

Es decir, es matriz biestocástica. Para ver que $P_\sigma$ es un punto extremo, por $(ii)$ podemos suponer que $P = tA+(1-t)B$, es decir, $p_{ij} = ta_{ij}+(1-t)b_{ij}$, con $0 < t < 1$ . Como $p_{ij} = 0$ si $j \neq \sigma(i)$ y $a_{ij}, b_{ij} \geq 0$, entonces $a_{ij} = b_{ij} = 0$ si $j \neq \sigma(i)$. Como $A$ y $B$ son matrices biestocásticas, esto implica que $a_{i\sigma(i)} = b_{i\sigma(i)} = 1$ para cada $1 \leq i \leq n$. Entonces $P = A = B$, lo cual demuestra que $P$ es un punto extremo. □

**Lema 1.2.8.** *Sea $N$ un conjunto finito. Si $I, J \subset N$, tales que $|I| + |J| = |N| + 1$ entonces $I \cap J \neq \emptyset$.*

*Demostración.* Supongamos que $I \cap J = \emptyset$. Como $I \cup J \subset N$, entonces $|I \cup J| \leq |N|$. Por otro lado $|I \cup J| = |I| + |J| = |N| + 1$, tenemos una contradicción. □

**Teorema 1.2.9 (Frobenius-König).** *Sea $A$ una matriz de tamaño $n \times n$. Toda diagonal de $A$ tiene un cero si y sólo si $A$ tiene una submatriz de tamaño $s \times t$ idénticamente cero, con $s + t = n + 1$.*

*Demostración.* Siguendo la demostración en [14]. Supongamos que $A$ tiene una submatriz nula de tamaño $s \times t$, con $s + t = n + 1$. Entonces sobre la intersección de renglones $i_1, i_2, \ldots, i_s$ y columnas $j_1, j_2, \ldots, j_t$ hay ceros. Luego, para toda $\sigma \in S_n$ tenemos que $\sigma(i_1), \ldots, \sigma(i_s)$ es nuevamente un conjunto con $s$ elementos. Así por el Lema 1.2.8 al menos un elemento $\sigma(i_l)$ se encuentra en el conjunto $\{j_1, \ldots, j_t\}$, luego la correspondiente diagonal $\{a_{1\sigma(1)}, \ldots, a_{n\sigma(n)}\}$ tiene un cero.

Ahora, supongamos que toda diagonal de $A$ tiene un cero. Procedamos por inducción sobre $n$, la dimensión del espacio. Si $n = 1$, el resultado se sigue.

Supongamos que se cumple para matrices de tamaño $n \times n$. Sea $A$ una matriz no nula de tamaño $(n+1)\times(n+1)$. Seleccionemos una entrada $a_{ij} = 0$ y sea $\widetilde{A}$ la matriz obtenida de quitar el renglón y la columna correspondientes a tal entrada. Como $\widetilde{A}$ es una matriz de tamaño $n\times n$ con todas sus diagonales con un cero, por hipótesis de inducción $\widetilde{A}$ tiene una submatriz nula de tamaño $s_1 \times t_1$ con $s_1 + t_1 = n + 1$. Consideremos la matriz original A y mediante permutaciones de renglones y columnas

$$A = \begin{array}{c c} & \begin{array}{ccc} s_1 & & t_1 \end{array} \\ \begin{array}{c} s_1 \\ \\ t_1 \end{array} & \begin{pmatrix} X & | & O \\ - & - & - \\ Y & | & Z \end{pmatrix} \end{array}$$

Toda diagonal en $X$ la podemos completar con las diagonales en $Z$ para obtener diagonales en $A$. Supongamos que $X$ tiene una diagonal sin ceros. Luego toda diagonal en $Z$ tiene un cero. Así, todas las diagonales en $X$ tienen un cero o bien, todas las diagonales en $Z$ tienen un cero. Digamos que todas las diagonales en $X$ tienen un cero, entonces por hipótesis de inducción $X$ tiene una submatriz idénticamente cero de tamaño $p \times q$ con $p + q = s_1 + 1$. Finalmente, $A$ tiene una matriz nula de tamaño $p \times (q + t_1)$ donde $p + (q + t_1) = t_1 + s_1 + 1 = n + 2$. □

**Corolario 1.2.10.** *Toda matriz $A$ biestocástica tiene una diagonal sin ceros.*

*Demostración.* Supongamos que $A$ tiene todas sus diagonales con un cero. Entonces por el Teorema de Frobenius-König, $A$ tiene una submatriz idénticamente cero de tamaño $s \times t$ con $s + t = n + 1$. Como $A$ es biestocástica se tiene que

$$\sum_{i\,j}^{n} a_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} = \sum_{i=1}^{n} 1 = n.$$

Pero la suma de los renglones y las columnas de $A$ correspondientes a los renglones y las columnas de la matriz nula no tienen términos comunes y suman $s + t > n$. Esta contradicción demuestra que $A$ tiene una diagonal sin ceros. □

**Teorema 1.2.11 (Birkhoff).** *El conjunto de matrices biestocásticas de tamaño $n \times n$ es un conjunto convexo y sus puntos extremos son matrices de permutación.*

*Demostración.* Siguiendo la referencia [1]. En el Teorema 1.2.7 demostramos una parte, falta ver que todo punto extremo es una matriz de permutación. Para esto demostraremos que toda matriz biestocástica $S$ puede escribirse como combinación lineal convexa de matrices de permutación.

Procedamos por inducción sobre $k$, el número de entradas positivas de la matriz biestocástica $S$. En general se tiene que $k \geq n$; si $k = n$, entonces $S$ es una matriz de permutación. Por el Corolario 1.2.10, $S$ tiene una diagonal sin ceros. Supongamos que $S$ no es matriz de permutación, sea $x = \min\{s_{1\sigma(1)}, \ldots, s_{n\sigma(n)}\}$ y $P_\sigma$ la matriz de permutación asociada a $\sigma$. Como $S$ no es de permutación entonces $x < 1$. Sea

$$T = \frac{1}{1-x}(S - xP_\sigma).$$

$T$ es una matriz biestocástica, pues:

$$t_{ij} = \frac{1}{1-x}(s_{ij} - x\delta_{\sigma(i)j}) = \begin{cases} \frac{1}{1-x}(s_{ij} - x) & \text{si } \sigma(i) = j \\ \frac{1}{1-x}s_{ij} & \text{si } \sigma(i) \neq j, \end{cases}$$

para ambos casos resulta que $t_{ij} \geq 0$. Ahora, tenemos que

$$\sum_i t_{ij} = \frac{1}{1-x}\sum_i s_{ij} - \frac{x}{1-x}\sum_i \delta_{\sigma(i)j} = \frac{1}{1-x} - \frac{x}{1-x} = 1, \qquad \forall\, j.$$

Y

$$\sum_j t_{ij} = \sum_j \frac{1}{1-x}(s_{ij} - x\delta_{\sigma(i)j}) = \frac{1}{1-x}\sum_j s_{ij} - \frac{x}{1-x} = 1, \qquad \forall\, i.$$

Además $T$ tiene una entrada nula más que $S$ pues suponiendo $x = s_{i\sigma(i)}$ para algún $1 \leq i \leq n$ y que $\sigma(i) = j$, entonces

$$t_{ij} = \frac{1}{1-x}(s_{ij} - x\delta_{\sigma(i)j}) = \frac{1}{1-s_{ij}}(s_{ij} - s_{ij}) = 0.$$

Así, por hipótesis de inducción $T = \sum_t \rho_t P_t$, donde $\sum_t \rho_t = 1$ y para cada $t$, $P_t$ es matriz de permutación. Luego

$$S = (1-x)T + xP_\sigma = xP_\sigma + (1-x)\sum_t \rho_t P_t,$$

con $\sum_t (1-x)\rho_t + x = (1-x)\sum_t \rho_t + x = 1$. $\square$

**Teorema 1.2.12.** *Una matriz $A \in \mathbf{M}_n$ es biestocástica si y sólo si $Ax \prec x$ $\forall x \in \mathbb{R}^n$.*

*Demostración.* Supóngase que $Ax \prec x$ $\forall x$. Tomando

$$x = e_j = (0, 0, ..., \underbrace{1}_{j}, ..., 0)$$

tenemos que $\sum_{i=1}^n a_{ij} = 1$ $\forall j = 1, 2, \ldots, n$ y además por la Proposición 1.1.3 para cada $1 \leq j \leq n$, se obtiene $a_{ij} \geq 0$ para $1 \leq i \leq n$.

Por otro lado, por $(iii)$ del Teorema 1.1.8 eligiendo $x = \mathbf{e}$ se tiene

$$\sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} - t \right| \leq \sum_{i=1}^n |1 - t| \qquad \text{para toda } t \in \mathbb{R}.$$

Si $t = 1$ obtenemos $\sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} - 1 \right| \leq 0$. Se sigue que $\left| \sum_{j=1}^n a_{ij} - 1 \right|$ es nulo para cada $i$, en consecuencia $\sum_{j=1}^n a_{ij} = 1$ para toda $i = 1, ..., n$. esto demuestra que $A$ es una matríz biestocástica.

Supongamos ahora que $A$ es biestocástica. Como las matrices de permutaciones son biestocásticas y '$\prec$' es relación de equivalencia siempre que $x = Py$ para alguna $P \in S_n$, sin pérdida de generalidad podemos suponer que $x_1 \geq ... \geq x_n$ y $y_1 \geq y_2 \geq ... \geq y_n$ con $y = Ax$.

Ahora, para $1 \leq k \leq n$ tenemos la identidad

$$\sum_{j=1}^k y_j = \sum_{j=1}^k \sum_{s=1}^n a_{js} x_s = \sum_{s=1}^n \sum_{j=1}^k a_{js} x_s.$$

Si escribimos $t_s = \sum_{j=1}^k a_{js}$ tenemos que $0 \leq t_s \leq 1$. Así que

$$\begin{aligned}\sum_{j=1}^k y_j - \sum_{j=1}^k x_j &= \sum_{j=1}^k \sum_{s=1}^n a_{js} x_s - \sum_{j=1}^k x_j = \sum_{s=1}^n \sum_{j=1}^k a_{js} x_s - \sum_{s=1}^k x_s \\ &= \sum_{s=1}^k t_s x_s - \sum_{s=1}^k x_s = \sum_{s=1}^k (t_s - 1)(x_s - x_k) + \sum_{s=k+1}^n t_s (x_s - x_k)\end{aligned}$$

es negativo para toda $1 \leq k \leq n$. Además cuando $k = n$ debemos tener igualdad, pues $A$ es biestocástica. Por tanto $y \prec x$. $\square$

Recordemos que para un operador $A$ su norma está definida como

$$\|A\| = \sup_{\|x\|=1} \|Ax\|.$$

**Proposición 1.2.13.** *Si $A$ es una matriz biestocástica, entonces, todos sus valores propios tienen módulo $\leq 1$, $1$ es valor propio de $A$ y $\|A\| = 1$.*

*Demostración.* Sea $A$ una matriz biestocástica, por la Proposición 1.2.6 es unital y, por lo tanto, 1 es valor propio de $A$ con vector propio asociado $\mathbf{e} = (1, 1, ..., 1)$.

Por el Teorema 1.2.12 tenemos que $Ax \prec x$, para toda $x \in \mathbb{R}^n$. Sea $x$ vector propio de $A$ con valor propio asociado $\lambda$, entonces $\lambda x \prec x$ y por el inciso $(iii)$ de la Proposición 1.1.8, lo anterior se cumple si y sólo si para toda $t \in \mathbb{R}$

$$\sum_{j=1}^{k} |(\lambda x)_j - t| \leq \sum_{j=1}^{k} |x_j - t|, \;\; 1 \leq k \leq n,$$

tomando $t = 0$ se obtiene que

$$|\lambda| \sum_{j=1}^{k} |x_j| \leq \sum_{j=1}^{k} |x_j|, \;\; 1 \leq k \leq n.$$

en consecuencia $|\lambda| \leq 1$. Finalmente, de la proposición 1.2.7, $|A|^2 = A^*A$ es biestocástica, luego por la proposición 1.2.6 cumple que es unital, entonces 1 es valor propio de $|A|^2$.

Ahora, veamos que $\alpha$ es valor propio de $|A|^2$ si y sólo si $\beta = \sqrt{\alpha}$ es valor propio de $|A|$. Como $\|A\| \leq 1$, usando el Lema de la raíz cuadrada A.2.4 tenemos que

$$|A| = I + c_1(I - A^*A) + c_2(I - A^*A)^2 + \dots$$

Sea $h$ vector propio de $|A|^2$, se sigue que

$$\begin{aligned} |A|h &= Ih + c_1(I - A^*A)h + c_2(I - A^*A)^2 h + \dots \\ &= 1h + c_1(1 - \alpha)h + c_2(1 - \alpha)^2 h + \cdots = (\alpha)^2 h. \end{aligned}$$

Y por A.2.5, tenemos

$$\|A\| = \| \, |A| \, \| = s_1(|A|) = 1.$$

□

Recordemos que $|x| = (|x_1|, ..., |x_n|)$ y para vectores $x$, $y \in \mathbb{R}^n$ escribiremos $x \leq y$ siempre que $x_i \leq y_i$ para toda $i = 1, ..., n$.

**Proposición 1.2.14.** *Si $A \in \mathbf{M}_n$ es matriz biestocástica, entonces*

$$|Ax| \leq A(|x|) \qquad \forall x \in \mathbb{R}^n.$$

*Demostración.* Nótese que $|Ax| = \left(\left|\sum_{j=1}^n a_{1j}x_j\right|, ..., \left|\sum_{j=1}^n a_{nj}x_j\right|\right)$. Para cada $k$ con $1 \leq k \leq n$, tenemos que

$$\left|\sum_{j=1}^n a_{kj}x_j\right| \leq \sum_{j=1}^n |a_{kj}x_j| = \sum_{j=1}^n a_{kj}|x_j|.$$

Esto demuestra la proposición. □

**Proposición 1.2.15.** *Toda submatriz cuadrada de una matriz biestocástica es una matriz bisubestocástica. Toda matriz bisubestocástica puede ser dilatada a una matriz biestocástica. Es decir, si $B_{n\times n}$ es matriz bisubestocástica entonces la dilatación $A$ puede ser elegida de tamaño a los más $2n \times 2n$. De hecho, si $R$ es la matriz diagonal cuya $i-$ésima entrada es la suma del $i-$ésimo renglón de $B$ y $C$ es la matriz diagonal cuya $j-$ésima entrada es la suma de la $j-$ésima columna de $B$, entonces*

$$A = \begin{pmatrix} B & I - R \\ I - C & B^* \end{pmatrix} \tag{1.10}$$

*es una matriz biestocástica.*

*Demostración.* Sea $A$ una matriz biestocástica. Consideremos la matriz $B$ tal que $b_{ij} = a_{ij}$ para $i \in \{i_1, \ldots, i_k\}$ y $j \in \{j_1, \ldots, j_k\}$, $k \leq n$; donde $\{i_1, \ldots, i_n\}$ son los renglones de $A$ y $\{j_1, \ldots, j_n\}$ sus columnas . Claramente $b_{ij} \geq 0$. Luego

$$\sum_{j=1}^k b_{ij} = \sum_{j=1}^k a_{ij} \leq \sum_{j=1}^n a_{ij} = 1, \qquad \text{para toda } i;$$

análogamente $\sum_i b_{ij} \leq 1$, para toda $j$, así $B$ es matriz bisubestocástica.

Ahora, sean $R = (r_{ij})$ donde $r_{ii} = \sum_{t=1}^n b_{it}$, $r_{ij} = 0$ para toda $i \neq j$ y $C = (c_{ij})$, donde $c_{jj} = \sum_{t=1}^n b_{tj}$ y $c_{ij} = 0$ para toda $i \neq j$. Consideremos la matriz $A$ como en (1.10). Tenemos que

$$\sum_{j=1}^{2n} a_{ij} = \sum_{j=1}^{n} b_{ij} + \sum_{j=n+1}^{2n} (I_{ij} - r_{ij}) = \sum_{j=1}^{n} b_{ij} + \left(1 - \sum_{j=1}^{n} b_{ij}\right) = 1 \quad \forall i = 1, \ldots, n$$

o bien,

$$\begin{aligned}\sum_{j=1}^{2n} a_{ij} = &\sum_{j=1}^{n}(I_{ij} - c_{ij}) + \sum_{j=n+1}^{2n} b_{ji} = 1 - \sum_{t=1}^{n} b_{ti} + \sum_{j=n+1}^{2n} b_{ji}\\ =&1 - \sum_{t=1}^{n} b_{ti} + \sum_{t=1}^{n} b_{ti} = 1 \qquad\qquad \forall i = n+1, \ldots, 2n.\end{aligned}$$

Análogamente, $\sum_i a_{ij} = 1$ para toda $j = 1, \ldots, 2n$. Así, $A$ es biestocástica. □

**Proposición 1.2.16.** *El conjunto de todas las matrices bisubestocásticas de tamaño $n \times n$ es un conjunto convexo y sus puntos extremos son matrices que tienen a lo más un 1 en cada renglón y cada columna y las entradas restantes cero.*

*Demostración.* Denotemos por $\mathfrak{B}$ al conjunto de matrices bisubestocásticas. Sean $A, B \in \mathfrak{B}$ y sea $0 \leq t \leq 1$, entonces

$$\begin{aligned}(i)\qquad & [tA + (1-t)B]_{ij} = ta_{ij} + (1-t)b_{ij} \geq 0. && \forall\ i, j.\\ (ii)\quad & \sum_{i=1}^{n}(tA + (1-t)B)_{ij} = t\sum_{i=1}^{n} a_{ij} + (1-t)\sum_{i=1}^{n} b_{ij}\\ & \qquad\qquad \leq t + (1-t) = 1 && \forall j,\end{aligned}$$

análogamente $\sum_j (tA + (1-t)B)_{ij} \leq 1$ para toda $i$. Así la matriz $Q = tA + (1-t)B$ es bisubestocástica.

Ahora, supongamos que $P$ es una matriz con a lo más un 1 en cada renglón y cada columna, veamos que es punto extremo. Supongamos

$$p_{ij} = ta_{ij} + (1-t)b_{ij},\ \forall i, j, \tag{1.11}$$

con $A, B \in \mathfrak{B}$ tales que $A \neq B$ y $0 < t < 1$. Si $p_{ij} = 0$, tenemos que $a_{ij} = 0$ implica que $b_{ij} = 0$. Luego $a_{ij}, b_{ij} > 0$. Entonces $0 = ta_{ij} + (1-t)b_{ij}$ no se puede cumplir para $0 < t < 1$ pues tenemos términos estrictamente positivos en el lado derecho. Si $p_{ij} = 1$ de (1.11), tenemos

$$1 = ta_{ij} + (1-t)b_{ij}$$

con $a_{ij}, b_{ij} > 0$ pues si $a_{ij} = 0$ entonces $b_{ij} > 1$ lo cual no puede suceder, análogamente si $b_{ij} = 0$. Si $a_{ij} = 1$ entonces $b_{ij} = 1$, luego $A = B = P$. Si $0 < a_{ij}, b_{ij} < 1$ tenemos $t = \frac{1-b_{ij}}{a_{ij}-b_{ij}}$. Supongamos que $a_{ij} - b_{ij} > 0$, entonces $1 - b_{ij} < a_{ij} - b_{ij}$, de donde $a_{ij} > 1$, lo cual no puede suceder. Si $a_{ij} - b_{ij} < 0$, entonces $0 < t = \frac{1-b_{ij}}{a_{ij}-b_{ij}} < 0$. Esto demuestra que $t = 0$ o bien $t = 1$ en (1.11).

Ahora, para demostrar que todo punto extremo de $\mathfrak{B}$ es una matriz con a lo más un 1 en cada renglón y cada columna, veremos que toda matriz $B \in \mathfrak{B}$ se puede escribir como combinación lineal convexa de tales matrices. Sea $B \in \mathfrak{B}$, por la Proposición 1.2.15 podemos dilatar a $B$ en

$$A = \begin{pmatrix} B & I-R \\ I-C & B^* \end{pmatrix}.$$

Luego, por el Teorema de Birkhoff $A = \sum_t \lambda_t P_t$, con $\sum_t \lambda_t = 1$ y $P_t$ matrices de permutación. Cada submatriz $Q_t$ de una matriz de permutación $P_t$ tiene a lo más un 1 en cada renglón y cada columna y tenemos

$$\sum_t \lambda_t P_t = \begin{pmatrix} \sum_t \lambda_t Q_t & X \\ Y & Z \end{pmatrix}.$$

Así, $B = \sum_t \lambda_t Q_t$. □

**Proposición 1.2.17.** *Una matriz $B$ con entradas no negativas es bisubestocástica si y sólo si existe una matriz biestocástica $A$ tal que $b_{ij} \le a_{ij}$ para toda $i, j = 1, \ldots, n$.*

*Demostración.* Supongamos que $A = (a_{ij})$ es matriz biestocástica tal que $b_{ij} \le a_{ij}$ para todas $i, j = 1, \ldots, n$, para una matriz $B = (b_{ij})$. Entonces

$$\sum_{i=1}^{n} b_{ij} \le \sum_{i=1}^{n} a_{ij} = 1, \qquad \forall\ j = 1, \ldots, n;$$

análogamente $\sum_j b_{ij} \le 1$ para toda $i = 1, \ldots, n$.

Supongamos ahora que $B$ es bisubestocástica. Por la Proposición 1.2.16, existen $\{Q_n\}_{n=1}^N$ puntos extremos del conjunto de matrices bisubestocásticas y $0 \leq t_k \leq 1$ tales que $\sum_k t_k = 1$ y

$$B = \sum_{n=1}^{N} t_n Q_n.$$

Para cada $Q_n$ existe una matriz biestocástica, de hecho una matriz de permutación $P_n$, tal que $q_{ij}^{(n)} \leq p_{ij}^{(n)}$, con $1 \leq i, j \leq n$, basta colocar los 1 que faltan. Tenemos que

$$b_{ij} = \sum_{n=1}^{N} t_n q_{ij}^{(n)} \leq \sum_{n=1}^{N} t_n p_{ij}^{(n)} := a_{ij} \qquad \forall\, i, j = 1, ..., n.$$

Por el Teorema de Birkhoff $A = (a_{ij})$ es una matriz biestocástica. $\square$

**Proposición 1.2.18.** *(i) Sean $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$, existe $0 \leq t \leq 1$ tal que*

$$(x_1, x_2) = (ty_1 + (1-t)y_2, (1-t)y_1, +ty_2)$$

*si y sólo si $x \prec y$.*

*(ii) Sean $x, y \in \mathbb{R}^n$ tales que $x$ se obtiene promediando cualesquiera dos coordenadas de $y$ como en $(i)$, dejando fijas a las demás coordenadas. Entonces $x \prec y$.*

*Demostración.* Demostremos $(i)$. Supóngase que $y_1 = \text{máx}\,(y_1, y_2)$, entonces

$$\begin{aligned} x_1 &= ty_1 + (1-t)y_2 \leq ty_1 + (1-t)y_1 = y_1 \\ x_2 &= (1-t)y_1 + ty_2 \leq (1-t)y_1 + ty_1 = y_1 \end{aligned}$$

Así tenemos que $\text{máx}(x_1, x_2) \leq \text{máx}(y_1, y_2)$ es decir $x \prec y$. El otro caso se demuestra de manera similar.

Recíprocamente, si $x \prec y$ entonces como $x_2^\downarrow \leq x_1^\downarrow \leq y_1^\downarrow$ y $x_1^\downarrow + x_2^\downarrow = y_1^\downarrow + y_2^\downarrow$, tenemos que $x_1^\downarrow = y_2^\downarrow + (y_1^\downarrow - x_2^\downarrow) \geq y_2^\downarrow$. Consecuentemente existe $t \in [0,1]$ tal que $x_1 = ty_1^\downarrow + (1-t)y_2^\downarrow$ y $x_2^\downarrow = y_1^\downarrow + y_2^\downarrow - ty_1^\downarrow - (1-t)y_2^\downarrow = ty_2^\downarrow + (1-t)y_1^\downarrow$. Esto demuestra el recíproco en el caso cuando $x_1^\downarrow = x_1$ y $y_1^\downarrow = y_1$, los casos restantes se demuestran de manera similar.

La demostración de $(ii)$ es similar a la primera parte de la demostración de $(i)$. $\square$

**Definición 1.2.19.** *Diremos que una transformación lineal $T$ sobre $\mathbb{R}^n$, es una $t-$**transformación** si existen $0 \leq t \leq 1$ e índices $i, j$ tales que:*

$$Ty = (y_1, ..., \underbrace{ty_i + (1-t)y_j}_{i}, ..., \underbrace{(1-t)y_i + ty_j}_{j}, ..., y_n). \tag{1.12}$$

**Proposición 1.2.20.** *Si $T$ es una $t-$transformación, entonces es combinación lineal convexa del operador identidad y alguna matriz de permutación $P$. Luago $T$ es biestocástica y $Ty \prec y$ para toda $y \in \mathbb{R}^n$.*

*Demostración.* Nótese que una $t-$transformación es una combinación convexa de la aplicación identidad y alguna permutación $P \in S_n$, en efecto,

$$\begin{aligned} Ty &= (y_1, ..., y_{i-1}, ty_i + (1-t)y_j, y_{i+1}, ..., y_{j-1}, (1-t)y_i + ty_j, y_{j+1}, ..., y_n) \\ &= (ty_1 + (1-t)y_1, ..., ty_i + (1-t)y_j, ..., (1-t)y_i + ty_j, ..., ty_n + (1-t)y_n) \\ &= (tId + (1-t)P)y. \end{aligned}$$

Luego por la Proposición 1.2.7 tenemos que $T$ es biestocástica. Así, del Teorema 1.2.12 se sigue el resultado. □

**Teorema 1.2.21.** *Para toda $x, y \in \mathbb{R}^n$, las siguientes afirmaciones son equivalentes:*

*(i)* $x \prec y$*;*

*(ii) El vector $x$ se obtiene a partir del vector $y$ mediante un número finito de $t-$transformaciones;*

*(iii) $x$ pertenece a la envolvente convexa de todos los vectores obtenidos al permutar las coordenadas de $y$;*

*(iv) $x = Ay$ para alguna matriz biestocástica $A$.*

*Demostración.* Probaremos $(i) \Rightarrow (ii)$ por inducción sobre la dimensión.

El caso $n = 2$ se sigue del inciso $(i)$ de la Proposición 1.2.18.

Supongamos que la conclusión se cumple para dimensiones hasta $n-1$. Sean $x, y \in \mathbb{R}^n$. Los vectores $x^{\downarrow}, y^{\downarrow}$ se pueden obtener a partir de $x, y$ mediante permutaciones $P \in S_n$. Cada permutación es producto de transposiciones $\tau \in S_n$ y cada transposición es una $t-$ transformación, pues tomando $t = 0$ en la definición 1.2.19 obtenemos

$$Ty = \tau y = (y_1, ..., y_{j-1}, y_k, y_{j+1}, ..., y_j, y_{k+1}, ..., y_n).$$

Sin pérdida de generalidad podemos suponer que $x_1 \geq x_2 \geq ... \geq x_n$ y que $y_1 \geq y_2 \geq ... \geq y_n$.

Si $x \prec y$, entonces $x_1 \leq y_1$ y $ny_n \leq \sum_{j=1}^n y_j = \sum_{j=1}^n x_j \leq nx_1$, de donde se ve que $y_n \leq x_1 \leq y_1$. Elijamos $1 \leq k \leq n$ tal que $y_k \leq x_1 \leq y_{k-1}$ entonces $x_1 = ty_1 + (1-t)y_k$, para alguna $0 \leq t \leq 1$. Sea $T_1$ la siguiente $t-$transformación

$$T_1 u = (tu_1 + (1-t)u_k, u_2, ..., u_{k-1}, (1-t)u_1 + tu_k, u_{k+1}, ..., u_n) \quad \forall u \in \mathbb{R}^n.$$

Nótese que la primer coordenada de $T_1 y$ es $x_1$. Sean

$$x' = (x_2, \ldots, x_n) \qquad y' = (y_2, ..., y_{k-1}, (1-t)y_1 + ty_k, y_{k+1}, ..., y_n).$$

Mostraremos que $x' \prec y'$. Como $y_1 \geq ... \geq y_{k-1} \geq ty_1 + (1-t)y_k = x_1 \geq x_2 \geq ... \geq x_n$, si $2 \leq m \leq k-1$ entonces se tiene $\sum_{j=2}^m y_j \geq \sum_{j=2}^m x_j$. Y para $k \leq m \leq n$ se tiene

$$\begin{aligned} \sum_{j=2}^m y'_j &= \sum_{j=2}^{k-1} y_j + ((1-t)y_1 + ty_k) + \sum_{j=k+1}^m y_j \\ &= \sum_{j=1}^m y_j - ty_1 + ty_k - y_k = \sum_{j=1}^m y_j - ty_1 - (1-t)y_k \\ &= \sum_{j=1}^m y_j - x_1 \geq \sum_{j=1}^m x_j - x_1 = \sum_{j=2}^m x'_j. \end{aligned}$$

Si $m = n$, tenemos $\sum_{j=2}^n x_j = \sum_{j=2}^n y_j$. Hemos demostrado que $x' \prec y'$.

Por la hipótesis de inducción, existe un número finito de $t-$ transformaciones $T_2, ..., T_r \in \mathbb{R}^{n-1}$ tales que $x' = (T_r \cdots T_2)y'$. Y podemos considerar que cada $T_j$ es una $t-$transformación en $\mathbb{R}^n$ que no mueve la primera coordenada, es decir

$$(T_r \cdots T_1)y = (T_r \cdots T_2)(x_1, y') = (x_1, x') = x.$$

Esto demuestra que $(i)$ implica $(ii)$.

Que $(ii) \Rightarrow (iii)$, se sigue de la Proposición 1.2.20, donde cada factor $T_j y = (tId + (1-t)P_j)y$ define una matriz biestocástica, luego por la Proposición 1.2.7 se tiene que $(iii) \Rightarrow (iv)$.

La implicación $(iv) \Rightarrow (i)$ es consecuencia del Teorema 1.2.12. $\square$

Recordemos que una matriz $U$ es unitaria si $U^*U = I = UU^*$.

**Proposición 1.2.22.** *(i) Si $U = (u_{ij})$ es una matriz unitaria, entonces la matriz $(|u_{ij}|^2)$ es biestocástica, diremos que es matriz uni-estocástica y se llamará ortoestocástica si $U$ es ortogonal real.*

*(ii) Si $x = Ay$ para alguna matriz biestocástica $A$, entonces existe una matriz ortoestocástica $B$ tal que $x = By$.*

*Demostración.* $(i)$ Supóngase que $U = (u_{ij})$ es matriz unitaria. De las identidades $U^*U = (\sum_{k=1}^{n} \overline{u_{ki}} u_{kj})_{ij} = I$ y $UU^* = (\sum_{k=1}^{n} u_{ik}\overline{u_{jk}})_{ij} = I$, se obtiene respectivamente

$$\sum_{k=1}^{n} \overline{u_{kj}} u_{kj} = \sum_{k=1}^{n} |u_{kj}|^2 = 1 \qquad \text{para toda } j = 1, ..., n \tag{1.13}$$

$$\sum_{k=1}^{n} u_{ik} \overline{u_{ik}} = \sum_{k=1}^{n} |u_{ik}|^2 = 1 \qquad \text{para toda } i = 1, ..., n. \tag{1.14}$$

Y también tenemos $\overline{u_{ij}} u_{ij} = |u_{ij}|^2 \geq 0$ para toda $i, j = 1, ..., n$. Por lo tanto,

$$A = \begin{pmatrix} |u_{11}|^2 & |u_{12}|^2 & \cdots & |u_{1n}|^2 \\ \vdots & & & \vdots \\ |u_{n1}|^2 & |u_{n2}|^2 & \cdots & |u_{nn}|^2 \end{pmatrix}$$

es biestocástica.

$(ii)$ Siguiendo la demostración en [10]. Supongamos que $x = Ay$ para alguna matriz biestocástica $A$, entonces por el Teorema 1.2.21 el vector $x$ se obtiene a partir de $y$ mediante un número número finito de $t-$transformaciones, es decir, $x = T_1 T_2 ... T_n y$. Para demostrar que $T_1 T_2 ... T_n = (W_{ij}^2)$ para alguna matriz ortogonal real $W$, procederemos por inducción sobre el número de $t-$transformaciones. Supóngase que $n = 1$, la dimensión de la matriz es arbitraria. De la Definición 1.2.19 tenemos que $T_1 = tI + (1-t)P_1$ para $0 \leq t \leq 1$ y alguna matriz de permutación $P_1$. Omitiendo las entradas donde $T_1$ actúa como identidad, escribiremos

$$T_1 = \begin{bmatrix} t & 1-t \\ 1-t & t \end{bmatrix}. \tag{1.15}$$

Definamos una matriz unitaria real $U$, que actúa como identidad en aquellas entradas donde $T_1$ lo hace y mediante la matriz

$$U \equiv \begin{bmatrix} \sqrt{t} & -\sqrt{1-t} \\ \sqrt{1-t} & \sqrt{t} \end{bmatrix} \tag{1.16}$$

en las entradas donde $T_1$ no actúa trivialmente. Tenemos que $U^*U = I = UU^*$ y que $T_1 = (U_{ij}^2)$. Para hacer el paso inductivo, supóngase que el producto de $n$ $t-$transformaciones para matrices de tamaño $d \times d$, con $d$ arbitraria, es ortoestocástico y considérese el producto $T_1T_2...T_{n+1}$. Supongamos que para $1 \leq k \leq n+1$ y cualquier $y \in \mathbb{C}^n$, la $t-$transformación $T_{n+2-k}$ actúa sobre la coordenada $k$ y alguna coordenada $d_k > k$ en $y$. Luego $T_{n+1}$ actúa sobre la primer coordenada y la coordenada $d_1 > 1$ de $y$. Sea $P$ la matriz correspondiente a la trasposición de las componentes 2 y $d_1$ de cualquier vector $y$, entonces

$$PT_{n+1}P = \begin{bmatrix} t & 1-t & 0 \\ 1-t & t & 0 \\ 0 & 0 & I_{d-2} \end{bmatrix} \tag{1.17}$$

donde $I_{d-2}$ es la matriz identidad de tamaño $(d-2) \times (d-2)$. Como el producto $T_1T_2...T_n$ actúa como la identidad en la primer columna y el primer renglón podemos definir la matriz $\Delta$ de tamaño $(d-1) \times (d-1)$ como

$$T_1T_2...T_n = \begin{bmatrix} 1 & 0 \\ 0 & \Delta \end{bmatrix}. \tag{1.18}$$

Si omitimos el primer renglón y la primer columna de la matriz $T_1T_2...T_n$ obtenemos $T_1'T_2'...T_n'$ de tamaño $(d-1)\times(d-1)$ tales que $T_1'T_2'...T_n' = \Delta$. Por hipótesis de inducción existe una matriz de tamaño $(d-1)\times(d-1)$ ortogonal $U = (U_{ij})$ tal que $\Delta_{ij} = U_{ij}^2$. Defíimanos la matriz $U'$ obtenida mediante $U' = P'UP'$ donde $P'$ es la matriz que traspone las primer coordenada y la coordenada $d_1 - 1$ de vectores $y$, de manera similarmente definimos $\Delta'$ mediante $\Delta' \equiv P'\Delta P'$. Entonces $\Delta'_{ij} = U'^2_{ij}$. También tenemos

$$PT_1T_2...T_nP = \begin{bmatrix} 1 & 0 \\ 0 & \Delta' \end{bmatrix}. \tag{1.19}$$

Multiplicando la ecuación anterior por $PT_{n+1}P$ obtenemos, de (1.15) y la identidad $P^2 = I$ que

$$PT_1T_2...T_{n+1}P = \begin{bmatrix} t & 1-t & 0 \\ (1-t)\vec{\delta} & t\vec{\delta} & \tilde{\Delta} \end{bmatrix}, \tag{1.20}$$

donde $\vec{\delta}$ es la primer columna de $\Delta'$ y $\tilde{\Delta}$ es la matriz de $(d-2) \times (d-1)$ que resulta cuando la primer columna de $\Delta'$ es removida. Sea $\tilde{U}$ la matriz de

$(d-2) \times (d-1)$ que resulta cuando la primer columna de $U'$ es removida y sea $\vec{u}$ la primer columna de $U'$. Definimos una matriz de $d \times d$ por

$$V = \begin{bmatrix} \sqrt{t} & -\sqrt{1-t} & 0 \\ \sqrt{1-t}\vec{u} & \sqrt{t}\vec{u} & \tilde{U} \end{bmatrix}. \tag{1.21}$$

$V$ es una matriz ortogonal. Para ver esto comprobamos que las columnas de $V$ son unitarias y ortogonales. Para la primer columna tenemos

$$\sqrt{t + (1-t)\vec{u} \cdot \vec{u}} = \sqrt{1} = 1, \tag{1.22}$$

y de manera similar para la segunda columna. Las columnas restantes son unitarias pues son columnas de la matriz unitaria $U'$. Para la ortogonalidad entre columnas, tomando las dos primeras

$$(\sqrt{t})(-\sqrt{1-t}) + (\sqrt{t})(\sqrt{1-t})\vec{u} \cdot \vec{u} = 0. \tag{1.23}$$

Para las columnas restantes sabemos que $\vec{u}$ es ortogonal a las columnas de $\tilde{U}$. En consecuencia tenemos $PT_1T_2...T_{n+1}P = ({V^2}_{ij})$, entonces definiendo $W = PVP$ tenemos que $W$ es una matriz ortogonal tal que $T_1T_2...T_{n+1} = ({W^2}_{ij})$, lo cual completa la inducción. Usando nuevamente el Teorema 1.2.21 se sigue el resultado. $\square$

Recordemos que una matriz $A$ es hermitiana si $A = A^*$, donde $A^* = \overline{A^t}$ es la matriz adjunta de $A$.

**Teorema 1.2.23 (Lema de Schur).** *Sea $A \in \mathbf{M}_n$ una matriz hermitiana; sea* $\operatorname{diag}(A)$ *el vector cuyas coordenadas corresponden a los elementos $a_{ii}$ de $A$ y sea $\lambda(A)$ el vector cuyas coordenadas corresponden a los valores propios de $A$ especificados en cualquier orden. Entonces*

$$\operatorname{diag}(A) \prec \lambda(A).$$

*Demostración.* Dada $A$ matriz hermitiana, por el Teorema espectral existe una matriz $U$ unitaria tal que diagonaliza a $A$, podemos escribir

$$A = U\operatorname{diag}(\lambda(A))U^*$$

de aquí tenemos

$$A = \begin{pmatrix} \sum_{j=1}^n u_{1j}\lambda_j\overline{u_{1j}} & \dots & \sum_{j=1}^n u_{1j}\lambda_j\overline{u_{nj}} \\ \vdots & & \vdots \\ \sum_{j=1}^n u_{nj}\lambda_j\overline{u_{1j}} & \dots & \sum_{j=1}^n u_{nj}\lambda_j\overline{u_{nj}} \end{pmatrix}. \tag{1.24}$$

Así que

$$\operatorname{diag}(A) = \left( \sum_{j=1}^{n} u_{1j}\lambda_j \overline{u_{1j}}, \ldots, \sum_{j=1}^{n} u_{nj}\lambda_j \overline{u_{nj}} \right), \tag{1.25}$$

entonces podemos escribir $a_{ii} = \sum_j |u_{ij}|^2 \lambda_j$. Esto implica que $\operatorname{diag}(A) = S\lambda(A)$, donde $S$ es la matriz $S = (|u_{ij}|^2)$, que es biestocástica por la parte $(i)$ de la Proposición 1.2.22. Ahora, una aplicación de la parte $(i)$ del Teorema 1.2.21 nos permite concluir que $\operatorname{diag}(A) \prec \lambda(A)$. □

**Teorema 1.2.24 (Principio del máximo de Ky Fan).** *Si $\lambda(A)^{\downarrow}$ denota el vector de valores propios ordenados en forma decreciente de una matriz hermitiana $A$, entonces para toda $k = 1, 2, ..., n$*

$$\sum_{j=1}^{k} \lambda_j^{\downarrow} = \text{máx} \sum_{j=1}^{k} \langle x_j, Ax_j \rangle, \tag{1.26}$$

*donde el máximo es tomado sobre todas las $k-$uplas de vectores ortonormales $\{x_1, \cdots, x_k\}$ en $\mathbb{C}^n$.*

*Demostración.* Sea $\{x_1, ..., x_k\}$ un conjunto ortonormal en $\mathbb{C}^n$, que podemos completar a una base $\{x_1, ..., x_k, ..., x_n\}$ en $\mathbb{C}^n$. Si $(a_{ij})$ es la matriz de $A$ respecto de esta base, entonces

$$Ax_j = \sum_{i=1}^{n} a_{ij} x_i \qquad \forall j. \tag{1.27}$$

Por lo tanto,

$$\langle x_k, Ax_l \rangle = \left\langle x_k, \sum_{i=1}^{n} a_{il} x_i \right\rangle = \sum_{i=1}^{n} a_{il} \langle x_k, x_l \rangle = a_{kl}. \tag{1.28}$$

Por el Lema de Schur, tenemos $\operatorname{diag}(A) = (\langle x_1, Ax_1 \rangle, \ldots, \langle x_n, Ax_n \rangle) \prec \lambda(A)$. Por la proposición 1.1.4 se sigue que

$$\sum_{j=1}^{k} \langle x_j, Ax_j \rangle \leq \sum_{j=1}^{k} \lambda_j^{\downarrow} \qquad 1 \leq k \leq n. \tag{1.29}$$

En particular, si $\{x_1, ..., x_k\}$ es un subconjunto ortonormal de vectores propios de $A$ tales que $Ax_j = \lambda_j x_j$, entonces

$$\sum_{j=1}^{k} \langle x_j, Ax_j\rangle^{\downarrow} = \sum_{j=1}^{k} \langle x_j, \lambda_j x_j\rangle^{\downarrow} = \sum_{j=1}^{k} \lambda_j^{\downarrow} \qquad 1 \leq k \leq n. \tag{1.30}$$

Es decir, la suma se satura alcanzando su máximo. Esto demuestra que

$$\text{máx} \sum_{j=1}^{k} \langle x_j, Ax_j\rangle = \sum_{j=1}^{k} \lambda_j^{\downarrow} \qquad 1 \leq k \leq n \tag{1.31}$$

$\square$

Nótese que lo anterior se cumple para toda $1 \leq k \leq n$ y sobre algún conjunto ortonormal $\{x_l\}_{l=1}^{k}$, tenemos

$$\sum_{j=1}^{k} \lambda_j^{\downarrow} = \text{máx} \sum_{j=1}^{k} \langle x_j, Ax_j\rangle \geq \sum_{j=1}^{k} \langle x_j, Ax_j\rangle^{\downarrow} = \sum_{j=1}^{k} a_{jj}^{\downarrow}. \tag{1.32}$$

Además si $k = n$ el máximo del lado izquierdo se alcanza en la base ortonormal $\{x_1, \cdots, x_n\}$ de los vectores propios de $A$ y la desigualdad anterior se convierte en igualdad. Así, el principio de mayorización de Ky-Fan es equivalente con el Lema de Schur.

**Proposición 1.2.25.** *Sean $A, B$ matrices hermitianas, entonces para toda $k = 1, ..., n$*

$$\sum_{j=1}^{k} \lambda_j^{\downarrow}(A+B) \leq \sum_{j=1}^{k} \lambda_j^{\downarrow}(A) + \sum_{j=1}^{k} \lambda_j^{\downarrow}(B). \tag{1.33}$$

*Demostración.* Usando el Teorema 1.2.24 y tomando máximos sobre un subconjunto ortornormal $\{x_1, \cdots, x_k\}$ en $\mathbb{C}^n$

$$\begin{aligned}\sum_{j=1}^{k} \lambda_j^{\downarrow}(A+B) &= \text{máx} \sum_{j=1}^{k} \langle x_j, (A+B)x_j\rangle = \text{máx} \sum_{j=1}^{k} \Big( \langle x_j, Ax_j\rangle + \langle x_j, Bx_j\rangle \Big) \\ &\leq \text{máx} \sum_{j=1}^{k} \langle x_j, Ax_j\rangle + \text{máx} \sum_{j=1}^{k} \langle x_j, Bx_j\rangle \\ &= \sum_{j=1}^{k} \lambda_j^{\downarrow}(A) + \sum_{j=1}^{k} \lambda_j^{\downarrow}(B).\end{aligned}$$

$\square$

Dada una matriz $A$ de tamaño $n \times n$, decimos que $B$ es la **raíz cuadrada** de la matriz $A$ si $B \cdot B = A$ y **el valor absoluto** $|A|$ se define mediante la relación $|A| = \sqrt{A^*A}$. A los valores propios de la matriz $|A|$, tomando en cuenta sus multiplicidades, se les llaman valores **singulares** de $A$, que denotaremos por $s_1(A), \ldots, s_n(A)$ y los tomaremos ordenados en forma decreciente.

**Proposición 1.2.26.**

*(a) Para cualquier matriz $A$, sea $\widetilde{A}$ la matriz:*

$$\widetilde{A} = \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}.$$

*Entonces, $\widetilde{A}$ es hermitiana y sus valores propios son los valores singulares de $A$ junto con sus negativos.*

*(b) Para cualesquiera matrices $A, B$ y todo $k = 1, 2, \ldots, n$ se cumple*

$$\sum_{j=1}^{k} s_j(A+B) \leq \sum_{j=1}^{k} s_j(A) + \sum_{j=1}^{k} s_j(B).$$

*Demostración.*

(a) Efectivamente $\widetilde{A}$ es hermitiana pues

$$(\widetilde{A})^* = \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}^* = \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix} = \widetilde{A}, \tag{1.34}$$

ya que $A^{**} = A$. Sea $\lambda$ un valor propio de $\widetilde{A}$, con vector propio $(X_1, X_2)$

$$\begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \lambda \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}. \tag{1.35}$$

Tenemos que $AX_2 = \lambda X_1$ y $A^*X_1 = \lambda X_2$, como $\widetilde{A}$ es hermitiana se sigue que $\lambda \in \mathbb{R}$, luego

$$|A|^2 X_2 = A^*AX_2 = \lambda A^*X_1 = \lambda^2 X_2.$$

Por tanto $|\lambda| = \sqrt{\lambda^2}$ es valor singular de $A$. Recíprocamente, si $W^*AQ = S$ es la descomposición en valores singulares de A, es decir $S = \operatorname{diag}(s_1(A), \ldots, s_n(A))$, $W$ es la matriz unitaria cuyas columnas $W_j$ son los vectores propios de $A^*A$

y $Q$ es la matriz unitaria cuyas columnas $Q_j$ son los vectores propios de $AA^*$ correspondientes a los valores propios $s_j^2(A)$, para $1 \leq j \leq n$. Entonces, para cada $j$ tenemos que

$$AQ_j = s_(A)W_j, \quad y \quad A^*W_j = s_j(A)Q_j,$$

es decir, $s_j(A)$ y $-s_j(A)$ son valores propios de $\widetilde{A}$ con vectores propios asociados $(W_j, Q_j)$ y $(-W_j, Q_j)$.

(b) Sean

$$\widetilde{A} = \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix} \qquad \widetilde{B} = \begin{pmatrix} 0 & B \\ B^* & 0 \end{pmatrix},$$

Por la Proposición 1.2.25 tenemos que

$$\sum_{j=1}^{k} \lambda_j^{\downarrow}(\widetilde{A} + \widetilde{B}) \leq \sum_{j=1}^{k} \lambda_j^{\downarrow}(\widetilde{A}) + \sum_{j=1}^{k} \lambda_j^{\downarrow}(\widetilde{B}).$$

Pero por el inciso anterior, para $1 \leq k \leq n$ podemos escribir

$$\sum_{j=1}^{k} s_j(A+B) \leq \sum_{j=1}^{k} s_j(A) + \sum_{j=1}^{k} s_j(B). \qquad \square$$

pues los negativos de los valores propios de $|A|$ aparecen después de la posición $n$.

En la proposición anterior, tomando $k = 1$ tenemos

$$s_1(A+B) \leq s_1(A) + s_1(B),$$

que es la desigualdad del triángulo para operadores sobre un espacio de dimensión finita, con la norma de operadores. Esto es

$$||A + B|| \leq ||A|| + ||B||.$$

# Capítulo 2

# Normas simétricas sobre $\mathbb{C}^n$

## 2.1. Normas en $\mathbb{C}^n$

Comencemos tomando las $p$-normas sobre $\mathbb{C}^n$. Para cada $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$

$$||x||_p = \left( \sum_{j=1}^{n} |x_j|^p \right)^{1/p} \qquad 0 \leq p < \infty$$
$$||x||_\infty = \max_{1 \leq i \leq n} |x_i|.$$

Para $1 \leq p \leq \infty$ estas funciones definen normas en $\mathbb{C}^n$.

Veamos que $||x||_\infty = \lim_{p\to\infty} ||x||_p$. Nótese que $|x_k| \leq \|x\|_\infty$ para toda $1 \leq k \leq n$, en consecuencia $\sum_{j=1}^{n} |x_j|^p \leq n\|x\|_\infty^p$. Como la función $u \mapsto u^{1/p}$ preserva el orden, se obtiene que $\left( \sum_{j=1}^{n} |x_j|^p \right)^{1/p} \leq (n\|x\|_\infty^p)^{1/p}$. Entonces

$$\liminf_{p\to\infty} \left( \sum_{j=1}^{n} |x_j|^p \right)^{1/p} \leq \liminf_{p\to\infty} n^{1/p}\|x\|_\infty = \|x\|_\infty. \tag{2.1}$$

Tenemos así una desigualdad. Ahora como siempre se cumple

$$\|x\|_\infty^p \leq |x_1|^p + \cdots + |x_n|^p,$$

tenemos que $\left(\|x\|_\infty^p\right)^{1/p} \leq \left( \sum_{j=1}^{n} |x_j|^p \right)^{1/p}$. Tomando el límite superior cuando $p \to \infty$ obtenemos la otra desigualdad. A continuación probaremos algunas propiedades de la familia de $p-$normas.

**Proposición 2.1.1.** *Para toda $x \in \mathbb{C}^n$ y toda $1 \leq p \leq \infty$ se cumplen:*

*(i) Propiedad del módulo o calibración: $\|x\|_p = \|\,|x|\,\|_p$.*

*(ii) Propiedad de monotonía: $\|x\|_p \leq \|y\|_p$ siempre que $|x| \leq |y|$.*

*(iii) Propiedad de simetría: $\|x\|_p = \|Px\|_p$ $\forall P$, donde $P$ es una matriz de permutación.*

*Demostración.* *(i)* Sea $x \in \mathbb{C}^n$ y $1 \leq p < \infty$ entonces

$$\|\,|x|\,\|_p = \left( \sum_{j=1}^{n} |\,|x_j|\,|^p \right)^{1/p} = \left( \sum_{j=1}^{n} |x_j|^p \right)^{1/p} = \|x\|_p.$$

Para $p = \infty$; $\|\,|x|\,\|_\infty = \text{máx}_{1 \leq i \leq n} |\,|x_i|\,| = \text{máx}_{1 \leq i \leq n} |x_i| = \|x\|_\infty$.

*(ii)* Sea $x \in \mathbb{C}^n$, supongamos que $|x_j| \leq |y_j|$ para toda $j = 1, 2, ..., n$. Entonces para toda $1 \leq p < \infty$ se cumple que $|x_j|^p \leq |y_j|^p$ para cada $j = 1, 2, ..., n$. Sumando las coordenadas tenemos que $\sum_j |x_j|^p \leq \sum_j |y_j|^p$. Como la función $u \mapsto u^{1/p}$ e s monótona tenemos que

$$\|x\|_p = \left( \sum_{j} |x_j|^p \right)^{1/p} \leq \left( \sum_{j} |y_j|^p \right)^{1/p} = \|y\|_p.$$

Para $p = \infty$, si $|x_j| \leq |y_j|$ para toda $j = 1, 2, ..., n$ entonces $\|x\|_\infty = \text{máx}_j |x_j| \leq \text{máx}_j |y_j| = \text{máx} \|y\|_\infty$.

*(iii)* Sean $x \in \mathbb{C}^n$ y $P \in S_n$, como $P$ es matriz de permutación entonces $Px = \big(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\big)$. Sea $1 \leq p < \infty$, calculamos

$$\|Px\|_p = \left( \sum_{j=1}^{n} |x_{\sigma(j)}|^p \right)^{1/p} = \left( \sum_{j=1}^{n} |x_j|^p \right)^{1/p} = \|x\|_p.$$

Ahora si $p = \infty$ calculamos $\|Px\|_\infty = \text{máx}_j |x_{\sigma(j)}| = \text{máx}_j |x_j| = \|x\|_\infty$. □

Observese que si suponemos el espacio de dimensión infinita y que está dotado de la $p$-norma, se trata del espacio $\ell_p$. Las primeras dos condiciones en la Proposición 2.1.1 se cumplen claramente pues las sucesiones son absolutamente convergentes. Por el Teorema de reordenamiento de series la última condición en la Proposición 2.1.1 se cumple, por lo tanto la proposición también es válida en $\ell_p$.

**Proposición 2.1.2.** *Una norma en $\mathbb{C}^n$ es calibrada si y sólo si es monótona.*

*Demostración.* Supongamos que $\|\cdot\|$ es monótona. Sea $x \in \mathbb{C}^n$ entonces $|x| \leq \big||x|\big|$ y por la monotonía de la norma tenemos que $\|x\| \leq \|\,|x|\,\|$. Por otro lado, $\big||x|\big| \leq |x|$, así por la monotonía de la norma se tiene que $\|\,|x|\,\| \leq \|x\|$ por lo tanto $\|\cdot\|$ es una norma calibrada.

Recíprocamente, supongamos que $\|\cdot\|$ es una norma calibrada. Sean $x, y \in \mathbb{C}^n$ tales que $|x_j| \leq |y_j|$ para toda $j = 1, 2, \ldots, n$, si tomamos $x_j = t_j y_j$ para $y_j \neq 0$ entonces $|t_j| \leq 1$. Luego

$$\begin{aligned}
\|x\| &= \|(t_1 y_1, \ldots, t_n y_n)\| \\
&= \left\| \left( \frac{1+t_1}{2} y_1 - \frac{1-t_1}{2} y_1, \ldots, \frac{1+t_1}{2} t_n y_n + \frac{1-t_1}{2} t_n y_n \right) \right\| \\
&\leq \frac{1+t_1}{2} \|(y_1, y_2, \ldots, t_n y_n)\| + \frac{1-t_1}{2} \|(-y_1, y_2, \ldots, t_n y_n)\| \\
&= \|(y_1, t_2 y_2 \ldots, t_n y_n)\| \leq \cdots \leq \|(y_1, y_2 \ldots, t_n y_n)\| \leq \|(y_1, y_2 \ldots, y_n)\|.
\end{aligned}$$

Donde para la primer desigualdad se aplicó la desigualdad del triángulo y la calibración sobre el primer sumando y se continuó con el procedimiento para las coordenadas restantes. $\square$

**Ejemplo 2.1.3.** *Consideremos las siguientes normas en $\mathbb{R}^2$.*

***a.*** $\|x\|_a = |x_1| + |x_2| + |x_1 - x_2|$ *cumple la propiedad de simetría pero no la calibración, pues*

$$\|Px\|_a = |x_2| + |x_1| + |x_2 - x_1| = \|x\|_a.$$

*Por otro lado tomando $x_1 = -1$, $x_2 = 1$*

$$\|\;|x|\;\|_a = \|(-1, 1)\| = |-1| + |1| + |\,|-1| - |1|\,| = 2 \neq 4 = \|x\|_a.$$

***b.*** $\|x\|_b = |x_1| + |x_1 - x_2|$ *no es simétrica ni calibrada. Tomando $x = (-1, 2)$*

$$\begin{aligned}
\|x\|_b = |-1| + |-1-2| = 4 \neq 5 = |2| + |2-(-1)| = \|Px\|_b \\
\|\;|x|\;\|_b = |\,|-1\,|| + |\,|-1| - |2|\,| = 2 \neq 4 = \|x\|_b.
\end{aligned}$$

***c.*** $\|x\|_c = 2|x_1| + |x_2|$ *no es simétrica pero sí es calibrada. Consideremos* $x = (1, 2)$

$$\begin{aligned}
\|Px\|_c = \|(2, 1)\|_c = 2|2| + |1| = 5 \neq 2|1| + |2| = \|x\|_c \\
\|\;|x|\;\|_c = 2|\;|x_1|\;| + |\;|x_2|\;| = \|x\|_c.
\end{aligned}$$

## 2.2. Normas simétricas y calibradas

Por la propiedad de calibración vista en la sección anterior, las normas simétricas y calibradas en $\mathbb{C}^n$ están determinadas por aquellas en $\mathbb{R}^n$.

**Definición 2.2.1.** *Una función $\Phi : \mathbb{R}^n \to \mathbb{R}_+$ es simétrica y calibrada si*

*(i)* $\Phi$ *es una norma.*

*(ii)* $\Phi(Px) = \Phi(x)$ *para toda* $x \in \mathbb{R}^n$ *y toda* $P \in S_n$*.*

*(iii)* $\Phi(\varepsilon_1 x_1, \ldots, \varepsilon_n x_n) = \Phi(x_1, \ldots, x_n)$*, donde* $\varepsilon_j = \pm 1$*,* $j = 1, 2, ..., n$*.*

*(iii)* $\Phi(\mathbf{e_1}) = \Phi(1, 0, \ldots, 0) = 1$ *($\Phi$ esta normalizada).*

Por la condición *(iii)* una función simétrica y calibrada está determinada por sus valores en $\mathbb{R}^n_+$

**Definición 2.2.2.** *Sea $x \in \mathbb{R}^n$ y supóngase que sus coordenadas son ordenadas de tal forma que $|x_1| \geq |x_2| \geq \cdots \geq |x_n|$. Para $k = 1, 2, ..., n$ definimos*

$$\Phi_{(k)}(x) = \sum_{j=1}^{k} |x_j|. \tag{2.2}$$

Es sencillo ver que es una norma y por las propiedades del módulo la función $\Phi_{(k)}$ es simétrica y calibrada; nos referiremos a esta función como la **norma de Ky Fan**. Además

$$\Phi_{(1)}(x) = \|x\|_\infty, \quad \Phi_{(n)}(x) = \|x\|_1.$$

Usaremos también $\|\cdot\|_{(k)}$ para denotar a la norma de Ky Fan, nótese el uso del paréntesis para no confundirlas con las $p-$normas.

**Proposición 2.2.3.**

*(i) Toda función simétrica y calibrada es continua.*

*(ii) Sea $\Phi$ una función simétrica y calibrada. Entonces para toda $x \in \mathbb{R}^n$, $\|x\|_\infty \leq \Phi(x) \leq \|x\|_1$.*

*(iii) Toda función $\Phi$ simétrica y calibrada es monótona en $\mathbb{R}^n_+$.*

*(iv) Sean $x, y \in \mathbb{R}^n$ tales que $|x| \leq |y|$, entonces $\Phi(x) \leq \Phi(y)$.*

*Demostración.*

*(i)* Dada una función $\Phi$ simétrica y calibrada, por $(i)$ de la Definición 2.2.1 $\Phi$ es una norma sobre $\mathbb{R}^n$. De aquí se obtiene la continuidad.

*(ii)* Sea $x \in \mathbb{R}^n$. Supongamos que $\|x\|_\infty = \text{máx}_{1 \leq j \leq n} |x_j| = x_k$, para alguna $k \in \{1, 2, \ldots, n\}$. Entonces de que $\Phi$ sea norma, la propiedad $(iv)$ de la Definición 2.2.1 y la Proposición 2.1.2 implican que

$$\|x\|_\infty = |x_k|\Phi(0, ..., 0, 1, 0, ..., 0) = \Phi(0, ..., 0, x_k, 0, ..., 0) \leq \Phi(x).$$

Por otro lado, nuevamente por las propiedades de una norma, la calibración y normalización tenemos que

$$\begin{aligned} \Phi(x) &= \Phi(x_1, ..., x_n) \leq \Phi(x_1, 0, ..., 0) + ... + \Phi(0, ..., x_n) \\ &= \Phi(|x_1|, ..., 0) + ... + \Phi(0, ..., |x_n|) \\ &= |x_1|\Phi(1, 0, \cdots, 0) + ... + |x_n|\Phi(0, \cdots, 1) = \|x\|_1, \end{aligned}$$

así tenemos $\|x\|_\infty \leq \Phi(x) \leq \|x\|_1$.

*(iii)* Sean $x, y \in \mathbb{R}^n_+$ tales que $x_j \leq y_j, \quad j = 1, \cdots n$, equivalentemente $|x| \leq |y|$. Así por la Proposición 2.1.2 se sigue el resultado.

*(iv)* Es consecuencia del inciso anterior. $\square$

Consideraremos funciones $\Phi : \mathbb{R}^n \to \mathbb{R}^m$, cuyo dominio es todo $\mathbb{R}^n$ o bien un subconjunto de este espacio que es convexo e invariante bajo la permutación de sus coordenadas. Retomando el concepto de la mayorización presentado en el capítulo anterior, ahora introducimos la siguiente

**Definición 2.2.4.** *Diremos que una función $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ es* ***isotona*** *si*

$$x \prec y \qquad \Rightarrow \qquad \Phi(x) \prec_\omega \Phi(y). \tag{2.3}$$

*y* ***fuertemente isotona*** *si*

$$x \prec_\omega y \qquad \Rightarrow \qquad \Phi(x) \prec_\omega \Phi(y). \tag{2.4}$$

*Si $m = 1$ la submayorización $\prec_\omega$ en el lado derecho de (2.3) se convierte en desigualdad. En este caso las funciones isotonas también se llaman Schur$-$* ***convexas*** *o bien $S-$**convexas**.*

**Definición 2.2.5.** *Una función $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ es **monótona creciente** si*

$$x \leq y \qquad \Rightarrow \quad \Phi(x) \leq \Phi(y).$$

**Definición 2.2.6.** *Una función $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ es **convexa** si para todos $x, y \in \mathbb{R}^n$*

$$\Phi(tx + (1-t)y) \leq t\Phi(x) + (1-t)\Phi(y) \qquad \text{para toda } 0 \leq t \leq 1.$$

**Definición 2.2.7.** *Una función real $f$ definida en un intervalo $I$ es convexa si dados $a_j \geq 0$ para $j = 1, 2, ..., n$ tales que $\sum_j a_j = 1$, entonces*

$$f\left(\sum_{j=1}^{n} a_j t_j\right) \leq \sum_{j=1}^{n} a_j f(t_j) \qquad \forall t_j \in I. \tag{2.5}$$

La función $f(t) = -\ln(t)$ es convexa, luego usando la desigualdad (2.5) y leyes de logaritmos tenemos

$$\ln\left(\sum_{j=1}^{n} a_j t_j\right) \geq \sum_{j=1}^{n} a_j \ln(t_j) = \sum_{j=1}^{n} \ln(t_j^{a_j}) = \ln\left(\prod_{j=1}^{n} t_j^{a_j}\right).$$

Aplicando la función exponencial obtenemos

$$\sum_{j=1}^{n} a_j t_j \geq \prod_{j=1}^{n} t_j^{a_j}. \tag{2.6}$$

Ésta es la desigualdad entre la media aritmética y la media geométrica.

**Teorema 2.2.8.** *(i) Sean $x, y$ vectores con coordenadas no negativas. Entonces $x \prec_\omega y$ si y sólo si $x = By$, para alguna matrix bisubestocástica $B$.*

*(ii) Sean $x, y \in \mathbb{R}^n$. Entonces $x \prec_\omega y$ si y sólo si existe un vector $u$ tal que $x \leq u$ y $u \prec y$.*

*Demostración.* Siguiendo la referencia [1]. Si $x, u \in \mathbb{R}^n$ y $x \leq u$, se sigue que $x \prec_\omega u$. Si además $u \prec y$, recordemos que la relación $\prec_\omega$ es transitiva, por lo tanto $x \prec_\omega y$.

Supongamos ahora que $x, y$ son vectores no negativos y $x = By$ para alguna matriz bisubestocástica $B$. Por la Proposición 1.2.17, podemos encontrar una matriz $A$ biestocástica tal que $b_{ij} \leq a_{ij}$ para todos $i, j$. Entonces

$x = By \le Ay$. Tomando $u = Ay$, por la Proposición 1.2.21 tenemos que $u \prec y$. Así, $x \prec_\omega y$.

Recíprocamente, sean $x, y$ vectores no negativos tales que $x \prec_\omega y$. Demostraremos que existe una matriz biestocástica $B$ tal que $x = By$. Si $x = 0$, podemos elegir $B = 0$, y si $x \prec y$ por el Teorema 1.2.21 podemos elegir $B$ biestocástica y el resultado se sigue. Supongamos que no pasa ninguno de los caso anteriores. Sea $r = \min\{x_1, \ldots, x_n\}$ y sea $s = \operatorname{Tr} y - \operatorname{Tr} x$. Tenemos que $s > 0$. Sea $m$ un entero positivo tal que $r \ge \frac{s}{m}$. Ahora, dilatamos los vectores $x, y$ en vectores en $u, v \in \mathbb{R}^{n+m}$, como

$$\begin{aligned} u &= (x_1, \ldots, x_n, s/m, \ldots, s/m) \\ v &= (y_1, \ldots, y_n, 0, \ldots, 0). \end{aligned}$$

Como $r \le x_j$ para todo $j = 1, \ldots, n$. Tenemos para algún $1 \le m_0 \le m$

$$\begin{aligned} \sum_{i=1}^{n+(m-m_0)} u_i^\downarrow &= \sum_{i=1}^{n} x_i^\downarrow + \sum_{i=n+1}^{n+(m-m_0)} \frac{s}{m} = \sum_{i=1}^{n} x_i^\downarrow + (m - m_0)\frac{\operatorname{Tr} y - \operatorname{Tr} x}{m} \\ &= \frac{m-m_0}{m}\sum_{i=1}^{n} y_j^\downarrow + \frac{m_0}{m}\sum_{i=1}^{n} u_i^\downarrow \le \frac{m-m_0}{m}\sum_{i=1}^{n} y_j^\downarrow + \frac{m_0}{m} \\ &= \sum_{i=1}^{n} y_i^\downarrow = \sum_{i=1}^{n+(m-m_0)} v_i^\downarrow. \end{aligned}$$

Por lo que $\sum_{i=1}^{k} u_i^\downarrow \le \sum_{i=1}^{k} v_i^\downarrow$ para toda $k = 1, \ldots, n+m$. Por otro lado

$$\sum_{i=1}^{n+m} u_i = \sum_{i=1}^{n} x_i + \sum_{i=n+1}^{n+m} \frac{s}{m} = \sum_{i=1}^{n} x_i + m\frac{\operatorname{Tr} y - \operatorname{Tr} x}{m} = \sum_{i=1}^{n+m} v_j.$$

Entonces, $u \prec v$. Así por el Teorema 1.2.21 existe una matriz $A$ biestocástica de tamaño $(n+m) \times (n+m)$ tal que $u = Av$. Luego, sea $B_{n\times n}$ submatriz de $A$ (situada en la esquina superior izquierda de $A$), por la Proposición 1.2.15, $B$ es bisubestocástica. En consecuencia $x = By$.

Sean $x, y \in \mathbb{R}^n$ tales que $x \prec_\omega y$. Escogemos $t > 0$ tal que $x + t\mathbf{e}, y + t\mathbf{e}$ son positivos, donde $\mathbf{e} = (1, 1, ..., 1)$. Nótese que seguimos teniendo $x + t\mathbf{e} \prec_\omega y + t\mathbf{e}$. Entonces, por lo mostrado anteriormente $x + t\mathbf{e} = B(y + t\mathbf{e})$, para alguna matriz bisubestocastica $B$. Luego, por la Proposición 1.2.17 existe una matriz biestocástica $A$ tal que $b_{ij} \le a_{ij}$ para todos $i, j = 1, \ldots, n$. Entonces $x + t\mathbf{e} = A(y + t\mathbf{e}) = Ay + t\mathbf{e}$ pues $A$ es biestocástica y preseva unidad. Entonces tomando $u = Ay$ tenemos $x \le u$ y $u \prec y$. $\square$

**Teorema 2.2.9.** *Sea $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ una función convexa. Supóngase que para toda $P \in S_n$ existe $P' \in S_m$ tal que*

$$\Phi(Px) = P'\Phi(x) \qquad \forall x \in \mathbb{R}^n. \tag{2.7}$$

*Entonces $\Phi$ es isotona. Si además $\Phi$ es monótona creciente entonces es fuertemente isotona.*

*Demostración.* Sean $x, y \in \mathbb{R}^n$ tales que $x \prec y$. Por la parte $(iii)$ del Teorema 1.2.21 existen matrices de permutación $P_1, ..., P_N \in S_n$ y números reales positivos $t_1, ..., t_N$ con $\sum_j t_j = 1$ tales que $x = \sum_j t_j P_j y$. Por la convexidad de $\Phi$ y la propiedad (2.7) tenemos que

$$\Phi(x) = \Phi\left(\sum_j t_j P_j y\right) \leq \sum_j t_j \Phi(P_j y) = \sum_j t_j P_j' \Phi(y)$$

Escribiendo $z = \sum_j t_j P_j' \Phi(y)$, se tiene que $\Phi(x) \leq z$ y por $(iii)$ del Teorema 1.2.21, $z \prec_\omega \Phi(y)$. Por lo tanto $\Phi(x) \prec_\omega \Phi(y)$.

Ahora, si $\Phi$ además es monótona creciente. Sean $x, y$ tales que $x \prec_\omega y$. Por el Teorema 2.2.8, existe $u \in \mathbb{R}^n$ tal que $x \leq u \prec y$. Así, por la monotonía tenemos que $\Phi(x) \leq \Phi(u)$ y $\Phi(u) \prec_\omega \Phi(y)$. Esto prueba que $\Phi$ es fuertemente isotona. □

**Proposición 2.2.10.** *Sean $x, y \in \mathbb{R}^n_+$. Si $x \prec_\omega y$ entonces $\Phi(x) \leq \Phi(y)$ para toda función $\Phi$ simétrica y calibrada.*

*Demostración.* Sea $\Phi$ función simétrica y calibrada. Por ser norma, $\Phi$ es convexa; luego se cumplen las condiciones del Teorema II.3.3 en [1]. Tenemos que $\Phi$ es fuertemente isotona. Por la definición 2.2.1 tenemos que la submayorización $\Phi(x) \prec_\omega \Phi(y)$ se vuelve la desigualdad, es decir $\Phi(x) \leq \Phi(y)$. □

**Proposición 2.2.11.** *Si para toda $k = 1, 2, ..., n$ y para toda $x, y \in \mathbb{R}^n$ $\Phi_{(k)}(x) \leq \Phi_{(k)}(y)$, entonces $\Phi(x) \leq \Phi(y)$ para toda función $\Phi$ simétrica y calibrada.*

*Demostración.* Supóngase que $\Phi_{(k)}(x) \leq \Phi_{(k)}(y)$ para toda $k = 1, 2, ..., n$ por la Definición 2.2.2 tenemos que $|x| \prec_\omega |y|$, entonces de la Proposición 2.2.10 se tiene que para toda $\Phi$ función simétrica y calibrada $\Phi(|x|) \leq \Phi(|y|)$. En consecuencia $\Phi(x) \leq \Phi(y)$ por la propiedad de calibración. □

**Proposición 2.2.12.** *Para cada* $k = 1, 2, ..., n$ *y cada* $x \in \mathbb{R}^n$

$$\Phi_{(k)}(x) = \text{mín}\left\{\|u\|_1 + k\|v\|_\infty : x = u + v\right\}. \tag{2.8}$$

*Demostración.* Supongamos que $x \in \mathbb{R}^n_+$ pues para calcular $\Phi_{(k)}(x)$ debemos tomar el orden en las coordenadas $|x_1| \geq |x_2| \geq \cdots \geq |x_n|$. Si $x = u + v$ entonces por las propiedades de norma y de la Definición 2.2.2

$$\begin{aligned}\Phi_{(k)}(x) =& \Phi_{(k)}(u+v) \leq \Phi_{(k)}(u) + \Phi_{(k)}(v) \leq \|u\|_1 + \Phi_{(k)}(v) \\ \leq& \|u\|_1 + k\|v\|_\infty.\end{aligned}$$

Ahora tomando

$$\begin{aligned}u =& (x_1^\downarrow - x_k^\downarrow, x_2^\downarrow - x_k^\downarrow, ..., x_k^\downarrow - x_k^\downarrow, 0, ..., 0) \\ v =& (x_k^\downarrow, x_k^\downarrow, ..., x_k^\downarrow, x_{k+1}^\downarrow, ..., x_n^\downarrow),\end{aligned}$$

tenemos que $u + v = x^\downarrow$ y

$$\begin{aligned}\|u\|_1 =& \sum_{j=1}^{n} \left|x_j^\downarrow - x_k^\downarrow\right| = \sum_{j=1}^{k} x_j^\downarrow - x_k^\downarrow = \Phi_{(k)}(x) - kx_k^\downarrow \\ \|v\|_\infty =& x_k^\downarrow,\end{aligned}$$

por lo tanto $\Phi_{(k)}(x) = \|u\|_1 + k\|v\|_\infty$ alcanzando el mínimo. Así $\Phi_{(k)}(x) = \text{mín}\{\|u\|_1 + k\|v\|_\infty : x = u + v\}$. □

**Teorema 2.2.13.** *Sean* $p, q$ *números reales tales que* $p > 1$ *y* $\frac{1}{p} + \frac{1}{q} = 1$. *Entonces para toda función simétrica y calibrada* $\Phi$ *y* $x, y \in \mathbb{R}^n$ *se cumple*

$$\Phi(|x \cdot y|) \leq \left[\Phi(|x|^p)\right]^{1/p} \left[\Phi(|y|^q)\right]^{1/q}. \tag{2.9}$$

*Demostración.* Para cada $j = 1, 2, \ldots, n$ tomando $t_1 = |x_j|^p$, $t_2 = |y_j|^q$, $a_1 = 1/p$, $a_2 = 1/q$; por la desigualdad (2.6) se tiene que

$$|x_j \cdot y_j| \leq \frac{|x_j|^p}{p} + \frac{|y_j|^q}{q} \qquad\qquad \forall\, j = 1, 2, ..., n.$$

Como $\Phi$ es función monótona y por las propiedades de norma tenemos

$$\Phi(|x \cdot y|) \leq \Phi\left(\frac{|x|^p}{p} + \frac{|y|^q}{q}\right) \leq \frac{\Phi(|x|^p)}{p} + \frac{\Phi(|y|^q)}{q}.$$

Ahora, para toda $t > 0$ reemplazamos a $x$ por $tx$ y $y$ por $t^{-1}y$, tenemos que

$$\Phi(|x \cdot y|) \leq \frac{\Phi(|tx|^p)}{p} + \frac{\Phi(|t^{-1}y|^q)}{q} = \frac{t^p}{p}\Phi(|x|^p) + \frac{t^{-q}}{q}\Phi(|y|^q).$$

Veamos que

$$[\Phi(|x|^p)]^{1/p}\,[\Phi(|y|^q)]^{1/q} = \min_{t>0}\left\{\frac{t^p}{p}\Phi(|x|^p) + \frac{t^{-q}}{q}\Phi(|y|^q)\right\}.$$

Sea $\varphi(t) = a\frac{t^p}{p} + b\frac{t^{-q}}{q}$, con $a, b, t > 0$, entonces

$$\varphi'(t) = at^{p-1} - bt^{-q-1}.$$

Entonces el punto crítico de esta función es $t = (a^{-1}b)^{1/pq} = a^{-1/pq}\,b^{1/pq}$. Como $\varphi''(t) = a(p-1)t^{p-2} + b(q+1)t^{-q-2} > 0$ para toda $t > 0$, tenemos que en $t = a^{-1/pq}\,b^{1/pq}$ la función tiene un mínimo. Evaluando en la función original obtenemos

$$\begin{aligned}\varphi(a^{-1/pq}\,b^{1/pq}) &= a\frac{(a^{-1/pq}\,b^{1/pq})^p}{p} + b\frac{(a^{-1/pq}\,b^{1/pq})^{-q}}{q}\\ &= \frac{a^{1/p}b^{1/q}}{p} + \frac{a^{1/p}b^{1/q}}{q} = a^{1/p}\,b^{1/q}.\end{aligned}$$

Por lo tanto si $a = \Phi(|x|^p)$, $b = \Phi(|y|^q)$ tenemos que

$$\Phi(|x \cdot y|) \leq \min_{t>0}\left\{\frac{t^p}{p}\Phi(|tx|^p) + \frac{t^{-q}}{q}\Phi(|t^{-1}y|^q)\right\} = (\Phi(|x|^p))^{1/p}\,(\Phi(|y|^q))^{1/q}.$$

$\square$

A la desigualdad (2.9) le llamaremos desigualdad de Hölder para funciones simétricas y calibradas. Si tomamos $p = 2$ obtenemos lo que llamaremos la desigualdad de Cauchy-Schwarz para funciones simétricas y calibradas

$$\Phi(|x \cdot y|) \leq \left(\Phi(|x|^2)\right)^{1/2}\left(\Phi(|y|^2)\right)^{1/2}.$$

Si $\Phi = \Phi_{(n)}$ la relación en (2.9) se vuelve la desigualdad de Hölder ya conocida para la $p$-norma en $\mathbb{C}^n$:

$$\sum_{j=1}^{n}|x_j\,y_j| \leq \left(\sum_{j=1}^{n}|x_j|^p\right)^{1/p}\left(\sum_{j=1}^{n}|y_j|^q\right)^{1/q}.$$

**Proposición 2.2.14.** *Sean* $p, q, r \in \mathbb{R}_+$ *tales que* $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$. *Entonces para cada función simétrica y calibrada* $\Phi$ *y cada* $x$, $y \in \mathbb{R}^n$ *se cumple*

$$[\Phi(|x \cdot y|^r)]^{1/r} \leq [\Phi(|x|^p)]^{1/p} [\Phi(|y|^q)]^{1/q}. \tag{2.10}$$

*Demostración.* Sean $t_1 = |x_j|^p$, $t_2 = |y_j|^q$, $a_1 = r/p$, $a_2 = r/q$ para $j = 1, 2, \ldots, n$, por la desigualdad (2.6) tenemos

$$|x_j \cdot y_j|^r = (|x_j|^p)^{r/p} (|y_j|^q)^{r/q} \leq \frac{r}{p}|x_j|^p + \frac{r}{q}|y_j|^q \qquad \forall\, j = 1, ..., n.$$

Por la monotonía de $\Phi$ y las propiedades de norma

$$\Phi(|x \cdot y|^r) \leq \frac{r}{p}\Phi(|x|^p) + \frac{r}{q}\Phi(|y|^q).$$

Ahora cambiando $x \to tx$, $y \to t^{-1}y$, tenemos

$$\Phi(|x \cdot y|^r) \leq \frac{rt^p}{p}\Phi(|x|^p) + \frac{rt^{-q}}{q}\Phi(|y|^q) \qquad \text{para todo } t \in \mathbb{R}_+.$$

Sea $\varphi(t) = \frac{rt^p}{p}a + \frac{rt^{-q}}{q}b$, entonces $\varphi'(t) = rt^{p-1}a - rt^{-q-1}b$ y tenemos que $t = (a^{-1}b)^{r/pq}$ es el único punto crítico, de hecho la función tiene un mínimo global en este punto, pues $\varphi''(t) = r(p-1)t^{p-2} + r(q+1)t^{-(q+2)}b > 0$ para toda $t \in \mathbb{R}_+$ . Evaluando la función original en el punto crítico obtenemos

$$\varphi((a^{-1}b)^{r/pq}) = \frac{r((a^{-1}b)^{r/pq})^p}{p}a + \frac{r((a^{-1}b)^{r/pq})^{-q}}{q}b = a^{r/p}\, b^{r/q},$$

es decir $\min_{t>0}\{\frac{rt^p}{p}a + \frac{rt^{-q}}{q}b\} = a^{r/p}\, b^{r/q}$. Así, tomando $a = \Phi(|x|^p)$ y $b = \Phi(|y|^q)$ se sigue el resultado. □

**Teorema 2.2.15 (Desigualdad de Minkowski).** *Sea* $\Phi$ *una función simétrica y calibrada, sea* $p \geq 1$, *entonces para toda* $x, y \in \mathbb{R}^n$ *se cumple*

$$[\Phi(|x + y|^p)]^{1/p} \leq [\Phi(|x|^p)]^{1/p} + [\Phi(|y|^p)]^{1/p}. \tag{2.11}$$

*Demostración.* Sea $p = 1$, por las propiedades de $|\cdot|$ y de norma para $\Phi$

$$\Phi(|x + y|) \leq \Phi(|x| + |y|) \leq \Phi(|x|) + \Phi(|y|).$$

Sean $p>1$ y $q>0$ tal que $\frac{1}{p}+\frac{1}{q}=1$. Sean $x,y\in\mathbb{R}^n$, entonces para cada $j=1,2,\ldots,n$; tenemos que

$$\begin{aligned}|x_j+y_j|^p =&|x_j+y_j||x_j+y_j|^{p-1}\\ \leq&|x_j||x_j+y_j|^{p-1}+|y_j||x_j+y_j|^{p-1},\end{aligned}$$

es decir $|x+y|^p\leq|x||x+y|^{p-1}+|y||x+y|^{p-1}$. Así para $\Phi$ función simétrica y calibrada, por la Proposición 2.2.3 y la desigualdad (2.9) tenemos

$$\begin{aligned}\Phi(|x+y|^p)&\leq\Phi(|x||x+y|^{p-1})+\Phi(|y||x+y|^{p-1})\\ &\leq[\Phi(|x|^p)]^{1/p}\,[\Phi(|x+y|^{q(p-1)})]^{1/q}+[\Phi(|y|^p)]^{1/p}\,[\Phi(|x+\quad y|^{q(p-1)})]^{1/q}\\ &=\left\{[\Phi(|x|^p)]^{1/p}+[\Phi(|y|^p)]^{1/p}\right\}(\Phi(|x+y|^p))^{1/q}.\end{aligned}$$

Luego, dividiendo entre $(\Phi(|x+y|^p))^{1/q}$ y como $\frac{1}{p}=1-\frac{1}{q}$ tenemos

$$(\Phi(|x+y|^p))^{1-1/q}=\frac{\Phi(|x+y|^p)}{(\Phi(|x+y|^p))^{1/q}}\leq[\Phi(|x|^p)]^{1/p}+[\Phi(|y|^p)]^{1/p}. \qquad \square$$

Nótese que si $\Phi=\Phi_{(n)}$ tenemos

$$\left(\sum_{j=1}^{n}|x_j+y_j|^p\right)^{1/p}\leq\left(\sum_{j=1}^{n}|x_j|^p\right)^{1/p}+\left(\sum_{j=1}^{n}|y_j|^p\right)^{1/p},$$

la cual es la desigualdad de Minkowski para $p$-normas sobre espacios de dimensión finita.

**Proposición 2.2.16.** *Sea $\Phi$ una función simétrica y calibrada, definamos $\Phi^{(p)}(x):=[\Phi(|x|^p)]^{1/p}$; entonces $\Phi^{(p)}$ también es una función simétrica y calibrada.*

*Demostración.*

($i$) Veamos que $\Phi^{(p)}$ es norma. Por ser $\Phi$ simétrica y calibrada cumple con el Teorema 2.2.15. Entonces $\Phi^{(p)}$ satisface la desigualdad del triángulo. Que $\Phi^{(p)}(\lambda x)=\lambda\Phi^{(p)}(x)$ y que $\Phi^{(p)}\geq 0$ se demuestran directamente.

($ii$) Sea $P\in S_n$, entonces $|Px|^p=(|x_{\sigma(1)}|^p,...,|x_{\sigma(n)}|^p)=P|x|^p$. Por la simetría de $\Phi$ tenemos

$$\Phi^{(p)}(Px)=[\Phi(|Px|^p)]^{1/p}=[\Phi(P|x|^p)]^{1/p}=[\Phi(|x|^p)]^{1/p}=\Phi^{(p)}(x).$$

$(iii)$ Sean $\xi_j = \pm 1$ para $j = 1, 2, ..., n$. Entonces

$$\begin{aligned}\Phi^{(p)}(\xi_1 x_1, ..., \xi_n x_n) =& [\Phi(|\xi_1 x_1|^p, ..., |\xi_n x_n|^p)]^{1/p}\\ =& [\Phi(|x_1|^p, ..., |x_n|^p)]^{1/p} = \Phi^{(p)}(x).\end{aligned}$$

$(iv)$ Calculamos:

$$\Phi^{(p)}(1, 0, ..., 0) = [\Phi(|(1, 0, ..., 0)|^p)]^{1/p} = [\Phi(|1|^p, 0, ..., 0)]^{1/p} = 1. \quad \square$$

**Proposición 2.2.17.** *Sean $r,\ p \geq 1$ y $\Phi_p$ la $p-$norma en $\mathbb{C}^n$, entonces $\Phi_p^{(r)} = \Phi_{p\,r}$.*

*Demostración.* Sea $x \in \mathbb{C}^n$, calculamos:

$$\Phi_p^{(r)}(x) = [\Phi_p(|x|^r)]^{1/r} = \left[\left(\sum_j |\,|x_j|^r|^p\right)^{1/p}\right]^{1/r} = \left[\sum_j (|x_j|^{pr})\right]^{1/pr} = \Phi_{pr}(x).$$

$\square$

**Definición 2.2.18.** *$\Psi$ se llama función **simétrica calibrada cuadrática** o simplemente **Q-norma** si para alguna función simétrica y calibrada $\Phi$ se tiene que $\Psi = \Phi^{(2)}$. Es decir*

$$\Psi(x) = \left[\Phi(|x|^2)\right]^{1/2}. \tag{2.12}$$

**Proposición 2.2.19.** $(i)$ *Una $p$-norma es una $Q$-norma si y sólo si $p \geq 2$.*

$(ii)$ *Para cada $k = 1, 2, ..., n$; $\Phi_{(k)}^{(p)}(x)$ es una $Q$-norma sí y solo si $p \geq 2$.*

*Demostración.*

*(i)* Sea $\Phi_p$ la $p$-norma en $\mathbb{C}^n$. Si $p \geq 2$, entonces $\Phi_{\frac{p}{2}}$ es una norma simétrica y calibrada; y tenemos que

$$\Phi_{\frac{p}{2}}\left(|x|^2\right)^{\frac{1}{2}} = \Phi_p(x).$$

Entonces $\Phi_p$ es una $Q$-norma.

Recíprocamente, si $\Phi_p(x) = \Phi(|x|^2)^{\frac{1}{2}}$ con $\Phi$ una función simétrica y calibrada, entonces

$$\Phi(|x|^2)^{\frac{1}{2}} = \Phi_p(x) = \Phi_{\frac{p}{2}}\left(|x|^2\right)^{\frac{1}{2}}.$$

Esto implica que $\Phi\big(|x|^2\big) = \Phi_{\frac{p}{2}}\big(|x|^2\big)$ para toda $x \in \mathbb{C}^n$. Pero una función simétrica y calibrada está completamente determinada por sus valores en $\mathbb{R}_+^n$. Entonces $\Phi = \Phi_{\frac{p}{2}}$ y como $\Phi_q$ no define una norma para $0 < q < 1$, se sigue que $p \geq 2$.

*(ii)* Nótese que

$$\Phi_{(k)}^{(p)}(x) = \left(\sum_{j=1}^{k} |x_j|^p\right)^{1/p} \qquad \text{para}\, k = 1, 2, ..., n,$$

es una $p$-norma para $k = 1, n$. El resultado se sigue por un argumento semejante al del inciso anterior. $\square$

**Definición 2.2.20.** *Para una norma $\Phi$ en $\mathbb{C}^n$, su función dual de $\Phi$ se define como*

$$\Phi'(x) = \sup_{\Phi(y)=1} |\langle x, y\rangle|. \tag{2.13}$$

**Proposición 2.2.21.** *Si $\Phi$ es una norma, entonces su función dual $\Phi'$ es una norma.*

*Demostración.* De la definición 2.2.20 se tiene $\Phi'(x) = \sup_{\Phi(y)=1} |\langle x, y\rangle| \geq |\langle x, y\rangle| \geq 0$. Sea $\lambda \in \mathbb{C}$, calculamos

$$\Phi'(\lambda x) = \sup_{\Phi(y)=1} |\langle \lambda x, y\rangle| = |\lambda| \sup_{\Phi(y)=1} |\langle x, y\rangle| = |\lambda|\Phi'(x).$$

Por las propiedades del supremo y del módulo tenemos

$$\begin{aligned}\Phi'(x+y) = \sup_{\Phi(u)=1} |\langle x+y, u\rangle| &= \sup_{\Phi(u)=1} |\langle x, u\rangle + \langle y, u\rangle| \\ &\leq \sup_{\Phi(u)=1} \{|\langle x, u\rangle| + |\langle y, u\rangle|\} \leq \sup_{\Phi(u)=1} |\langle x, u\rangle| + \sup_{\Phi(u)=1} |\langle y, u\rangle| \\ &= \Phi'(x) + \Phi'(y)\end{aligned}$$

$\square$

Nótese que por las propiedades del supremo $\Phi'$ sigue siendo una norma incluso cuando $\Phi$ sea una función sobre $\mathbb{C}^n$ que no cumple la desigualdad del triángulo.

**Proposición 2.2.22.** *Si $\Phi$ es una función simétrica y calibrada entonces $\Phi'$ también lo es.*

*Demostración.* Por la proposición anterior sólo basta ver que cumple las propiedades de simetría, calibración y normalización. Sea $P \in S_n$, sabemos que $P^{-1} \in S_n$, entonces

$$\begin{aligned}\Phi'(Px) = \sup_{\Phi(y)=1} |\langle Px, y\rangle| &= \sup_{\Phi(y)=1} |\langle x, P^*y\rangle| = \sup_{\Phi((P^*)^{-1}z)=1} |\langle x, z\rangle| \\ &= \sup_{\Phi(z)=1} |\langle x, z\rangle| = \Phi'(x).\end{aligned}$$

Sean $\xi_j = \pm 1$, para $j = 1, 2, ..., n$, denotemos por $\xi$ a la matriz diag $(\xi_1, ..., \xi_n)$. Entonces

$$\begin{aligned}\Phi'(\xi_1 x_1, ..., \xi_n x_n) &= \sup_{\Phi(y)=1} |\langle \xi x, y\rangle| = \sup_{\Phi(y)=1} |\langle x, \xi^* y\rangle| \\ &= \sup_{\Phi(\xi^{-1}z)=1} |\langle x, z\rangle| = \sup_{\Phi(z)=1} |\langle x, z\rangle| = \Phi'(x).\end{aligned}$$

Sea $\mathbf{e}_1 = (1, 0, ..., 0)$, entonces

$$\Phi'(\mathbf{e}_1) = \sup_{\Phi(y)=1} |\langle \mathbf{e}_1, y\rangle| = \sup_{\Phi(y)=1} |y_1| = 1.$$

Pues por la Proposición 2.2.3 inciso *(ii)* se tiene que $|y_1| \leq \|y\|_\infty \leq \Phi(y)$ y estamos tomando $\Phi(y) = 1$; de donde $\sup_{\Phi(y)=1} |y_1| \leq 1$. Por otra parte como $\Phi(\mathbf{e}_1) = 1$ el supremo se alcanza en $y_1 = 1$. □

**Proposición 2.2.23.** *Dada una norma $\Phi$ en $\mathbb{C}^n$ se cumple*

$$|\langle x, y\rangle| \leq \Phi(x)\Phi'(y) \qquad \textit{para toda } x, y \in \mathbb{C}^n. \tag{2.14}$$

*Demostración.* Sea $z \in B_\Phi = \{x : \Phi(x) = 1\}$; de la definición 2.2.20 tenemos $\Phi(z)\Phi'(x) = \sup_{\Phi(z)=1} |\langle x, z\rangle| \geq |\langle x, z\rangle|$. Sea $z = \frac{x}{\Phi(x)}$ entonces $z \in B_\Phi$, por lo tanto

$$\left|\left\langle y, \frac{x}{\Phi(x)}\right\rangle\right| \leq \sup_{\Phi(z)=1} |\langle y, z\rangle| = \Phi'(y).$$

Entonces, por la linealidad del producto interior tenemos

$$\frac{1}{\Phi(x)}|\langle y, x\rangle| = \frac{1}{\Phi(x)}|\langle x, y\rangle| \leq \Phi'(y).$$

Multiplicando por $\Phi(x)$ se sigue el resultado, siempre que $\Phi(x) \neq 0$. Para el caso $\Phi(x) = 0$, la relación (2.14) se cumple con igualdad pues por las propiedades de norma de $\Phi$ se tiene que $\Phi(x) = 0$ si y sólo si $x = 0$. □

**Proposición 2.2.24.** *Sean $\Phi_p$ una $p$-norma, $p, q \in \mathbb{R}$ tales que $1 \leq p \leq \infty$ y $\frac{1}{p} + \frac{1}{q} = 1$, entonces $\Phi'_p = \Phi_q$*

*Demostración.* Consideremos $(\mathbb{C}^n, \|\cdot\|_p)^* = (\mathbb{C}^n, \|\cdot\|_q)$. Sea $\rho : \mathbb{C}^n \to \mathbb{C}$ un funcional lineal acotado; por el Lema de Riesz existe un único $y \in \mathbb{C}^n$ tal que $\rho(x) = \langle x, y\rangle$. Por otra parte tenemos

$$\|y\|_q = \|\rho\| = \sup_{\|x\|_p=1} |\rho(x)| = \sup_{\Phi_p(x)=1} |\langle x, y\rangle| = \Phi'_p(y).$$ □

**Proposición 2.2.25.** *Sean $\Phi\,,\Psi$ normas en $\mathbb{C}^n$. Si*

$$\Phi(x) \leq c\Psi(x) \qquad\qquad \forall\, x \in \mathbb{C}^n \text{ y alguna } c > 0, \tag{2.15}$$

*entonces*

$$\Psi'(x) \leq c\Phi'(x) \qquad\qquad \forall\, x \in \mathbb{C}^n. \tag{2.16}$$

*Demostración.* Si $\Psi(x) = 1$, entonces $\Phi(x) \leq c$. Esto demuestra que $B_\Psi \subset B_\Phi$, con $B_\Phi = \{x : \Phi(x) \leq c\}$ y $B_\Psi = \{x : \Psi(x) = 1\}$. Luego

$$\begin{aligned}\Psi'(x) &= \sup_{\Psi(y)=1} |\langle x, y\rangle| \leq \sup_{\Phi(y)\leq c} |\langle x, y\rangle| = \sup_{\Phi(c^{-1}y)\leq 1} |\langle x, y\rangle| \\ &= \sup_{\Phi(z)\leq 1} |\langle x, cz\rangle| = c \sup_{\Phi(z)\leq 1} |\langle x, z\rangle| = c\Phi'(x).\end{aligned}$$ □

**Definición 2.2.26.** *Una función $\Phi$ simétrica y calibrada es una $Q'$-norma si es el dual de una $Q$-norma.*

Por las proposiciones 2.2.19 y 2.2.24 tenemos que para $1 \leq p \leq 2$ las $p$-normas son ejemplos de $Q'$-normas.

**Proposición 2.2.27.**

*(i) Si $\Phi$ es una norma tal que $\Phi = \Phi'$, entonces $\Phi$ es la norma euclidiana.*

*(ii) Si $\Phi$ es una $Q$-norma tal que es $Q'$-norma también, entonces $\Phi$ es la norma euclidiana.*

*Demostración.*

*(i)* Como $\Phi$ es norma, se cumple la Proposición 2.2.23, es decir

$$|\langle x, y\rangle| \leq \Phi(x)\Phi'(y), \text{ para toda } x, y \in \mathbb{C}^n.$$

Tomando $y = x$ y del hecho que $\Phi = \Phi'$ tenemos que $(\Phi_2(x))^2 = |\langle x, x\rangle| \leq \Phi(x)\Phi(x) = \Phi(x)^2$, de donde $\Phi_2(x) \leq \Phi(x)$. De la Proposición 2.2.25, tomando $c = 1$ se tiene que $\Phi(x) = \Phi'(x) \leq \Phi_2'(x) = \Phi_2(x)$, se obtiene así la otra desigualdad.

*(ii)* Como $\Phi$ es $Q-$norma entonces $\Phi = \varphi^{(2)}$, para una función simétrica y calibrada $\varphi$. Como $\Phi$ es $Q'$-norma entonces $\Phi = (\psi^{(2)})'$, para una función

simétrica y calibrada $\psi$. En cualquier caso tenemos funciones simétricas y calibradas, entonces por la Proposición 2.2.3 se tiene que

$$\Phi(x) = [\varphi(|x|^2)]^{1/2} = \varphi^{(2)}(x) \leq \Phi_1^{(2)}(x) = \Phi_2(x).$$

Por otro lado, del hecho que $\psi^{(2)}(x) \leq \Phi_1^{(2)}(x) = \Phi_2(x)$, por la Proposición 2.2.25 tomando $c = 1$ tenemos que $\Phi_2(x) \leq (\psi^{(2)}(x))' = \Phi(x)$. Por lo tanto $\Phi$ es la norma euclideana. □

**Proposición 2.2.28.** *Para cada $k = 1, 2, \ldots, n$ la función dual de la norma de Ky Fan está dada por*

$$\Phi'_{(k)}(x) = \text{máx}\left\{\|x\|_\infty, \frac{1}{k}\|x\|_1\right\} \qquad \textit{para toda } x \in \mathbb{R}^n. \tag{2.17}$$

*Demostración.* Como $\Phi_{(k)}$ es simétrica y calibrada, por la Proposición 2.2.22 $\Phi'_{(k)}$ también lo es. Así, por la Proposición 2.2.3 se tiene que

$$\|x\|_\infty \leq \Phi'_{(k)}(x) \leq \|x\|_1.$$

Por otro lado, usando la Definición 2.2.2 se cumple que $\Phi_{(k)}(x) \leq k\|x\|_\infty$, entonces de la Proposición 2.2.25 se tiene que $\Phi'_{(k)}(x) \geq k^{-1}(\|x\|_\infty)' = k^{-1}\|x\|_1$. Entonces

$$\|x\|_\infty \leq \Phi'_{(k)}(x) \qquad \text{y} \qquad k^{-1}\|x\|_1 \leq \Phi'_{(k)}(x),$$

de donde $\text{máx}\left\{\|x\|_\infty, \frac{1}{k}\|x\|_1\right\} \leq \Phi'_{(k)}(x)$.

Ahora, supongamos que

$$\|x\|_\infty \leq \frac{1}{k}\|x\|_1 \qquad \text{para toda } 1 \leq k \leq n, \tag{2.18}$$

implica que

$$\|x\|_\infty \leq \frac{1}{k}\Phi_{(k)}(x) \qquad \text{para toda } 1 \leq k \leq n. \tag{2.19}$$

Tenemos así que: $(i)$ $\text{máx}\{\|x\|_\infty, \frac{1}{k}\|x\|_1\} = \frac{1}{k}\|x\|_1$, para toda $1 \leq k \leq n$; y $(ii)$ la desigualdad $\|x\|_\infty \leq \frac{1}{k}\|x\|_1$, para toda $1 \leq k \leq n$ junto con la Proposición 2.2.25 implican que $\Phi'_{(k)}(x) \leq \frac{1}{k}(\|x\|_\infty)' = \frac{1}{k}\|x\|_1$ para toda $1 \leq k \leq n$. De esto se sigue que

$$\Phi'_{(k)}(x) \leq \frac{1}{k}\|x\|_1 = \text{máx}\{\|x\|_\infty, \frac{1}{k}\|x\|_1\} \qquad \text{para toda } 1 \leq k \leq n.$$

Resta demostrar que (2.18) implica (2.19). Procederemos por inducción sobre $n$, la dimensión del espacio. Para $n = 1$, no hay nada que demostrar. Supongamos que (2.18) implica (2.19), vale para toda $n$ y veamos que se cumple para $(n+1)$. Supongamos

$$k|x_1| \leq \|x\|_1 = \sum_{j=1}^{n+1} |x_j|, \qquad \text{para toda } 1 \leq k \leq n+1.$$

Luego

$$(k-1)|x_1| \leq \sum_{j=2}^{n+1} |x_j| = \sum_{j=1}^{n} |x_{j+1}|, \qquad \text{para toda } 1 \leq (k-1) \leq n.$$

Usando la hipótesis de inducción obtenemos que

$$(k-1)|x_1| \leq \sum_{j=1}^{k-1} |x_{j+1}| \qquad \text{para toda } 1 \leq k \leq n.$$

Sumando $|x_1|$ en ambos lados se obtiene

$$k|x_1| \leq \sum_{j'=0}^{k-1} |x_{j'+1}| = \sum_{j=1}^{k} |x_j| = \Phi_{(k)}(x) \qquad \text{para toda } 1 \leq k \leq n+1.$$

Como se quería. □

**Proposición 2.2.29.** *Consideremos* $1 = \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n \geq 0$. *Dada una función* $\Phi$ *simétrica y calibrada en* $\mathbb{R}^n$*, la función*

$$\Psi(x) = \Phi(\alpha_1 |x_1|^{\downarrow}, \ldots, \alpha_n |x_n|^{\downarrow}) \tag{2.20}$$

*es simétrica y calibrada.*

*Demostración.*

*(i)* $\Psi(x) = \Phi(\alpha_1 |x_1|^{\downarrow}, \ldots, \alpha_n |x_n|^{\downarrow}) \geq 0$ para toda $x \in \mathbb{R}^n$.
Sea $\lambda \in \mathbb{R}$, entonces

$$\begin{aligned}\Psi(\lambda x) &= \Phi(\alpha_1 |\lambda x_1|^{\downarrow}, \ldots, \alpha_n |\lambda x_n|^{\downarrow}) = \Phi(|\lambda| \alpha_1 |x_1|^{\downarrow}, \ldots, |\lambda| \alpha_n |x_n|^{\downarrow}) \\ &= |\lambda| \Phi(\alpha_1 |x_1|^{\downarrow}, \ldots, \alpha_n |x_n|^{\downarrow}) = |\lambda| \Psi(x).\end{aligned}$$

Sean $x, y \in \mathbb{R}^n$; por las propiedades de monotonía y de norma de $\Phi$ se tiene que

$$\begin{aligned}\Psi(x+y) =&\Phi(\alpha_1|x_1+y_1|^{\downarrow}, \dots, \alpha_n|x_n+y_n|^{\downarrow})\\ \leq&\Phi(\alpha_1|x_1|^{\downarrow}+\alpha_1|y_1|^{\downarrow}, \dots, \alpha_n|x_n|^{\downarrow}+\alpha_n|y_n|^{\downarrow})\\ \leq&\Phi(\alpha_1|x_1|^{\downarrow}, \dots, \alpha_n|x_n|^{\downarrow})+\Phi(\alpha_1|y_1|^{\downarrow}, \dots, \alpha_n|y_n|^{\downarrow})\\ =&\Psi(x)+\Psi(y).\end{aligned}$$

Esto demuestra que $\Psi$ es una norma.

*(ii)* Sea $P \in S_n$ y entonces

$$\Psi(Px) = \Phi[P(\alpha_1|x_1|^{\downarrow}, \dots, \alpha_n|x_n|^{\downarrow})] = \Phi(\alpha_1|x_1|^{\downarrow}, \dots, \alpha_n|x_n|^{\downarrow}) = \Psi(x).$$

Esto demuestra que $\Psi$ es simétrica.

*(iii)* $\Psi$ es calibrada. Sea $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ donde $\xi_k = \pm 1$ para $k = 1, 2, \dots, n$, entonces

$$\Psi(\xi x) = \Phi(\alpha_1|\xi_1 x_1|^{\downarrow}, \dots, \alpha_n|\xi_n x_n|^{\downarrow}) = \Phi(\alpha_1|x_1|^{\downarrow}, \dots, \alpha_n|x_n|^{\downarrow}) = \Psi(x)$$

*(iv)* $\Psi$ es normalizada. Sea $\mathbf{1} = (1, 0, \dots, 0)$ y sabemos que $\alpha_1 = 1$

$$\begin{aligned}\Psi(\mathbf{e}_1) &= \Psi(1, 0, \dots, 0) = \Phi(\alpha_1|1|, 0, \dots, 0) = \alpha_1\Phi(1, 0, \dots, 0)\\ &= \Phi(1, 0, \dots, 0) = 1.\end{aligned}$$

□

## 2.3. Normas en $\mathbb{C}^n$ unitariamente invariantes

En esta sección $\mathbb{C}^n$ denotará el espacio de Hilbert $\mathbb{C}^n$ dotado del producto interno $\langle \cdot, \cdot \rangle$ y con la norma asociada $\|\cdot\|$. Si $A$ es un operador lineal sobre $\mathbb{C}^n$ denotaremos por $\|A\|$ a la norma de dicho operador, la cual se define como

$$\|A\| = \sup_{\|x\|=1} \|Ax\|. \tag{2.21}$$

Para cualquier operador $A$ sobre $\mathbb{C}^n$, su módulo se define por $|A| = (A^*A)^{1/2}$, donde $(A^*A)^{1/2}$ es la raíz cuadrada positiva del operador positivo $A^*A$.

**Proposición 2.3.1.** *Si $U$, $V$ son operadores unitarios sobre $\mathbb{C}^n$ entonces*

$$|U\,A\,V| = V^*|A|V. \tag{2.22}$$

*y $s(A) = s(UAV)$. Además*

$$\|A\| = \|U\,A\,V\|. \tag{2.23}$$

*Demostración.* Sean $U$ *y* $V$ operadores unitarios, entonces

$$\begin{aligned}|U\,A\,V|^2 =&(U\,A\,V)^*(U\,A\,V) = (V^*\,A^*\,U^*)(U\,A\,V) = V^*\,A^*\,Id\,A\,V\\ =&V^*\,|A|^2\,V = V^*\,|A|\,VV^*\,|A|\,V = (V^*\,|A|\,V)^2.\end{aligned}$$

Como tenemos operadores positivos se sigue que $|U\,A\,V| = V^*|A|\,V$. Ahora, sea $\lambda$ valor propio de $|UAV|$ con vector propio $x$, de la relación (2.22) tenemos

$$V^*|A|V\,x = |U\,A\,V|x = \lambda x,$$

como $V$ es unitario se sigue que $\lambda$ es valor propio de $|A|$ con vector propio asociado $y = Vx$.

Para $x \in \mathbb{C}^n$, de la relación (2.22) resulta que

$$\begin{aligned}\|U\,A\,Vx\|^2 &= \|V^*|A|\,Vx\|^2 = \langle V^*|A|\,Vx, V^*|A|\,Vx\rangle\\ &= \langle V\,V^*|A|\,Vx, |A|\,Vx\rangle = \langle |A|\,Vx, |A|Vx\rangle.\end{aligned}$$

Como $V$ es unitario se tiene que $\|Vx\| = \|x\|$; además $\|\,|A|\,\| = \|A\|$, por lo tanto

$$\|A\| =\|U\,A\,V\|. \qquad \square$$

Una norma en los operadores sobre $\mathbb{C}^n$ que satisface la propiedad (2.23) se llama **unitariamente invariante**. Denotaremos a tales normas mediante $|||\cdot|||$ y las normalizaremos de manera que $|||\mathrm{diag}\,(1,0,\cdots,0)||| = 1$.

El siguiente Teorema relaciona las normas unitariamente invariantes con las funciones simétricas y calibradas, mediante los valores singulares.

**Teorema 2.3.2.** *(i) Sea $\Phi$ una función simétrica y calibrada sobre $\mathbb{R}^n$. Definimos la función $|||\cdot|||_\Phi$ sobre $\mathbf{M}_n$ como sigue*

$$|||A|||_\Phi = \Phi(s(A)). \tag{2.24}$$

*Entonces $|||\cdot|||_\Phi$ es una norma unitariamente invariante.*

*(ii) Sea $|||\cdot|||$ una norma unitariamente invariante. Definimos la función $\Phi_{|||\cdot|||}$ sobre $\mathbb{R}^n$ mediante*

$$\Phi_{|||\cdot|||}(x) = |||\mathrm{diag}\,(x)|||. \tag{2.25}$$

*Donde* $\mathrm{diag}\,(x)$ *es la matriz diagonal con entradas $x_1, x_2, \ldots, x_n$. Entonces $\Phi_{|||\cdot|||}$ es una función simétrica y calibrada.*

*Demostración.*

*(i)* De la proposición anterior tenemos que $s(U\,A\,V) = s(A)$. Entonces

$$|||U\,A\,V|||_\Phi = \Phi(s(U\,A\,V)) = \Phi(s(A)) = |||A|||_\Phi.$$

Como $\Phi$ es función simétrica y calibrada, entonces para toda $A \in \mathbf{M}_n$ se tiene que $|||A|||_\Phi = \Phi(s(A)) \geq 0$. Sean $A \in \mathbf{M}_n$ y $\alpha \in \mathbb{C}$ entonces

$$|||\alpha A|||_\Phi = \Phi(s(\alpha A)) = \Phi(|\alpha| s(A)) = |\alpha|\Phi(s(A)) = |\alpha|\; |||A|||_\Phi.$$

Veamos que para toda $A, B \in \mathbf{M}_n$ se cumple

$$|||A + B|||_\Phi \leq |||A|||_\Phi + |||B|||_\Phi.$$

Por la Proposición 1.2.26 tenemos que

$$\sum_{j=1}^{k} s_j(A+B) \leq \sum_{j=1}^{k} s_j(A) + \sum_{j=1}^{k} s_j(B),$$

de donde $s(A+B) \prec_\omega s(A) + s(B)$. Así, por la Proposición 2.2.10 tenemos que

$$\Phi(s(A+B)) \leq \Phi(s(A) + s(B)) \leq \Phi(s(A)) + \Phi(s(B)).$$

Esto demuestra que $|||\cdot|||_\Phi$ es una norma.

*(ii)* Como $|||\cdot|||$ es norma, entonces para toda $x \in \mathbb{C}^n$ se tiene que $\Phi_{|||\cdot|||}(x) = |||\text{diag}\,(x)||| \geq 0$. Sean $x \in \mathbb{C}^n$ y $\alpha \in \mathbb{C}$ resulta que

$$\Phi_{|||\cdot|||}(\alpha x) = |||\text{diag}\,(\alpha x)||| = |||\alpha\,\text{diag}\,(x)||| = |\alpha|\; |||\text{diag}\,(x)||| = |\alpha|\; \Phi_{|||\cdot|||}(x).$$

Sean $x, y \in \mathbb{C}^n$ tenemos

$$\begin{aligned}\Phi_{|||\cdot|||}(x+y) &= |||\text{diag}\,(x+y)||| = |||\text{diag}\,(x) + \text{diag}\,(y)||| \\ &\leq |||\text{diag}\,(x)||| + |||\text{diag}\,(y)||| = \Phi_{|||\cdot|||}(x) + \Phi_{|||\cdot|||}(y).\end{aligned}$$

Nótese que las matrices $P \in S_n$ y las matrices de la forma

$$\xi = \begin{pmatrix} e^{i\theta_1} & \dots & 0 \\ & \ddots & \\ 0 & \dots & e^{i\theta_n} \end{pmatrix}$$

son unitarias para $\theta_j \in \mathbb{R}$, $j = 1, \cdots, n$. Sea $P_\sigma \in S_n$ entonces como $|||\cdot|||$ cumple la identidad (2.23) tenemos que

$$\begin{aligned}\Phi_{|||\cdot|||}(Px) =& |||\operatorname{diag}(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)})||| = |||P \operatorname{diag}(x) Id||| \\ =& |||\operatorname{diag}(x)||| = \Phi_{|||\cdot|||}(x).\end{aligned}$$

Sean $\theta_j = 0, \pi$, entonces $\xi_j = \pm 1$ para $j = 1, 2, \ldots, n$. Sea $x \in \mathbb{C}^n$, se tiene que

$$\Phi_{|||\cdot|||}(\xi x) = |||\operatorname{diag}(\xi x)||| = |||\xi \operatorname{diag}(x) Id||| = |||\operatorname{diag}(x)||| = \Phi_{|||\cdot|||}(x).$$

Sea $\mathbf{e}_1 = (1, 0, \ldots, 0)$, calculamos

$$\Phi_{|||\cdot|||}(\mathbf{e}_1) = |||\operatorname{diag}(\mathbf{e}_1)||| = 1. \qquad \square$$

Las funciones simétricas y calibradas construidas en la sección anterior conducen a varios ejemplos de normas unitariamente invariantes. Dos clases de tales normas son

(1) Las $p$-normas de Schatten. Para $1 \leq p \leq \infty$

$$\|A\|_p = \|s(A)\|_p = \left[\sum_{j=1}^{n} (s_j(A))^p\right]^{1/p}. \tag{2.26}$$

$$\|A\|_\infty = \|s(A)\|_\infty = s_1(A) = \|A\|. \tag{2.27}$$

(2) Las $k$-normas de Ky Fan

$$\|A\|_{(k)} = \Phi_{(k)}(s(A)) = \sum_{j=1}^{k} s_j(A) \qquad \text{para } 1 \leq k \leq n. \tag{2.28}$$

Hemos visto que para la norma de Ky Fan $\|A\|_{(1)} = \|A\|_\infty = \|A\|$ y $\|A\|_{(n)} = \|A\|_1 = \|\,|A|\,\|_1 = \operatorname{Tr}(|A|)$, donde $\operatorname{Tr}(A)$ denota la traza del operador $A$. En la base que diagonaliza a $|A|$ tenemos que

$$\operatorname{Tr}(|A|^p)^{\frac{1}{p}} = \Big(\sum_{j=1}^{n} s_j(A)^p\Big)^{\frac{1}{p}}.$$

Entonces, como la traza es invariante bajo cambios de base, esto demuestra que

$$\|A\|_p = \operatorname{Tr}(|A|^p)^{\frac{1}{p}}.$$

Para $A$, $B$ matrices complejas $n \times n$, el producto interno se define mediante

$$\langle A, B\rangle_{HS} = \operatorname{Tr}\big(A^*B\big).$$

Este producto interno provee a $\mathbf{M}_n$ con una estructura de espacio de Hilbert. La norma inducida por este producto interno es

$$\|A\|_{HS} = \operatorname{Tr}\big(A^*A\big)^{\frac{1}{2}} = \operatorname{Tr}(|A|^2)^{\frac{1}{2}} = \Big(\sum_j (s_j(A))^2\Big)^{\frac{1}{2}} = \|A\|_2,$$

si usamos la base que diagonaliza a $|A|$ para calcular la traza. Entonces la norma Hilbert-Schmidt coincide con la $2-$norma de Schatten. Esta norma también se llama norma de Frobenius. Si la matriz asociada a $A$ tiene entradas $a_{ij}$ entonces $\|A\|_2 = \Big(\sum_{ij} |a_{ij}|^2\Big)^{1/2}$. Así la norma $\|\cdot\|_2$ es la norma euclidiana del operador $A$ pensando a éste como vector en $\mathbb{C}^{n^2}$. En el próximo capítulo usaremos la estructura de espacio de Hilbert inducida por el producto Hilbert-Schmidt.

El siguiente Teorema nos dice la importancia de las normas de Ky Fan.

**Teorema 2.3.3 (Dominación de Ky Fan).** *Sean $A, B \in \mathbf{M}_n$; si $\|A\|_{(k)} \leq \|B\|_{(k)}$ para $k = 1, 2, \ldots, n$. Entonces*

$$|||A||| \leq |||B|||$$

*para todas las normas unitariamente invariantes $|||\cdot|||$.*

*Demostración.* Sea $|||\cdot|||$ una norma unitariamente invariante y $\Phi := \Phi_{|||\cdot|||}$ la función simétrica y calibrada asociada. Por la descomposición en valores singulares, para cualquier matriz $A$ se tiene que $\operatorname{diag}(s(A)) = W^*AQ$, donde $W$ y $Q$ matrices unitarias, en consecuencia $|||A||| = |||W^*AQ||| = |||\operatorname{diag}(s(A))|||$. Por otra parte, del Teorema 2.3.2 tenemos que $|||A||| = |||A|||_\Phi = \Phi_{|||\cdot|||}(s(A))$.

Sean $A, B$ matrices de tamaño $n \times n$. Si $\|A\|_{(k)} \leq \|B\|_{(k)}$ para $k = 1, 2, \ldots, n$; por la ecuación (2.28) tenemos que

$$\sum_{j=1}^{k} s_j(A) \leq \sum_{j=1}^{k} s_j(B) \qquad \text{para toda } 1 \leq k \leq n.$$

Entonces $s(A) \prec_\omega s(B)$. Por la Proposición 2.2.10 se tiene que

$$|||A||| = \Phi_{|||\cdot|||}(s(A)) \leq \Phi_{|||\cdot|||}(s(B)) = |||B|||. \qquad \square$$

Recordemos que de la Proposición 2.2.3 se cumple que $\|x\|_\infty \leq \Phi(x) \leq \|x\|_1$ para toda $x \in \mathbb{C}^n$ y toda función simétrica y calibrada $\Phi$. Entonces por el Teorema 2.3.2 tenemos que $\|A\|_\infty = \|A\| \leq |||A||| \leq \|A\|_{(n)} = \|A\|_1$, para toda $A \in \mathbf{M}_{n\times n}$ y toda norma unitariamente invariante $||| \cdot |||$.

La dualidad en el espacio de normas unitariamente invariantes se define de la siguiente manera. Si $|||\cdot|||$ es una norma untariamente invariante, entonces su norma dual es

$$|||A|||' = \sup_{|||B|||=1} |\mathrm{Tr}(A^*B)|.$$

Es fácil demostrar que $||| \cdot |||$ es una norma unitariamente invariante.

**Proposición 2.3.4.** *Si $\Phi$ es función simétrica y calibrada en $\mathbb{R}^n$ y $\| \cdot \|_\Phi$ la correspondiente norma unitariamente invariante en $M_n(\mathbb{C})$. Entonces $\|\cdot\|'_\Phi = \| \cdot \|_{\Phi'}$.*

*Demostración.* Nótese que

$$|\mathrm{Tr}(A^*B)| \leq \mathrm{Tr}|A^*B| = \sum_{j=1}^n s_j(A^*B).$$

Por la Proposición IV.2.5 en [1], tenemos que

$$\sum_{j=1}^n s_j(A^*B) \leq \sum_{j=1}^n s_j(A)s_j(B).$$

Entonces

$$\|A\|'_\Phi \leq \Phi'(s(A)) = \|A\|_{\Phi'}.$$

Recíprocamente,

$$\begin{aligned}
\|A\|_{\Phi'} = \Phi'(s(A)) &= \sup\Big\{ \sum_{j=1}^n s_j(A)y_j : \ y \in \mathbb{R}^n,\ \Phi(y) = 1\Big\} \\
&= \sup\{\mathrm{Tr}\big(\mathrm{diag}(s(A))\mathrm{diag}(y)\big) : \|\mathrm{diag}(y)\|_\Phi = 1\} \\
&\leq \|\mathrm{diag}(s(A))\|'_\Phi = \|A\|'_\Phi.
\end{aligned}$$

□

**Corolario 2.3.5.** *Dadas $A$ y $B$ matrices complejas $n \times n$, entonces* *(i)* $|\mathrm{Tr}\,(A^*B)| \leq |||A|||\,|||B|||'$.

*(ii)* $\|A\|_p' = \|A\|_q$, *si* $\frac{1}{p} + \frac{1}{q} = 1$.

*Demostración.* Es una consecuencia inmediata de la proposición anterior. $\square$

Denotaremos por $L_p(\mathbb{C}^n)$ al espacio de Banach $\big(M_n(\mathbb{C}), \|\cdot\|_p\big)$ y $L_p(\mathbb{C}^n)^*$ denotará a su dual.

**Proposición 2.3.6.** *Si* $\frac{1}{p} + \frac{1}{q} = 1$*, la transformación* $\mathrm{Tr}\,(A\,\cdot) : L_q(\mathbb{C}^n) \to L_p(\mathbb{C}^n)^*$ *es una isometría suprayectiva.*

*Demostración.* Por el inciso $(i)$ de la proposición anterior se tiene que $\mathrm{Tr}(AB) \leq \|A\|_p\|B\|_q$ para todo $B \in L_q(\mathbb{C}^n)$ Entonces $\mathrm{Tr}(A\cdot)$ es un funcional lineal con norma $\|\mathrm{Tr}(A\cdot)\| \leq \|A\|_q$.

Recíprocamente, si $f$ es un funcional lineal sobre $L_p(\mathbb{C}^n)$ entonces tenemos que la función definida para $u, v \in \mathbb{C}^n$ mediante $c(u,v) := f(|v\rangle\langle u|)$, es una forma sesquilineal sobre $\mathbb{C}^n \times \mathbb{C}^n$ y además

$$|c(u,v)| = |f(|v\rangle\langle u|)| \leq \|f\|\||v\rangle\langle u|\|_p = \|f\|\|u\|\|v\|\big(\mathrm{Tr}|\ |\hat{v}\rangle\langle\hat{u}|\ |^p\big)^{\frac{1}{p}},$$

donde $\hat{u} = \frac{u}{\|u\|}$. Pero no es difícil ver que para vectores unitarios se tiene $\big(\mathrm{Tr}|\ |\hat{v}\rangle\langle\hat{u}|\ |^p\big)^{\frac{1}{p}} = 1$. Entonces tenemos que

$$|c(u,v)| \leq \|f\|\|u\|\|v\|.$$

Es decir la forma sesquilineal $c$ es acotada en $\mathbb{C}^n$. Por lo tanto existe un operador $A$ tal que $f(|v\rangle\langle u|) = c(u,v) = \langle A^*u, v\rangle = \mathrm{Tr}\big(A^*|v\rangle\langle u|\big)$ para todo $u, v \in \mathbb{C}^n$. Esto implica que $f(x) = \mathrm{Tr}(A^*x)$ para todo $x \in M_n(\mathbb{C})$. $\square$

# Capítulo 3

# Desigualdades tipo Minkowski y subaditividad fuerte de la entropía cuántica

## 3.1. Normas $L^p(L^q)$ y concavidad/convexidad

El contenido de este capítulo se encuentra en las referencias [4] y [5]. Hemos ordenado el material de una manera un poco diferente y completado todos los detalles en las demostraciones. Iniciamos con la definición de ciertos funcionales $\Phi_{p,q}$, $\Psi_{p,q}$, $\Upsilon_{p,q}$ y determinaremos condiciones sobre $p$, $q$ para las cuales se obtiene concavidad o convexidad.

Denotaremos por $\mathbb{H}_n$ al conjunto de matrices hermitianas de tamaño $n \times n$ y usaremos $\mathbb{H}_n^+$ para el conjunto de matrices positivas. Para $m, n \in \mathbb{Z}^+$, podemos identificar operadores sobre $\mathbb{C}^n \otimes \mathbb{C}^m$ con matrices de tamaño $mn \times mn$ y usaremos $\mathbb{H}_{nm}$ para referirnos al conjunto de matrices hermitianas sobre el producto tensorial $\mathbb{C}^n \otimes \mathbb{C}^m$.

**Definición 3.1.1.** *Para números $p$, $q > 0$ y un entero positivo $m$, definimos $\Phi_{p,q}$ sobre el producto cartesiano de $m-$copias de $\mathbb{H}_n^+$, mediante*

$$\Phi_{p,q}(A_1, \ldots, A_m) = \left\| \left( \sum_{j=1}^{m} A_j^p \right)^{1/p} \right\|_q . \tag{3.1}$$

**Definición 3.1.2.** *Para números $p,\, q > 0$ y cualesquiera dos enteros positivos $m,\, n$. Definimos $\Psi_{p,\,q}$ sobre el conjunto $\mathbb{H}^+_{mn}$ mediante*

$$\Psi_{p,\,q}(A) = \|(\operatorname{Tr}_2 A^p)^{1/p}\|_q, \tag{3.2}$$

*donde* $\operatorname{Tr}_2(\cdot)$ *denota el operador de traza parcial (ver Apéndice).*

**Definición 3.1.3.** *Para cualquier matriz fija $B \in \mathbf{M}_n(\mathbb{C})$ y números $p,\, q > 0$. Definimos $\Upsilon_{p,\,q,\,B}$ sobre el conjunto $\mathbb{H}^+_n$, mediante*

$$\Upsilon_{p,\,q,\,B}(A) = \operatorname{Tr}\left[(B^* A^p\, B)^{q/p}\right]. \tag{3.3}$$

En adelante, escribiremos simplemente $\Upsilon_{p,\,q}$ en lugar de $\Upsilon_{p,\,q,\,B}$.

**Definición 3.1.4.** *Sean $U$, $V$ espacios vectoriales. Una función $f : U \times V \to \mathbb{R}$ es conjuntamente cóncava si, para $A_1, A_2 \in U$, $B_1, B_2 \in V$ y $0 \le t \le 1$ se cumple que*

$$f\big(tA_1 + (1-t)A_2, tB_1 + (1-t)B_2\big) \ge t\, f(A_1, B_1) + (1-t) f(A_2, B_2).$$

*Decimos que $f$ es conjuntamente convexa si $-f$ es conjuntamente cóncava.*

**Definición 3.1.5.** *Si $V, W$ son espacios vectoriales sobre $\mathbb{R}$, $f : V \to W$ y $k \in \mathbb{R}$. Decimos que $f$ es una función homogénea de grado $k$ si*

$$f(\lambda\, v) = \lambda^k f(v) \qquad \textit{para toda } \lambda \ge 0 \textit{ y toda } v \in V. \tag{3.4}$$

**Proposición 3.1.6.** *Si $f : V \to \mathbb{R}$ una función homogénea de grado $1$. Entonces $f$ es convexa si y sólo si el conjunto de nivel $\Omega = \{x \in V : f(x) \le 1\}$ es convexo.*

*Demostración.* Supongamos que $f$ es convexa. Sean $x, y \in \Omega$, entonces $f(x) \le 1$ y $f(y) \le 1$. De la convexidad de la función

$$f(tx + (1-t)y) \le tf(x) + (1-t)f(y) \le t + (1-t) = 1.$$

Así $tx + (1-t)y \in \Omega$, es decir el conjunto es convexo.

Recíprocamente, supóngase que $\Omega$ es un conjunto convexo. Nótese que $\Omega$ es absorbente, pues para $x \in V \setminus \Omega$ se tiene que $f(x) > 1$, entonces tomando $\lambda = \frac{1}{f(x)}$ tenemos que $\lambda x = \frac{x}{f(x)} \in \Omega$.

Definamos $P_\Omega : V \to \mathbb{R}$, mediante

$$P_\Omega(x) = \text{ínf}\{t > 0 : \frac{x}{t} \in \Omega\} = \text{ínf}\{t > 0 : f(x) \leq t\}.$$

Veamos que éste funcional es homogéneo. Sea $\alpha > 0$, entonces

$$\begin{aligned} P_\Omega(\alpha\, x) &= \text{ínf}\{t > 0 : f(\alpha\, x) \leq t\} = \text{ínf}\{t > 0 : f(x) \leq \frac{t}{\alpha}\} \\ &= \text{ínf}\{\alpha t' > 0 : f(x) \leq t'\} = \alpha\, \text{ínf}\{t' > 0 : f(x) \leq t'\} = \alpha\, P_\Omega(x). \end{aligned}$$

Sea $G(x) = \{t > 0 : f(x) \leq t\}$. $G(x) \neq \emptyset$, pues el mismo $f(x)$ es un elemento del conjunto si $f(x) > 0$. En caso contrario, si $f(x) < 0$ cualquier $t > 0$ funciona. Así, tomar ínfimo tiene sentido.

Sean $s, r > 0$ tales que $a = rx \in \Omega$ y $b = s(-x) \in \Omega$, pues este conjunto es absorbente. Tenemos que $0 = \frac{sa+rb}{s+r} \in \Omega$. Si $x \in \Omega$ y $t < 1$; entonces $tx = (1-t) \cdot 0 + xt \in \Omega$.

Sea $t \in G(x)$ y $s > 1$, entonces $\frac{f(x)}{t} \leq 1$, por lo tanto $\frac{x}{t} \in \Omega$. Más aún, $\frac{x}{st} \in \Omega$, de donde $st \in G(x)$. Tenemos que $\text{ínf}\, G(x) = P_\Omega(x)$.

Si $P_\Omega(x) \in G(x)$, tenemos que $G(x) = [P_\Omega(x), \infty)$, es una semirecta cerrada; de lo contrario $G(x) = (P_\Omega, \infty)$ es abierto.

Sean $s > P_\Omega(x)$ y $r > P_\Omega(y)$, así que $\frac{x}{s}, \frac{y}{r} \in \Omega$. Luego, por ser $\Omega$ convexo

$$\frac{x+y}{r+s} = \frac{s}{r+s}\frac{x}{s} + \frac{r}{r+s}\frac{y}{t} \in \Omega.$$

Además $r + s \geq P_\Omega(x+y)$. Así que, tomando ínfimo sobre $r, s$ se cumple que

$$P_\Omega(x) + P_\Omega(y) \leq P_\Omega(x+y).$$

Usando esto, supongamos que $x, y \in \Omega$. Sean $f(x) > 0$ y $f(y) > 0$, de lo contrario usamos la homogeneidad de la función. Entonces, usando la homogeneidad y la sublinealidad de $P_\Omega$ obtenemos que

$$\begin{aligned} f(tx + (1-t)y) &\leq P_\Omega(tx + (1-t)y) \leq tP_\Omega(x) + (1-t)P_\Omega(y) \\ &= t\, \text{ínf}\{s > 0 : f(x) \leq s\} + (1-t)\, \text{ínf}\{s' > 0 : f(y) \leq s'\} \\ &\leq tf(x) + (1-t)f(y). \end{aligned}$$

□

**Proposición 3.1.7.** *Sea $B \in \mathbf{M}_n$ fija. Para toda $A \in \mathbb{H}_n^+$ la función $\Upsilon_{p,q}$ es una homogénea de grado $q \geq 1$.*

*Demostración.* Sea $\lambda \geq 0$. Entonces

$$\begin{aligned}\Upsilon_{p,q}(\lambda A) &= \mathrm{Tr}\left[(B^*(\lambda A)^p B)^{q/p}\right] = \mathrm{Tr}\left[(B^*\lambda^p A^p B)^{q/p}\right] = \\ &= \mathrm{Tr}\left[\lambda^q (B^* A^p B)^{q/p}\right] = \lambda^q \mathrm{Tr}\left[(B^* A^p B)^{q/p}\right] = \lambda^q\, \Upsilon_{p,q}(A). \quad \square\end{aligned}$$

A continuación daremos la prueba de la concavidad/convexidad del funcional $\Upsilon_{p,q}$, la cual se divide en dos casos, a saber $1 \leq q \leq p \leq 2$ y $1 \leq p \leq 2$ para $p < q$.

**Lema 3.1.8.** *Para $B \in \mathbf{M}_n(\mathbb{C})$ fija. Si $1 \leq p \leq 2$ y $q > p$, entonces*

$$\Upsilon_{p,q}(A) = \mathrm{Tr}\left[(B^*\, A^p\, B)^{q/p}\right],$$

*es convexa sobre $\mathbb{H}_n^+$.*

*Demostración.* Sea $A \in \mathbb{H}_n^+$, se tiene que $A^p \geq 0$. Sea $h \in \mathbb{C}^n$, entonces

$$\langle h, B^* A^p B\, h\rangle = \langle Bh, A^p B\, h\rangle \geq 0.$$

Sean $r = q/p \geq 1$ y $r'$ tales que $\frac{1}{r} + \frac{1}{r'} = 1$. Tenemos que $\|B^* A^p B\|_r = [Tr(B^* A^p B)^r]^{1/r} = \left[\Upsilon_{p,q}(A)\right]^{1/r}$. Por otro lado, por dualidad podemos escribir

$$\|B^* A^p B\|_r = \sup_{\substack{\|Y\|_{r'} \leq 1 \\ Y \geq 0}} \mathrm{Tr}\,[B^* A^p B\, Y].$$

Dado que, para $1 \leq p \leq 2$, la función $A \to A^p$ es convexa en operadores; así también $B^* A^p B$ lo es. Como tomamos $Y \geq 0$ tenemos que $Tr[B^* A^p B\, Y]$ es convexa. Así $\|B^* A^p B\|_r$ es el supremo de una familia de funciones convexas y por tanto es convexa. Como $\Upsilon_{p,q}(A) = (\|B^* A^p B\|_r)^r$ y para $r \geq 1$ la $r$-ésima potencia de una función positiva y convexa es convexa, se obtiene que $\Upsilon_{p,q}(A)$ es convexa. $\square$

Para el caso $q < p$ la demostración es más complicada y necesitamos los siguientes lemas.

**Lema 3.1.9.** *Para toda $r > 1$ y toda $c, x > 0$ se tiene que $\frac{1}{r}c^r + \frac{r-1}{r}x^r \geq c\,x^{r-1}$ y*

$$c = \frac{1}{r} \inf\left\{\frac{c^r}{x^{r-1}} + (r-1)x : x > 0\right\}.$$

*Demostración.* De la relación entre la media aritmética y la media geométrica (2.6) tomando $t_1 = c^r$, $t_2 = x^r$ y $a_1 = r^{-1}$, $a_2 = 1 - r^{-1} = \frac{r-1}{r}$ y como $x > 0$, tenemos que

$$c \leq \frac{1}{r}\frac{c^r}{x^{r-1}} + \frac{r-1}{r}x.$$

Consideremos la función $f(x) = \frac{c^r}{x^{r-1}} + (r-1)x$. Nótese que $f(c) = c$, de donde

$$c = \frac{1}{r}\,\text{ínf}\left\{\frac{c^r}{x^{r-1}} + (r-1)x : x > 0\right\}.$$

Además si $0 < r < 1$, entonces $f''(c) < 0$ y alcanza un valor máximo. Entonces $c$ sería un supremo. $\square$

**Lema 3.1.10.** *Si $B \in \mathcal{B}(\mathcal{H})$, fija y $A \in \mathbb{H}_n^+$. Entonces los operadores $B^*A^pB$ y $A^{p/2}B\,B^*A^{p/2}$ tienen el mismo conjunto de valores propios, además*

$$\Upsilon_{p,q}(A) = \operatorname{Tr}\left[(A^{p/2}B\,B^*A^{p/2})^{\frac{q}{p}}\right].$$

*Demostración.* Sean $B \in \mathcal{B}(\mathcal{H})$ fija y $A \in \mathbb{H}_n^+$. Nótese que

$$B^*\,A^p\,B = B^*\,A^{p/2}A^{p/2}\,B = (A^{p/2}\,B)^*A^{p/2}\,B \geq 0, \tag{3.5}$$

$$A^{p/2}B\,B^*A^{p/2} = A^{p/2}\,B(A^{p/2}\,B)^* \geq 0. \tag{3.6}$$

Tenemos que ambos operadores son positivos, tienen espectro real y son funciones del mismo operador, $A^{p/2}B$. Sean $f_1(z) = z \cdot \overline{z}$, $f_1(z) = \overline{z} \cdot z$, tenemos que $f_2(z) = \overline{f_1(z)}$. Entonces por el Teorema del mapeo espectral

$$\begin{aligned}\sigma(f_2(A^{p/2}B)) &= f_2(\sigma(A^{p/2}B)) = |\sigma(A^{p/2}B)|^2\\ &= f_1(\sigma(A^{p/2}B)) = \sigma(f_1(A^{p/2}B)).\end{aligned}$$

Ahora, usando la base que diagonaliza a $A^{p/2}B$ tenemos

$$\begin{aligned}\Upsilon_{p,q}(A) &= \operatorname{Tr}\left[(B^*A^p\,B)^{q/p}\right] = \sum_j f_2(\lambda_j)^{q/p} = \sum_j |\lambda_j|^{2q/p}\\ &= \sum_j f_1(\lambda_j)^{q/p} = \operatorname{Tr}\left[(A^{p/2}B\,B^*A^{p/2})^{q/p}\right].\end{aligned}$$

$\square$

En el siguiente Lema demostramos una fórmula que nos da una variante de $\Upsilon_{p,q}$.

**Lema 3.1.11.** *Sea* $A \in \mathbb{H}_n^+$, $r = p/q > 1$. *Entonces*

$$\Upsilon_{p,q}(A) = \frac{1}{r} \operatorname{ínf} \left\{ \operatorname{Tr}\left[A^{p/2} B \frac{1}{X^{r-1}} B^* A^{p/2} + (r-1)X\right] \right\},$$

*donde el ínfimo es tomado sobre todas las matrices* $X \in \mathbb{H}_n^+$. *Si el ínfimo se reemplaza por supremo la fórmula es válida para* $\frac{p}{q} < 1$.

*Demostración.* Sea $C = (A^{p/2}\,B)^*$. Por continuidad podemos asumir que $C^*C > 0$. Tomamos, para toda $X > 0$

$$\operatorname{Tr}\left[C^* \frac{1}{X^{r-1}} C + (r-1)X\right]. \tag{3.7}$$

Para $r > 1$, existe un minimizador $X$. Sea $Y = X^{1-r}$, como es un cambio de variable minimizar la expresión (3.7) respecto a $X$ es lo mismo que minimizar $\operatorname{Tr}\left[C^* Y\, C + (r-1) Y^{\frac{-1}{r-1}}\right]$ respecto a $Y$. Equivalentemente por ciclicidad y linealidad de la traza, se tiene que

$$\operatorname{Tr}\left[C\,C^*\,Y + (r-1)Y^{\frac{-1}{r-1}}\right]. \tag{3.8}$$

Como $X > 0$, se sigue que $Y > 0$, podemos reemplazar $Y$ por $Y + tD$, con $D$ autoadjunto. Entonces $\operatorname{Tr}\left[C\,C^*\,(Y+tD) + (r-1)(Y+tD)^{\frac{-1}{r-1}}\right]$. Derivando respecto a $t$, según la regla del Teorema A.3.4 se tiene que

$$\begin{aligned}
\frac{d}{dt}\operatorname{Tr}\left[C\,C^*\,(Y+tD) + (r-1)(Y+tD)^{\frac{-1}{r-1}}\right]\Big|_{t=0} \\
&= \operatorname{Tr}\left[CC^*D - (Y+tD)^{\frac{-r}{r-1}} D\right]\Big|_{t=0} \\
&= \operatorname{Tr}\left[CC^*D - Y^{\frac{-r}{r-1}} D\right] \\
&= \operatorname{Tr}\left[(CC^* - Y^{\frac{-r}{r-1}})D\right] = \operatorname{Tr}\left[D(CC^* - Y^{\frac{-r}{r-1}})\right].
\end{aligned}$$

Ahora,

$$0 = \operatorname{Tr}\left[D(CC^* - Y^{\frac{-r}{r-1}})\right] = \left\langle D, CC^* - Y^{\frac{-r}{r-1}} \right\rangle,$$

para todo $D$ autoadjunto. Por lo tanto, $CC^* = Y^{\frac{-r}{r-1}}$, así que $Y = (CC^*)^{-\frac{r-1}{r}}$. Sustituyendo en (3.8)

$$\begin{aligned}
\operatorname{Tr}\left[C\,C^*\,Y + (r-1)Y^{\frac{-1}{r-1}}\right] &= \operatorname{Tr}\left[C\,C^*\,(CC^*)^{-\frac{r-1}{r}} + (r-1)\left((CC^*)^{-\frac{r-1}{r}}\right)^{\frac{-1}{r-1}}\right] \\
&= \operatorname{Tr}\left[(CC^*)^{1-\frac{-1}{r-1}} + (r-1)(CC^*)^{1/r}\right] \\
&= r\operatorname{Tr}\left[(CC^*)^{1/r}\right] \\
&= r\operatorname{Tr}\left[(B\,A^{1/2}A^{1/2}B)^{1/r}\right] = r\,\Upsilon_{p,q}(A).
\end{aligned}$$

Y se sigue el resultado. La correspondiente fórmula para $\frac{p}{q} < 1$ se demuestra de manera análoga. $\square$

**Lema 3.1.12.** *Si $f(x, y)$ es una función conjuntamente convexa. Entonces $g(x) = \inf_y f(x, y)$ es convexa. El enunciado sigue siendo válido cuando cambiamos convexa por cóncava y tomamos supremo en lugar de ínfimo.*

*Demostración.* Sean $x_0$, $x_1$ y $0 < \lambda < 1$. Sea $\varepsilon > 0$. Elegimos $y_0$, $y_1$ tales que

$$f(x_0, y_0) < g(x_0) + \varepsilon \quad y \quad f(x_1, y_1) < g(x_1) + \varepsilon.$$

Entonces

$$\begin{aligned} g\big((1-\lambda)x_0 + \lambda x_1\big) &= \inf_y f\big((1-\lambda)x_0 + \lambda x_1, y\big) \\ &\leq f\big((1-\lambda)x_0 + \lambda x_1, (1-\lambda)y_0 + \lambda y_1\big) \\ &\leq (1-\lambda)f(x_0, y_0) + \lambda f(x_1, y_1) \\ &< (1-\lambda)g(x_0) + \lambda g(x_1) + 2\varepsilon. \end{aligned}$$

$\square$

Un caso especial de este Lema se tiene para una función convexa en $\mathbb{R}^n \times \mathbb{R}^n$ dada por $f(x, y) = h(x-y) + g(y)$, donde $h$, $g$ son funciones convexas en $\mathbb{R}^n$. Entonces $\inf_y f(x, y)$ se llama convolución ínfima de las funciones $h, g$. La concavidad de la función, cuando cambiamos ínfimo por supremo, aparece en el siguiente hecho físico: Cuando se combinan dos sistemas físicos y permite el flujo de calor entre ellos de tal forma que la energía total $x$ se conserva, la energía se distribuye de tal forma que la entropía es maximizada. Entonces la entropía dada por

$$S(x) = \sup_y \{S_1(x-y) + S_2(y)\},$$

es nuevamente una función cóncava de la entropía total de $x$, como toda entropía debe ser.

**Lema 3.1.13.** *Si $R_1$, $R_2$, $S_1$, $S_2$, $T_1$, $T_2$ son operadores positivos sobre un espacio de Hilbert $\mathcal{H}$ y supongamos que $R_1R_2 = R_2R_1$, $S_1S_2 = S_2S_1$, $T_1T_2 = T_2T_1$ además*

$$R_1 \geq S_1 + T_1, \qquad R_2 \geq S_2 + T_2.$$

*Entonces, para $0 \leq t \leq 1$*

$$R_1^t R_2^{1-t} \geq S_1^t S_2^{1-t} + T_1^t T_2^{1-t}. \tag{3.9}$$

*Demostración.* Sea $\mathbf{E}$ el conjunto de $t \in [0,1]$ para los cuales la desigualdad (3.9) se cumple; $\mathbf{E}$ es un conjunto cerrado y claramente contiene a 0, 1.

Veamos en primer lugar que de $\frac{1}{2}$ está en $\mathbf{E}$, luego usaremos esto para mostrar que es un conjunto convexo. Sean $x$, $y \in \mathcal{H}$, entonces usando la desigualdad del triángulo, la desigualdad de Schwarz y la relación (2.6) tenemos que

$$\begin{aligned}
&\Big|\Big\langle x, (S_1^{1/2}S_2^{1/2} + T_1^{1/2}T_2^{1/2})y\Big\rangle\Big| \\
&\qquad\qquad \leq \big|\big\langle x, S_1^{1/2}S_2^{1/2}y\big\rangle\big| + \big|\big\langle x, T_1^{1/2}T_2^{1/2}y\big\rangle\big| \\
&\qquad\qquad \leq \|S_1^{1/2}x\|\|S_2^{1/2}y\| + \|T_1^{1/2}x\|\|T_2^{1/2}y\| \\
&\qquad\qquad \leq \Big[\|S_1^{1/2}x\|^2 + \|T_1^{1/2}x\|^2\Big]^{1/2}\Big[\|S_2^{1/2}y\|^2 + \|T_2^{1/2}y\|^2\Big]^{1/2} \\
&\qquad\qquad = \Big(\langle x, S_1x\rangle + \langle x, T_1x\rangle\Big)^{1/2}\Big(\langle y, S_2x\rangle + \langle y, T_2y\rangle\Big)^{1/2} \\
&\qquad\qquad = \Big(\langle x, S_1 + T_1\, x\rangle\Big)^{1/2}\Big(\langle y, S_2 + T_2\, y\rangle\Big)^{1/2} \\
&\qquad\qquad \leq \big(\langle x, R_1\, x\rangle\big)^{1/2}\big(\langle y, R_2\, y\rangle\big)^{1/2}.
\end{aligned}$$

Como $T_i + S_i \leq R_i$ para $i = 1, 2$; se tiene que $T_i^* + S_i^* \leq R_i^*$, para $i = 1, 2$. Entonces $R_i$ está acotado por abajo y por tanto es invertible. Así, tomando $u$, $v \in \mathcal{H}$

$$\begin{aligned}
&\Big|\big\langle u, R_1^{-1/2}(S_1^{1/2}S_2^{1/2} + T_1^{1/2}T_2^{1/2})R_2^{-1/2}\, v\big\rangle\Big| \\
&\qquad\qquad = \Big|\big\langle R_1^{-1/2}u, (S_1^{1/2}S_2^{1/2} + T_1^{1/2}T_2^{1/2})R_2^{-1/2}\, v\big\rangle\Big| \\
&\qquad\qquad \leq \Big[\langle R_1^{-1/2}u, R_1^{1/2}u\rangle\langle R_2^{-1/2}v, R_2^{1/2}v\rangle\Big]^{1/2} \\
&\qquad\qquad = \big[\langle u, u\rangle\langle v, v\rangle\big]^{1/2} = 1.
\end{aligned}$$

En consecuencia, $\big\|R_1^{-1/2}(S_1^{1/2}S_2^{1/2} + T_1^{1/2}T_2^{1/2})R_2^{-1/2}\big\| \leq 1$. Como los operadores dentro de la norma son positivos y como $R_1, R_2$ conmutan, entonces el producto $R_1^{-1/4}R_1^{-1/4}\Big[(S_1^{1/2}S_2^{1/2} + T_1^{1/2}T_2^{1/2})R_2^{-1/4}R_2^{-1/4}\Big]$ es normal; además tienen los mismos valores propios. De la normalidad, tenemos que el radio espectral $\rho$ coincide con la norma. Así

$$\begin{aligned}
&\left\|R_2^{-1/4}R_1^{-1/4}(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2})R_2^{-1/4}R_1^{-1/4}\right\| \\
&\qquad = \rho\left(R_2^{-1/4}R_1^{-1/4}(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2})R_2^{-1/4}R_1^{-1/4}\right) \\
&\qquad = \rho\left(R_1^{-1/4}(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2})R_2^{-1/4}R_1^{-1/4}R_2^{-1/4}\right) \\
&\qquad = \rho\left(R_1^{-1/4}(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2})R_2^{-1/2}R_1^{-1/4}\right) \\
&\qquad = \rho\left(R_1^{-1/2}(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2})R_2^{-1/2}\right) \\
&\qquad \leq \left\|R_1^{-1/2}(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2})R_2^{-1/2}\right\| \leq 1.
\end{aligned}$$

Por lo tanto $R_2^{-1/4}R_1^{-1/4}(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2})R_2^{-1/4}R_1^{-1/4} \leq I$; de donde

$$(S_1^{1/2}S_2^{1/2}+T_1^{1/2}T_2^{1/2}) \leq R_1^{1/2}R_2^{1/2}.$$

Esto muestra que $\frac{1}{2} \in \mathbf{E}$.

Ahora, supongamos que $p, q \in E$, es decir

$$\begin{aligned}
R_1^p R_2^{1-p} &\geq S_1^p S_2^{1-p} + T_1^p T_2^{1-p} \\
R_1^q R_2^{1-q} &\geq S_1^q S_2^{1-q} + T_1^q T_2^{1-q}.
\end{aligned}$$

Por lo mostrado anteriormente, tenemos que

$$\begin{aligned}
\left(R_1^p R_2^{1-p}\right)^{1/2}\left(R_1^q R_2^{1-q}\right)^{1/2} \geq \left(S_1^p S_2^{1-p}\right)^{1/2}\left(S_1^q S_2^{1-q}\right)^{1/2} \\
+ \left(T_1^p T_2^{1-p}\right)^{1/2}\left(T_1^q T_2^{1-q}\right)^{1/2}.
\end{aligned}$$

Por la conmutatividad y que $(AB)^{1/2} = A^{1/2}B^{1/2}$ en este caso, tenemos que $\frac{1}{2}(p+q)$ está en $\mathbf{E}$. Además $p = \alpha p + (1-\alpha)p$, $q = \beta q + (1-\beta)q$ entonces $\frac{1}{2}(p+q) = \left(\frac{\alpha+\beta}{2}\right)p + \left(1 - \frac{\alpha+\beta}{2}\right)q$. Por lo tanto el conjunto $\mathbf{E}$ es convexo. $\square$

**Teorema 3.1.14 (Lieb).** *Si $X \in \mathcal{B}(\mathcal{H})$ es fijo. Para cada número real $0 \leq t \leq 1$, la función*

$$f(A,B) = \operatorname{Tr}\left[X^* A^t X B^{1-t}\right],$$

*es conjuntamente cóncava para parejas de operadores positivos.*

*Demostración.* Sean $A, B \in \mathcal{B}(\mathcal{H})$ operadores positivos. Sean $L_A$, $R_B$ los operadores de multiplicación definidos para toda $X$ mediante $L_A(X) = AX$ y $R_B(X) = XB$. Son operadores positivos pues

$$\langle X, L_A(X)\rangle = \mathrm{Tr}\,(X^* L_A(X)) = \sum_j \langle e_j, X^*AX\, e_j\rangle = \sum_j \langle Xe_j, AX\, e_j\rangle \geq 0.$$

Análogamente para $R_B(X)$.

Supongamos que $A_1$, $A_2$, $B_1$, $B_2$ son operadores positivos; sean $A = A_1 + A_2$ y $B = B_1 + B_2$. Sean $L_A$, $L_{A_1}$, $L_{A_2}$ los operadores de multiplicación por la izquierda inducidos por $A$, $A_1$, $A_2$. Análogamente, sean $R_B$, $R_{B_1}$, $R_{B_1}$, los operadores de multiplicación por la derecha inducidos por $B$, $B_1$, $B_2$. Tenemos que $L_A = L_{A_1} + L_{A_2}$ y $R_B = R_{B_1} + R_{B_2}$; además conmutan entre sí, pues $L_{A_1}R_{B_1}(X) = (A_1X)B_1 = R_{B_1}L_{A_1}$. Entonces, del Lema anterior tenemos

$$L_{A_1}^t R_{B_1}^{1-t} + L_{A_2}^t R_{B_2}^{1-t} \leq L_A^t R_B^{1-t} \qquad \text{para } 0 \leq t \leq 1.$$

Equivalentemente, usando el producto Hilbert-Schmidt

$$\begin{aligned}\mathrm{Tr}\left(X^*A_1^tXB_1^{1-t}\right) + \mathrm{Tr}\left(X^*A_2^tXB_2^{1-t}\right) \leq& \mathrm{Tr}\left(X^*A^tXB^{1-t}\right)\\ =& \mathrm{Tr}\left(X^*(A_1+A_2)^tX(B_1+B_2)^{1-t}\right).\end{aligned}$$

En consecuencia, por linealidad de la traza

$$\frac{1}{2}f(A_1, B_1) + \frac{1}{2}f(A_2, B_2) \leq f\left(\frac{A_1+A_2}{2}, \frac{B_1+B_2}{2}\right).$$

Luego, para $0 \leq \lambda \leq 1$

$$\begin{aligned}&\lambda f(A_1, B_1) + (1-\lambda) f(A_2, B_2)\\ &\quad = \lambda \mathrm{Tr}\left(X^*A_1^tXB_1^{1-t}\right) + (1-\lambda)\mathrm{Tr}\left(X^*A_2^tXB_2^{1-t}\right)\\ &\quad = \frac{1}{2}\mathrm{Tr}\left(X^*(2\lambda A_1)^tX(2\lambda B_1)^{1-t}\right) + \frac{1}{2}\mathrm{Tr}\left(X^*(2(1-\lambda)A_2)^tX(2(1-\lambda)B_2)^{1-t}\right)\\ &\quad \leq \mathrm{Tr}\left(X^*\Big(\frac{2\lambda A_1 + 2(1-\lambda)A_2}{2}\Big)^t X\Big(\frac{2\lambda B_1 + 2(1-\lambda)B_2}{2}\Big)^{1-t}\right)\\ &\quad = \mathrm{Tr}\left(X^*\big(\lambda A_1 + (1-\lambda)A_2\big)^t X\big(\lambda B_1 + (1-\lambda)B_2\big)^{1-t}\right)\\ &\quad = f(\lambda A_1 + (1-\lambda)A_2, \lambda B_1 + (1-\lambda)B_2).\end{aligned}$$

□

**Proposición 3.1.15.** *Existe un isomorfismo natural $\varphi$ entre los espacios $\mathcal{H} \otimes \mathcal{K}$ y $\mathcal{L}(\mathcal{H}, \mathcal{K})$ de manera que al tensor elemental $h \otimes k$ le corresponde la transformación que envía a cada $u \in \mathcal{H}$ en $\langle h, u\rangle k$. Denotaremos a esta transformación mediante el símbolo $|k\rangle\langle h|$. Ésta es una trasformación lineal de rango uno y todas las transformaciones de rango uno pueden obtenerse de esta manera.*

*Demostración.* La demostración de que la transformación $h \otimes k \to |k\rangle\langle h|$ es un isomorfismo es directa. Sea $S$ una aplicación lineal de rango uno de la forma que se pide en el enunciado. Sea $0 \neq k \in \mathbf{Ran}(S)$ y $u \in H$ arbitraria. Entonces $S(u) = f(u)k$. Es fácil ver que $f \in \mathcal{H}^*$, pues $f(u + \alpha v)k = S(u + \alpha v) = S(u) + \alpha S(v)$ entonces $f$ es lineal. $S$ es acotada, i.e., $\|S(u)\| \leq c\|u\|$ lo cual implica que $|f(u)| \leq \frac{c}{\|k\|}\|u\|$. Por el Lema de Riesz se tiene que existe $h \in \mathcal{H}$ tal que $f(u) = \langle h, u\rangle k$. □

**Proposición 3.1.16.** *El espacio $\mathcal{L}(\mathcal{H}, \mathcal{K})$ es un espacio de Hilbert con el producto interno $\langle A, B\rangle = \operatorname{Tr} A^*B$. El conjunto de proyectores $|f_j\rangle\langle e_i|$ para $1 \leq i \leq m$ y $1 \leq j \leq n$ forman una base ortonormal para este espacio. Entonces la función $\varphi$ mencionada en la proposición anterior es un isomorfismo entre espacios de Hilbert, es decir $\langle\varphi^{-1}(A), \varphi^{-1}(B)\rangle = \langle A, B\rangle$*

*Demostración.* Sea $e_1, ..., e_n$ una base ortonormal para $\mathcal{H}$ y $f_1, ..., f_m$ una base ortonormal para $\mathcal{K}$. Entonces $\{e_i \otimes f_j\}$ es una base para $\mathcal{H} \otimes \mathcal{K}$. Este elemento se define como la forma bilineal

$$(e_i \otimes f_j)(x, y) = \langle e_i, x\rangle_{\mathcal{H}}\langle f_j, y\rangle_{\mathcal{K}}.$$

Identifíquese cada elemento de $\mathcal{L}(\mathcal{H}, \mathcal{K})$ con su matriz respecto a esa base; sea

$$\varphi(e_i \otimes f_j) = |f_j\rangle\langle e_i| \qquad \text{para toda } 1 \leq i \leq n \text{ y } 1 \leq j \leq m, \tag{3.10}$$

donde $|f_j\rangle\langle e_i|$ actúa sobre $x \in \mathcal{H}$ como $|f_j\rangle\langle e_i|x = \langle e_i, x\rangle f_j$. La matriz asociada a este operador es $E_{ji}$, la cual tiene 1 en el $j$-ésimo renglón y la $i$-ésima columna y el resto de sus elementos son cero.

Podemos escribir $A = (a_{ij}) = \sum_{ij} a_{ji}E_{ji}$, entonces

$$\begin{aligned}
\varphi^{-1}(A) =& \varphi^{-1}\left(\sum_{i,j} a_{ji}E_{ji}\right) = \sum_{ij} a_{ji}\varphi^{-1}(E_{ji}) = \sum_{ij} a_{ji}(f_i \otimes e_j) \\
=& \sum_i \left(\sum_j a_{ji}f_j\right) \otimes e_i = \sum_i (Ae_i) \otimes e_i.
\end{aligned}$$

Ahora

$$
\begin{aligned}
\left\langle \varphi^{-1}(A), \varphi^{-1}(B) \right\rangle_{\otimes} &= \left\langle \sum_i (Ae_i) \otimes e_i, \sum_{i'} (Be_{i'}) \otimes e_{i'} \right\rangle_{\otimes} \\
&= \sum_{i\, i'} \langle Ae_i, Be_{i'} \rangle_{\mathcal{K}} \, \langle e_i, e_{i'} \rangle_{\mathcal{H}} = \sum_i \langle Ae_i, Be_i \rangle_{\mathcal{K}} \\
&= \mathrm{Tr}\, [A^* \, B] = \langle A, B \rangle_{\mathcal{L}(\mathcal{H},\mathcal{K})}. \qquad \square
\end{aligned}
$$

Denotaremos simplemente por $\mathcal{L}(\mathcal{H})$ a $\mathcal{L}(\mathcal{H}, \mathcal{H})$.

**Proposición 3.1.17.** *Sea $\widetilde{\varphi}$ la trasformación inducida por el isomorfismo $\varphi : \mathcal{H} \otimes \mathcal{H}^* \to \mathcal{L}(\mathcal{H})$ de manera que el siguente diagrama conmute*

$$
\begin{tikzcd}
\mathcal{H} \otimes \mathcal{H}^* \arrow[r, "\varphi"] \arrow[d, "A\otimes B^*"'] & \mathcal{L}(\mathcal{H}) \arrow[d, "\widetilde{\varphi}(A\otimes B^*)"] \\
\mathcal{H} \otimes \mathcal{H}^* \arrow[r, "\varphi"'] & \mathcal{L}(\mathcal{H}).
\end{tikzcd}
$$

*Entonces, para toda $X \in \mathcal{L}(\mathcal{H})$, la correspondencia $\widetilde{\varphi}$ cumple*

*(i)* $\widetilde{\varphi}(A \otimes B^*)(X) = \varphi(A \otimes B^* \varphi^{-1}(X))$.

*(ii)* $\widetilde{\varphi}(A \otimes B^*)(X) = A\, X\, B$.

*Demostración.*

*(i)* Tomamos $X \in \mathcal{L}(\mathcal{H})$. Entonces $\varphi^{-1}(X) \in \mathcal{H} \otimes \mathcal{H}^*$ y $A \otimes B^* \varphi^{-1}(X) \in \mathcal{H} \otimes \mathcal{H}^*$ aplicamos $\varphi$ y obtenemos $\varphi(A \otimes B^* \varphi^{-1}(X)) = \widetilde{\varphi}(A \otimes B^*)(X)$ pues el diagrama conmuta.

*(ii)* Si $\{e_i\}_{i=1}^n$ es una base de $\mathcal{H}$, entonces $\widetilde{\varphi}(A \otimes B^*)(X) = \varphi(A \otimes B^* \varphi^{-1}(X)) = \varphi(A \otimes B^* \varphi^{-1}(\sum_i X e_i \otimes e_i)) = \sum_i \varphi(AXe_i \otimes B^* e_i) = \sum_i |AXe_i\rangle\langle B^* e_i| = AX\, (\sum_i |e_i\rangle\langle e_i|)\, B = A\, X\, B$. $\square$

**Proposición 3.1.18.** *Si $A$, $B \in \mathcal{L}(\mathcal{H})$ son operadores positivos. La concavidad de $A^t \otimes B^{1-t} : \mathcal{H} \otimes \mathcal{H}^* \to \mathcal{H} \otimes \mathcal{H}^*$ es equivalente a la concavidad de $X \mapsto \mathrm{Tr}\,(X A^t X^* B^{1-t})$.*

*Demostración.* Como $A$, $B$ son positivos entonces son autoadjuntos así como sus potencias. Luego por la proposición anterior

$$\begin{aligned}\operatorname{Tr}\left(X\,A^t\,X^*\,B^{1-t}\right) &= \left\langle X, A^t\,X^*\,B^{1-t}\right\rangle_{\mathcal{L}(\mathcal{H})} = \left\langle X, \widetilde{\varphi}(A^t\otimes B^{1-t})(X)\right\rangle_{\mathcal{L}(\mathcal{H})}\\ &= \left\langle \varphi^{-1}(X), \varphi^{-1}\left[\widetilde{\varphi}(A^t\otimes B^{1-t})(X)\right]\right\rangle_{\mathcal{H}\otimes\mathcal{H}^*}\\ &= \left\langle \varphi^{-1}(X), \varphi^{-1}\left[\varphi\left(A^t\otimes B^{1-t}\varphi^{-1}(X)\right)\right]\right\rangle_{\mathcal{H}\otimes\mathcal{H}^*}\\ &= \left\langle \varphi^{-1}(X), A^t\otimes B^{1-t}\varphi^{-1}(X)\right\rangle_{\mathcal{H}\otimes\mathcal{H}^*}.\end{aligned}$$

Así la concavidad/convexidad de una función equivale a la de la otra. □

**Lema 3.1.19.** *La transformación sobre $\mathbb{H}_n^+\times\mathbb{H}_n^+$ definida mediante*

$$(A,X)\to\operatorname{Tr}\left(A^{p/2}\,B^*\frac{1}{X^{r-1}}BA^{p/2}\right), \tag{3.11}$$

*es conjuntamente convexa para toda $1\le r\le p\le 2$ y conjuntamente cóncava para toda $0<p<r<1$*

*Demostración.* Escribiremos la expresión (3.11) de una forma más conveniente. Definamos

$$Z=\begin{bmatrix} A & 0\\ 0 & X\end{bmatrix},\qquad\qquad K=\begin{bmatrix} 0 & 0\\ B & 0\end{bmatrix}.$$

Entonces

$$\begin{aligned}Z^p\,K^*\,Z^{1-r}\,K &= \begin{bmatrix} A^p & 0\\ 0 & X^p\end{bmatrix}\begin{bmatrix} 0 & B^*\\ 0 & 0\end{bmatrix}\begin{bmatrix} A^{1-r} & 0\\ 0 & X^{1-r}\end{bmatrix}\begin{bmatrix} 0 & 0\\ B & 0\end{bmatrix}\\ &= \begin{bmatrix} A^p\,B^*X^{1-r}B & 0\\ 0 & 0\end{bmatrix}.\end{aligned}$$

Luego, por la ciclicidad de la traza se tiene que

$$\begin{aligned}\operatorname{Tr}\left(K\,Z^p\,K^*\,Z^{1-r}\right) &= \operatorname{Tr}\left(Z^p\,K^*\,Z^{1-r}\,K\right)\\ &= \operatorname{Tr}\left(\begin{bmatrix} A^p\,B^*X^{1-r}B & 0\\ 0 & 0\end{bmatrix}\right) = \operatorname{Tr}\left(A^p\,B^*X^{1-r}B\right)\\ &= \operatorname{Tr}\left(A^{p/2}\,A^{p/2}\,B^*X^{1-r}B\right) = \operatorname{Tr}\left(A^{p/2}\,B^*X^{1-r}B\,A^{p/2}\right)\\ &= \operatorname{Tr}\left(A^{p/2}\,B^*\frac{1}{X^{r-1}}B\,A^{p/2}\right),\end{aligned}$$

por lo tanto, de la Proposición 3.1.18 se tiene la convexidad (concavidad). □

**Teorema 3.1.20.** *Para* $1 \leq q \leq p \leq 2$, $\Upsilon_{p,q}$ *es convexa sobre* $\mathbb{H}_n^+$. *Para* $0 \leq p \leq q \leq 1$, $\Upsilon_{p,q}$ *es cóncava en* $\mathbb{H}_n^+$.

*Demostración.* Por el Lema 3.1.19, la transformación $\mathrm{Tr}\left[A^{1/2}\, B^* \frac{1}{X^{r-1}} B\, A^{p/2}\right]$, es conjuntamente convexa para $1 \leq r \leq p \leq 2$. Por los lemas 3.1.11 y 3.1.12, tomando $r = p/q$ tenemos que $\Upsilon_{p,q}(A) = \inf_X f(A, X)$, donde $f(A, X)$ es conjuntamente convexa en $A$ y $X$. Luego, por el Lema 3.1.12, se sigue que $\Upsilon_{p,q}$ es convexa. La concavidad se demuestra de la misma forma. □

**Teorema 3.1.21.** *Para* $1 \leq p \leq 2$ *y toda* $q \geq 1$, $\Phi_{p,q}$ *es conjuntamente convexa en* $(\mathbb{H}_n^+)^m$. *Para* $0 \leq p \leq q \leq 1$, $\Phi_{p,q}$ *es conjuntamente cóncava en* $(\mathbb{H}_n^+)^m$.

*Demostración.* Definamos las matrices de tamaño $mn \times mn$ como sigue

$$\mathscr{A} = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{bmatrix} \qquad y \qquad \mathscr{B} = \begin{bmatrix} I & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ I & 0 & \dots & 0 \end{bmatrix}.$$

Entonces,

$$\begin{aligned} \mathscr{B}^* \mathscr{A}^p \mathscr{B} &= \begin{bmatrix} I & I & \dots & I \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} A_1^p & & 0 \\ & \ddots & \\ 0 & & A_m^p \end{bmatrix} \begin{bmatrix} I & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ I & 0 & \dots & 0 \end{bmatrix} \\ &= \begin{bmatrix} A_1^p + \dots + A_m^p & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix}. \end{aligned}$$

Así,

$$\left(\mathscr{B}^* \mathscr{A}^p \mathscr{B}\right)^{q/p} = \begin{bmatrix} (A_1^p + \dots + A_m^p)^{q/p} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix}.$$

Entonces, $\mathrm{Tr}\left[\left(\mathscr{B}^* \mathscr{A}^p \mathscr{B}\right)^{q/p}\right] = \mathrm{Tr}\left[\left(\sum_{j=1}^m A_j^p\right)^{q/p}\right] = \left(\Phi_{p,q}(A_1, \dots, A_m)\right)^q$. Por lo tanto, del Teorema 3.1.20, el lado izquierdo es cóncava (convexa) en $\mathcal{A}$ y el lado derecho es homogéneo de grado $q$. El resultado se sigue de la Proposición 3.1.6. □

**Lema 3.1.22.** *Para $p > 2$,*

$$\Phi_{p,q}(A_1, A_2, \ldots, A_m) = \left\| \left( \sum_{j=1}^{m} A_j{}^p \right)^{1/p} \right\|_q , \tag{3.12}$$

*no es convexa ni cóncava, para cualesquiera valores $p \neq q$.*

*Demostración.* Sean $A, B \in \mathbf{H}_n^+$. Consideremos

$$\Phi_{p,q}(tA, B) = \left\| (t^p A^p + B^p)^{1/p} \right\|_q .$$

Tenemos que

$$\Phi_{p,q}(tA, B)_{t=0} = \left( \left\| (t^p A^p + B^p)^{1/p} \right\|_q \right)_{t=0} = \|B\|_q$$

y por el Teorema A.3.4

$$\begin{aligned}
&\frac{d}{dt} \left[ Tr(t^p A^p + B^p)^{q/p} \right]^{\frac{1}{q}} \Big|_{t=0} \\
&\qquad = \frac{1}{q} \left[ Tr(t^p A^p + B^p)^{q/p} \right]^{\frac{1}{q}-1} \Big|_{t=0} \frac{d}{dt} \left( Tr(t^p A^p + B^p)^{q/p} \right) \Big|_{t=0} \\
&\qquad = \frac{1}{q} \|B\|_q^{1-q} \frac{d}{dt} \left( Tr \left[ t\, (t^{p-1} A^p) + B^p \right]^{q/p} \right) \Big|_{t=0} \\
&\qquad = \frac{1}{q} \|B\|_q^{1-q} Tr \left[ t^{p-1} A^p \left( \frac{q}{p} (t^p A^p + B^p)^{q/p-1} \right)_{t=0} \right] \\
&\qquad = \frac{1}{q} \|B\|_q^{1-q} \frac{q t^{p-1}}{p} Tr \left[ A^p B^{q-p} \right] = \frac{t^{p-1}}{p} \|B\|_q^{1-q} Tr \left[ A^p B^{q-p} \right] .
\end{aligned}$$

Así, la expansión de Taylor para $\Phi_{p,q}$ es

$$\Phi_{p,q}(tA, B) = \|B\|_q + \frac{t^p}{p} \|B\|_q^{1-q} Tr \left[ A^p B^{q-p} \right] + \mathcal{O}(t^{2p}).$$

Ahora, tomando $A$ como $A_1$, $A_2$ y $\frac{A_1+A_2}{2}$ tenemos que

$$\frac{1}{2}\Phi_{p,q}(tA_1, B) + \frac{1}{2}\Phi_{p,q}(tA_2, B) - \Phi_{p,q}\left( t\frac{A_1 + A_2}{2}, B \right) \tag{3.13}$$

$$\begin{aligned}
= \frac{t^p}{p} \|B\|_q^{1-q} \Big[ \frac{1}{2} Tr(A_1^p B^{q-p}) + \frac{1}{2} Tr(A_2^p B^{q-p}) \qquad (3.14) \\
- Tr\Big( \left( \frac{A_1 + A_2}{2} \right)^p B^{q-p} \Big) \Big] + \mathcal{O}(t^{2p}).
\end{aligned}$$

Ahora si $p > 2$, entonces la función $X \mapsto X^p$ no es convexa en operadores (ver Ejemplo V.5.4, pg. 147, en [1]). Entonces podemos encontrar $A_1, A_2 \in \mathbf{H}_n^+$ y un vector unitario $v$ tal que

$$\frac{1}{2}\langle v, A_1^p v\rangle + \frac{1}{2}\langle v, A_2^p v\rangle - \left\langle v, \left(\frac{A_1 + A_2}{2}\right)^p v \right\rangle < 0. \tag{3.15}$$

Sin embargo, como $X \mapsto \operatorname{Tr} X^p$ es convexo, por ser norma y cumplir con la Proposición 1.2.26, existe necesariamente otro vector unitario $w$ tal que si reemplazamos $v$ por $w$ la desigualdad en (3.15) se invierte.

Por lo tanto, para $q > p$, tomamos $B = |v\rangle\langle v|$, sustituyendo en (3.14) y recordando que $\operatorname{Tr}\,(X\,|v\rangle\langle v|) = \langle v, X\, v\rangle$. Entonces la cantidad a la derecha en (3.14) no es otra cosa que (3.15), y así el lado izquierdo de (3.14) es estrictamente negativo. Esto demuestra que $\Phi_{p,q}$ no puede ser convexo para tales $p, q$. Como (3.15) es un sumando de una traza y $\operatorname{Tr} X^p$ es convexa, entonces existe $w$ tal que cambiando $v$ por $w$, la expresión en (3.15) cambia de signo, así vemos que la concavidad no se cumple.

Para $q < p$, tomamos $B = (\varepsilon I + |v\rangle\langle v|)^{-1}$ para algún $\varepsilon$ pequeño, tal que $B^{q-p}$ es escencialmente la proyección ortogonal sobre $span(\{v\})$, entonces procedemos como anteriormente se hizo. $\square$

Una consecuencia de la concavidad/convexidad de $\Phi_{p,q}$ se obtiene escribiendo a la traza parcial como un promedio. Sea $A$ un operador positivo sobre $\mathcal{H}_1 \otimes \mathcal{H}_2$. Supongamos que la dimensión de $\mathcal{H}_2$ es $N$. Fijemos una base ortonormal para $\mathcal{H}_2$, digamos $\{e_1, e_2\, \dots, e_N\}$. Con respecto a ésta base definimos los operadores unitarios y autoadjuntos

$$\begin{aligned} U_{ij} &= I - |e_i\rangle\langle e_i| - |e_j\rangle\langle e_j| + |e_i\rangle\langle e_j| + |e_j\rangle\langle e_i| \\ V_{ij} &= I - 2|e_i\rangle\langle e_i|, \end{aligned}$$

con $i \neq j$. Observamos que $(U_{ij})$ es una matriz de permutación, la cual traspone las entradas $(i, j), (j, i)$. Por otro lado, $(V_{ij})$ es un operador que cambia el signo a la entrada $(i, i)$. En ambos casos tenemos operadores unitarios y autoadjuntos.

Sea $\mathcal{G}$ el subgrupo del grupo de operadores unitarios en $\mathcal{H}_2$ que es generado por la familia $\{I, (U_{ij}), (V_{ij})\}$. Cada operador $W$ en este grupo actúa mediante

$$W\, e_j = (-1)^{s(j)} e_{\sigma(j)},$$

donde $\sigma$ es una permutación sobre un conjunto de $N$ elementos y $s$ es alguna aplicación $s : \{1, 2, \ldots, N\} \to \{0, 1\}$. El orden del grupo es $\#(s) \cdot \#(S_N) = 2^N N!$.

**Proposición 3.1.23.** *Si $X \in \mathcal{B}(\mathcal{H})$ es tal que $XW = WX$ para todo $W \in \mathcal{G}$. Entonces $X = a\,I$.*

*Demostración.* Sea $\{e_1, e_2, \ldots, e_N\}$ base ortonormal de $\mathcal{H}_2$. Para $j$ fijo

$$X\,W\,e_j = X\left((-1)^{s(j)} e_{\sigma(j)}\right) = (-1)^{s(j)} x\,e_{\sigma(j)} = (-1)^{s(j)} \sum_{k'} (x_{k'\sigma(j)}) e_{k'}.$$

Por otra parte,

$$W X\,e_j = W \sum_k x_{kj} e_k = \sum_k x_{kj} W e_k = \sum_k x_{kj} (-1)^{s(k)} e_{\sigma(k)}.$$

Sea $k' = \sigma(k)$. Como $WX = XW$ comparamos

$$(-1)^{s(j)} \sum_k (x_{\sigma(k)\sigma(j)}) e_{\sigma(k)} - \sum_k x_{kj} (-1)^{s(k)} e_{\sigma(k)} = 0.$$

De donde, para toda $\sigma \in S_N$ y para toda $k$ tenemos que $(1)^{s(j)} x_{\sigma(k)\sigma(j)} = (-1)^{s(k)} x_{kj}$. Si $k$ es tal que $s(j) \neq s(k)$, entonces para toda $\sigma$ se tiene $-x_{\sigma(k)\sigma(j)} = x_{kj}$. Tomando $\sigma = id$, entonces $x_{kj} = 0$. Si $j = k$, entonces $\sigma(j) = \sigma(k)$ y en consecuencia $x_{\sigma(j)\sigma(j)} = x_{jj}$. Así $X = a\,I$. $\square$

**Lema 3.1.24.** *Sea $\mathcal{G}$ el grupo generado por $\{I, (U_{ij}), (V_{ij})\}$ como se definieron anteriormente. Sea $A \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$. Entonces*

$$\frac{1}{2^N N!} \sum_{W \in \mathcal{G}} (I \otimes W^*) A (I \otimes W) = \frac{1}{N} (Tr_2 A) \otimes I. \tag{3.16}$$

*Demostración.* Sea $e_i \otimes f_j \in \mathcal{H}_1 \otimes \mathcal{H}_2$, donde $e_i$, $f_j$ son elementos básicos;

entonces

$$\begin{aligned}
\frac{1}{2^N\,N!}&\Big\langle e_i\otimes f_j, \sum_{W\in\mathcal{G}}(I\otimes W^*)A(I\otimes W)e_i\otimes_j \Big\rangle\\
&= \frac{1}{2^N\,N!}\sum_{W\in\mathcal{G}}\Big\langle e_i\otimes W f_j, A(e_i\otimes W f_j)\Big\rangle\\
&= \frac{1}{2^N\,N!}\sum_{j}\Big\langle e_i\otimes(-1)^{s(j)}f_{\sigma(j)}, A\big(e_i\otimes(-1)^{s(j)}f_{\sigma(j)}\big)\Big\rangle\\
&= \frac{2^N\,(N-1!)}{2^N\,N!}\sum_{j}\big\langle e_i\otimes f_j, A\,(e_i\otimes f_j)\big\rangle = \frac{1}{N}\big\langle e_i, Tr_2A\,e_i\big\rangle\\
&= \frac{1}{N}\big\langle e_i\otimes f_j, Tr_2A\,e_i\otimes f_j\big\rangle = \frac{1}{N}\big\langle e_i\otimes f_j, Tr_2A\otimes I(\,e_i\otimes f_j)\big\rangle.
\end{aligned}$$

Para el caso general usamos la linealidad del producto interior y de la traza sobre elementos en el producto tensorial $h=\sum_{ij}\alpha_{ij}e_j\otimes f_j$. $\square$

**Teorema 3.1.25.** *Para* $1\le p\le 2$ *y* $q\ge 1$*,* $\Psi_{p,q}$ *es convexa sobre* $\mathbb{H}^+_{mn}$*. Para* $0\le p\le q\le 1$*,* $\Psi_{p,q}$ *es cóncava en* $\mathbb{H}^+_{mn}$*.*

*Demostración.* Primero veamos que para algún $X$ y $U$ operador unitario se tiene que $(U^*X\,U)^p=\underbrace{(U^*X\,U)(U^*X\,U)\cdots(U^*X\,U)}_{p}=U^*X^p\,U$. Ahora, usando la base que diagonaliza al operador $Tr_2A^p$, la traza normalizada y la identidad (3.16), tenemos

$$\begin{aligned}
\mathrm{Tr}_1\Big[(Tr_2A^p)^{q/p}\Big] &= \mathrm{Tr}\Big[\Big(\frac{1}{N^p}\mathrm{Tr}_2A^p\otimes I\Big)^{q/p}\Big] = \mathrm{Tr}\Big[\Big(\frac{N^{1-p}}{N}Tr_2A^p\otimes I\Big)^{q/p}\Big]\\
&= N^{\frac{q}{p}-1}\mathrm{Tr}\Big[\Big(\frac{1}{N}\mathrm{Tr}_2A^p\otimes I\Big)^{q/p}\Big]\\
&= N^{\frac{q}{p}-1}\mathrm{Tr}\Big[\frac{1}{2^N\,N!}\sum_{W\in\mathcal{G}}(I\otimes W^*)A^p(I\otimes W)\Big]^{q/p}\\
&= N^{\frac{q}{p}-1}\frac{1}{2^N\,N!}\mathrm{Tr}\Big[\sum_{W\in\mathcal{G}}\big((I\otimes W^*)A(I\otimes W)\big)^p\Big]^{q/p}\\
&= \frac{N^{\frac{q}{p}-1}}{2^N\,N!}\Big[\Phi_{p,q}\big(\{(I\otimes W^*)A(I\otimes W)\}_{W\in\mathcal{G}}\big)\Big]^q.
\end{aligned}$$

Por lo tanto del Teorema 3.1.21 se sigue el resultado. $\square$

## 3.2. Desigualdades tipo Minkowski

Usaremos las propiedadesd de concavidad/convexidad de los funcionales $\Phi_{p,q}$, $\Psi_{p,q}$ y $\Upsilon_{p,q}$ para obtener desigualdades del tipo Minkowski para la traza de operadores. A continuación presentamos la desigualdad de Minkowski para funciones medibles no-negativas sobre espacios producto de medida sigma finitos.

**Teorema 3.2.1 (Desigualdad de Hölder).** *Sean $p, q$ indices tales que $1 \leq p < \infty$ y $\frac{1}{p} + \frac{1}{q} = 1$. Sean $f \in \mathrm{L}^p(\Omega)$ y $g \in \mathrm{L}^q(\Omega)$. Entonces el producto puntual dado por $f \cdot g(x) = f(x)g(x)$ está en $\mathrm{L}^1(\Omega)$ y*

$$\left| \int_\Omega f \cdot g \, d\mu \right| \leq \int_\Omega |f||g| d\mu \leq \|f\|_p \|g\|_q. \tag{3.17}$$

**Teorema 3.2.2 (Desigualdad de Minkowski).** *Supongamos que $\Omega$ y $\Gamma$ son espacios $\sigma$-finitos de medidas $\mu$ y $\nu$ respectivamente. Sea $f$ una función no-negativa definida sobre $\Omega \times \Gamma$, $\mu \times \nu$-medible. Sea $1 \leq p < \infty$, entonces*

$$\left[ \int_\Omega \left[ \int_\Gamma f(x,y) \nu(dy) \right]^p \mu(dx) \right]^{1/p} \leq \int_\Gamma \left[ \int_\Omega f(x,y)^p \mu(dx) \right]^{1/p} \nu(dx). \tag{3.18}$$

*Demostración.* Consideremos

$$\int_\Omega f(x,y)^p \mu(dx), \qquad H(x) := \int_\Gamma f(x,y) \nu(dy).$$

Por el Teorema de Fubini y por ser $f$ $\mu \times \nu$-medible, las funciones anteriores son medibles.

Supongamos que $f > 0$ en un conjunto de $\mu \times \nu$-medida positiva. Escribimos el lado izquierdo de (3.18) como

$$\begin{aligned}
&\int_\Omega H(x)^p \mu(dx) \\
&= \int_\Omega \left[ \int_\Gamma f(x,y) \nu(dy) \right]^p \mu(dx) = \int_\Omega \left[ \int_\Gamma f(x,y) \nu(dy) \right]^{p-1} H(x) \mu(dx) \\
&= \int_\Omega \left[ \int_\Gamma f(x,y) \nu(dy) \right] H^{p-1}(x) \mu(dx) = \int_\Gamma \left[ \int_\Omega f(x,y) H^{p-1} \mu(dx) \right] (x) \nu(dy).
\end{aligned}$$

Por la desigualdad de Hölder tenemos que

$$\int_\Omega H(x)^p \mu(dx) \leq \int_\Gamma \left[\int_\Omega f(x,y)^p \mu(dx)\right]^{1/p} \left[\int_\Omega H(x)^p \mu(dx)\right]^{\frac{p-1}{p}} \nu(dy).$$

Entonces

$$\begin{aligned}
&\left[\int_\Omega \left[\int_\Gamma f(x,y)\nu(dy)\right]^p \mu(dx)\right]^{1/p} \\
&\quad = \left[\int_\Omega H(x)^p \mu(dx)\right]^{1/p} = \left[\int_\Omega H(x)^p \mu(dx)\right]^{1-\frac{p-1}{p}} \\
&\quad = \frac{\int_\Omega H(x)^p \mu(dx)}{\left[\int_\Omega H(x)^p \mu(dx)\right]^{\frac{p-1}{p}}} \leq \int_\Gamma \left[\int_\Omega f(x,y)^p \mu(dx)\right]^{1/p} \nu(dx).
\end{aligned}$$

□

**Teorema 3.2.3.** *Sean $(X, d\mu), (Y, d\nu)$ espacios de medida sigma-finitos. Entonces, para cualquier función medible no-negativa $f$ definida sobre $X \times Y$ y $1 \leq q \leq p$, se tiene que*

$$\left[\int_Y \left[\int_X f^q(x,y) d\mu\right]^{p/q} d\nu\right]^{1/p} \leq \left[\int_X \left[\int_Y f^p(x,y) d\nu\right]^{q/p} d\mu\right]^{1/q}. \tag{3.19}$$

*Demostración.* Escribimos

$$\left[\int_Y \left[\int_X f^p(x,y) d\mu\right]^{p/q} d\nu\right]^{q/p}.$$

Tomando $s = \frac{p}{q}$, aplicamos el Teorema anterior

$$\left[\int_Y \left[\int_X f^q(x,y) d\mu\right]^{s} d\nu\right]^{1/s} \leq \int_X \left[\int_Y f^{sq}(x,y) d\nu\right]^{1/s} d\mu.$$

Por lo tanto

$$\left[\int_Y \left[\int_X f^q(x,y) d\mu\right]^{p/q} d\nu\right]^{q/p} \leq \int_X \left[\int_Y f^p(x,y) d\nu\right]^{q/p} d\mu.$$

□

En esta parte el símbolo $\mathrm{Tr}_1$ también denotará la traza usual para operadores sobre $\mathcal{H}_1$.

**Teorema 3.2.4 (Desigualdad Tipo Minkowski I).** *Sea $A \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ positivo. Entonces, para toda $p \geq 1$*

$$\big(\mathrm{Tr}_{\,2}(\mathrm{Tr}_{\,1} A)^p\big)^{1/p} \leq \mathrm{Tr}_{\,1}(\mathrm{Tr}_{\,2} A^p)^{1/p}. \tag{3.20}$$

*Si $0 < p \leq 1$, la desigualdad se invierte.*

*Demostración.* Sea $A \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ positivo. Demostraremos primero que existe un operador positivo $B \in \mathcal{B}(\mathcal{H}_2)$ tal que $\big(\mathrm{Tr}_{\,2}(\mathrm{Tr}_{\,1} A)^p\big)^{1/p} = \mathrm{Tr}_{\,2}\big(B \mathrm{Tr}_{\,1} A\big)$.

Sea $\{\beta_j\} \subset \mathcal{H}_2$ la base que diagonaliza a $Tr_1 A$ y defínase

$$B = \sum_j b_j |\beta_j\rangle\langle\beta_j|,$$

donde $b_j = \alpha\, \lambda_j$ y $\alpha = (\sum_j \lambda_j^q)^{-1/q}$, con $\frac{1}{p} + \frac{1}{q} = 1$, donde $\lambda_j$ son los valores propios de $\mathrm{Tr}_{\,1} A$. Entonces

$$\begin{aligned}
(\mathrm{Tr}\ B^q)^{1/q} &= \Big(\mathrm{Tr}\ \sum_j b_j^q |\beta_j\rangle\langle\beta_j|\Big)^{1/q} = \Big(\sum_j b_j^q \mathrm{Tr}\ |\beta_j\rangle\langle\beta_j|\Big)^{1/q} \\
&= \Big(\sum_j b_j^q\Big)^{1/q} = \Big(\sum_j (\alpha\lambda_j)^q\Big)^{1/q} \\
&= \alpha\Big(\sum_j \lambda_j^q\Big)^{1/q} = 1.
\end{aligned}$$

Luego,

$$\begin{aligned}
\big(\mathrm{Tr}_{\,2}(\mathrm{Tr}_{\,1} A)^p\big)^{1/p} &= \big(\mathrm{Tr}_{\,2} \sum_j \lambda_j^p |\beta_j\rangle\langle\beta_j|\big)^{1/p} = \Big(\sum_j \langle\beta_j, \lambda_j^p \beta_j\rangle\Big)^{1/p} \\
&= \Big(\sum_j \lambda_j^p\Big)^{1/p} = \Big(\sum_j \lambda_j^p\Big)^{1/p}\Big(\sum_j b_j^q\Big)^{1/q} \\
&= \sum_j b_j \lambda_j.
\end{aligned}$$

Esta última relación por la condición para igualdad en Hölder. Por otro lado

$$\begin{aligned}\sum_j b_j\lambda_j &= \sum_j b_j\lambda_j\langle\beta_j,\beta_j\rangle = \sum_j b_j\langle\beta_j, \mathrm{Tr}_1 A\beta_j\rangle \\ &= \sum_j \langle B\beta_j, \mathrm{Tr}_1 A\beta_j\rangle = \sum_j \langle\beta_j, B\,\mathrm{Tr}_1 A\beta_j\rangle = \mathrm{Tr}_2\big(B\mathrm{Tr}_1 A\big).\end{aligned}$$

Así, $\big(\mathrm{Tr}_2(\mathrm{Tr}_1 A)^p\big)^{1/p} = \mathrm{Tr}_2\big(B\mathrm{Tr}_1 A\big)$. Ahora, de la definición de traza parcial A.4.1, tenemos

$$\begin{aligned}\mathrm{Tr}_2\big(B\,\mathrm{Tr}_1 A\big) &= \sum_j\sum_i \langle\alpha_i\otimes B\beta_j, A\,\alpha_i\otimes\beta_j\rangle \\ &= \sum_{ij}\big\langle(I\otimes B)(\alpha_i\otimes\beta_j), A\,\alpha_i\otimes\beta_j\big\rangle \\ &= \sum_{ij}\big\langle\alpha_i\otimes\beta_j, (I\otimes B)A\,\alpha_i\otimes\beta_j\big\rangle = \mathrm{Tr}_{1,2}((I\otimes B)A).\end{aligned}$$

De las identidades anteriores tenemos que

$$\big(\mathrm{Tr}_2(\mathrm{Tr}_1 A)^p\big)^{1/p} = \mathrm{Tr}_2\big(B\mathrm{Tr}_1 A\big) = \mathrm{Tr}_{1,2}((I\otimes B)A) \tag{3.21}$$

$$= \sum_{ij}\langle\alpha_i\otimes B\beta_j, A\,\alpha_i\otimes\beta_j\rangle \tag{3.22}$$

$$= \sum_{ij}\langle\alpha_i\otimes b_j\beta_j, A\,\alpha_i\otimes\beta_j\rangle \tag{3.23}$$

$$= \sum_{ij} b_j\langle\alpha_i\otimes\beta_j, A\,\alpha_i\otimes\beta_j\rangle \tag{3.24}$$

$$\le \sum_i\Big(\sum_j b_j^q\Big)^{1/q}\Big(\sum_j\big\langle\alpha_i\otimes\beta_j, A(\alpha_i\otimes\beta_j)\big\rangle^p\Big)^{1/p} \tag{3.25}$$

$$\le \sum_i\Big(\sum_j\big\langle\alpha_i\otimes\beta_j, A(\alpha_i\otimes\beta_j)\big\rangle^p\Big)^{1/p}. \tag{3.26}$$

Ahora, veamos que para cada $i, j$ se cumple la siguiente desigualdad

$$\langle\alpha_i\otimes\beta_j, A\,\alpha_i\otimes\beta_j\rangle \le \langle\alpha_i\otimes\beta_j, A^p\,\alpha_i\otimes\beta_j\rangle^{1/p}.$$

Usando el Teorema espectral, escribimos $A = \sum_{kl} \lambda_{kl} |\alpha_{kl}\rangle\langle\alpha_{kl}|$. Entonces

$$\begin{aligned}
\langle \alpha_i \otimes \beta_j, A\, \alpha_i \otimes \beta_j \rangle &= \Big\langle \alpha_i \otimes \beta_j, \sum_{kl} \lambda_{kl} |\alpha_{kl}\rangle\langle\alpha_{kl}|\, \alpha_i \otimes \beta_j \Big\rangle \\
&= \sum_{kl} \lambda_{kl} |\langle \alpha_i \otimes \beta_j, \alpha_{kl}\rangle|^2 \cdot 1 \leq \Big( \sum_{kl} \lambda_{kl}^p |\langle \alpha_i \otimes \beta_j, \alpha_{kl}\rangle|^{2p} \Big)^{1/p} \\
&= \Big( \sum_{kl} \lambda_{kl}^p \big|\langle \alpha_i \otimes \beta_j, \alpha_{kl}\rangle\big| \big|\langle \alpha_i \otimes \beta_j, \alpha_{kl}\rangle\big|^{2p-1} \Big)^{1/p} \\
&\leq \Big( \sum_{kl} \lambda_{kl}^p \big|\langle \alpha_i \otimes \beta_j, \alpha_{kl}\rangle\big| \|\alpha_{kl}\|^{2p-1} \, \|\alpha_i \otimes \beta_j\|^{2p-1} \Big)^{1/p} \\
&= \Big( \sum_{kl} \lambda_{kl}^p \langle \alpha_i \otimes \beta_j, |\alpha_{kl}\rangle\langle\alpha_{kl}|\alpha_i \otimes \beta_j \rangle \Big)^{1/p} \\
&= \Big( \Big\langle \alpha_i \otimes \beta_j, \sum_{kl} \lambda_{kl}^p |\alpha_{kl}\rangle\langle\alpha_{kl}|\alpha_i \otimes \beta_j \Big\rangle \Big)^{1/p} \\
&= \big( \langle \alpha_i \otimes \beta_j, A^p \alpha_i \otimes \beta_j \rangle \big)^{1/p}.
\end{aligned}$$

Usando esto en (3.26) tenemos

$$\begin{aligned}
\sum_i \Big( \sum_j \big\langle \alpha_i \otimes \beta_j, A(\alpha_i \otimes \beta_j) \big\rangle^p \Big)^{1/p} &\leq \sum_i \Big( \sum_j \big\langle \alpha_i \otimes \beta_j, A^p(\alpha_i \otimes \beta_j) \big\rangle \Big)^{1/p} \\
&= \sum_i \Big[ \langle \alpha_i, \mathrm{Tr}_2 A^p \alpha_i \rangle \Big]^{1/p} \\
&= \sum_i \langle \alpha_i, (\mathrm{Tr}_2 A^p)^{1/p} \alpha_i \rangle = \mathrm{Tr}_1 (\mathrm{Tr}_2 A^p)^{1/p}
\end{aligned}$$

donde $\{\alpha_i\}$ es la base de vectores propios de $Tr_2\, A^p$. Así, siguiendo la cadena de desigualdades se sigue que

$$\big(\mathrm{Tr}_2 (\mathrm{Tr}_1\, A)^p\big)^{1/p} \leq \mathrm{Tr}_1 (\mathrm{Tr}_2\, A^p)^{1/p}.$$

Supongamos ahora que $0 < p \leq 1$. Definimos $r = 1/p$ y sea $B = A^p$, de donde $A = B^r$. Como $r > 1$, aplicamos la desigualdad probada e intercambiando el orden de las trazas tenemos

$$(\mathrm{Tr}_1 (\mathrm{Tr}_2 B)^r)^{1/r} \leq \mathrm{Tr}_2 (\mathrm{Tr}_1 B^r)^{1/r}$$

Sustituyendo $B = A^p$, se tiene que

$$(\mathrm{Tr}_1(\mathrm{Tr}_2 A^p)^{1/p})^p \leq \mathrm{Tr}_2(\mathrm{Tr}_1 A)^p.$$

□

**Teorema 3.2.5 (Desigualdad Tipo Minkowski II).** *Para $1 \leq p \leq 2$ y todo $A$ operador positivo sobre $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$,*

$$\mathrm{Tr}_3(\mathrm{Tr}_2(\mathrm{Tr}_1 A^p))^{1/p} \leq \mathrm{Tr}_{1,3}[(\mathrm{Tr}_2 A^p)^{1/p}]. \tag{3.27}$$

*Para $0 < p < 1$,la desigualdad se invierte.*

*Demostración.* Supongamos que la dimensión de $\mathcal{H}_1$ es $N$. Tomando la traza normalizada y la base de valores propios de $Tr_1 A^p$, el lado izquierdo de 3.27 lo podemos escribir en términos de la definición 3.1.2

$$\begin{aligned}\mathrm{Tr}_3(\mathrm{Tr}_2(\mathrm{Tr}_1 A^p))^{1/p} &= \mathrm{Tr}_{1,3}\Big[\mathrm{Tr}_2\Big(\frac{1}{N} I_{\mathcal{H}_1} \otimes \mathrm{Tr}_1 A\Big)^p\Big]^{1/p} \\ &= \Psi_{p,1}\Big(\frac{1}{N} I_{\mathcal{H}_1} \otimes \mathrm{Tr}_1 A\Big),\end{aligned}$$

donde la pareja de espacios para aplicar $\Psi_{p,1}$ son $\mathcal{H}_2$ y $\mathcal{H}_1 \otimes \mathcal{H}_3$. Luego

$$\begin{aligned}&\Psi_{p,1}\Big(\frac{1}{N} I_{\mathcal{H}_1} \otimes Tr_1 A\Big) \\ &\qquad = \Psi_{p,1}\Big(\frac{1}{2^N\, N!}\sum_{W\in\mathcal{G}}(W^* \otimes I_{\mathcal{H}_2})A(W^* \otimes I_{\mathcal{H}_2})\Big) \\ &\qquad \leq \frac{1}{2^N\, N!}\sum_{W\in\mathcal{G}} \Psi_{p,1}\Big((W^* \otimes I_{\mathcal{H}_2})A(W^* \otimes I_{\mathcal{H}_2})\Big) \\ &\qquad = \frac{1}{2^N\, N!}\sum_{W\in\mathcal{G}} Tr_{1,3}\Big(Tr_2\big((W^* \otimes I_{\mathcal{H}_2})A^p(W^* \otimes I_{\mathcal{H}_2})\big)\Big)^{1/p} \\ &\qquad = \frac{1}{2^N\, N!}\sum_{W\in\mathcal{G}} Tr_{1,3}\Big((W^* \otimes I_{\mathcal{H}_2})Tr_2 A^p(W^* \otimes I_{\mathcal{H}_2})\Big)^{1/p} \\ &\qquad = \frac{1}{2^N\, N!}\sum_{W\in\mathcal{G}} Tr_{1,3}\Big((W^* \otimes I_{\mathcal{H}_2})(Tr_2 A^p)^{1/p}(W^* \otimes I_{\mathcal{H}_2})\Big) \\ &\qquad = Tr_3 Tr_1 (Tr_2 A^p)^{1/p}.\end{aligned}$$

Lo cual demuestra el Teorema. □

## 3.3. Subaditividad fuerte de la entropía de von Neumann

El propósito de esta sección es demostrar la subaditividad fuerte de la entropía de von Neumann.

Consideremos una función $f(p) = \rho^p$, para un estado $\rho$. podemos escribir $\rho = \sum_j \lambda_j^p |e_j\rangle\langle e_j|$, entonces

$$\begin{aligned}\frac{d}{dp}(\mathrm{Tr}\,\rho^p)\Big|_{p=1} &= \frac{d}{dp}\Big[\mathrm{Tr}\Big(\sum_j \lambda_j^p |e_j\rangle\langle e_j|\Big)\Big]_{p=1} = \frac{d}{dp}\Big[\sum_j \lambda_j^p \mathrm{Tr}\left(|e_j\rangle\langle e_j|\right)\Big]_{p=1}\\ &= \frac{d}{dp}\Big[\sum_j \lambda_j^p\Big]_{p=1} = \sum_j \lambda_j \log \lambda_j = -S(\rho).\end{aligned}$$

Consideremos $f(x) = a^{x+1}$, y escribamos su expansión en serie de McLaurin, es decir

$$f(x) = a + xa\ln a + a(\ln a)^2 \frac{x^2}{2!} + \dots.$$

Si tomamos ahora un operador positivo $A$ y $\varepsilon > 0$ suficientemente pequeño, tenemos

$$A^{1\pm\varepsilon} = A \pm \varepsilon\, A\log\, A \pm \varepsilon^2 \frac{A(\log\, A)^2}{2!} \pm \dots \tag{3.28}$$

**Teorema 3.3.1.** *(Subaditividad fuerte de la entropía de von Neumann) Sea $\rho_{123}$ un estado sobre el espacio de Hilbert de dimensión finita $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$. Entonces*

$$S(\rho_{13}) + S(\rho_{23}) \geq S(\rho_{123}) + S(\rho_3),$$

*donde $S$ es la entropía de von Neumann.*

*Demostración.* Tomando $A = \rho_{123} \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3)$. Aplicamos traza parcial sobre el segundo factor en (3.28),

$$\mathrm{Tr}_2(\rho^{1+\varepsilon}) = \mathrm{Tr}_2\rho + \varepsilon\,\mathrm{Tr}_2\rho\log\,\rho + \mathcal{O}(\varepsilon^2) = \rho_{1,3} + \varepsilon\,\mathrm{Tr}_2\rho\log\,\rho + \mathcal{O}(\varepsilon^2).$$

Por otro lado, sabemos que $\frac{1}{1+\varepsilon} = \sum_{n\geq 0}(-\varepsilon)^n = 1 - \varepsilon + \mathcal{O}(\varepsilon^2)$. Así

$$\begin{aligned}\Big[\mathrm{Tr}_2(\rho^{1+\varepsilon})\Big]^{\frac{1}{1+\varepsilon}} =& \Big[\rho_{1,3} + \varepsilon\,\mathrm{Tr}_2\rho\log\,\rho + \mathcal{O}(\varepsilon^2)\Big]^{1-\varepsilon+\mathcal{O}(\varepsilon^2)}\\ =&\rho_{1,3} + \varepsilon\,\mathrm{Tr}_2\rho\log\,\rho\\ &- \varepsilon\left(\rho_{1,3} + \varepsilon\,\mathrm{Tr}_2\rho\log\,\rho\right)\log\left(\rho_{1,3} + \varepsilon\,\mathrm{Tr}_2\rho\log\,\rho\right) + \mathcal{O}(\varepsilon^2)\\ =&\rho_{1,3} + \varepsilon\mathrm{Tr}_2\rho\log\rho - \varepsilon\rho_{1,3}\log\rho_{1,3} + \mathcal{O}(\varepsilon^2),\end{aligned}$$

tomando $\mathrm{Tr}_1$ y $\mathrm{Tr}_3$ obtenemos que.

$$\mathrm{Tr}_3\,\mathrm{Tr}_1\Big[\mathrm{Tr}_2(\rho^{1+\varepsilon})\Big]^{\frac{1}{1+\varepsilon}} = 1-\varepsilon S(\rho)+\varepsilon S(\rho_{1,3})+\mathcal{O}(\varepsilon^2).$$

Consideremos ahora $\mathrm{Tr}_1\rho = \rho_{2,3}$, así como se hizo anteriormente, tenemos que

$$\rho_{2,3}^{1+\varepsilon} = \rho_{2,3}+\varepsilon\rho_{2,3}\log\rho_{2,3}+\mathcal{O}(\varepsilon^2).$$

Luego

$$\mathrm{Tr}_3\Big[\mathrm{Tr}_2\big(\rho_{2,3}^{1+\varepsilon}\big)\Big]^{\frac{1}{1+\varepsilon}} = 1-\varepsilon S(\rho_{2,3})+\varepsilon S(\rho_3)+\mathcal{O}(\varepsilon^2).$$

Ahora, aplicando el Teorema 3.2.4, tomando $q=1$ y $p=1+\varepsilon$, tenemos

$$1-\varepsilon S(\rho)+\varepsilon S(\rho_{1,3})+\mathcal{O}(\varepsilon^2)\le 1-\varepsilon S(\rho_{2,3})+\varepsilon S(\rho_3)+\mathcal{O}(\varepsilon^2)$$

Equivalentemente

$$1-S(\rho)+S(\rho_{1,3})+\frac{\mathcal{O}(\varepsilon^2)}{\varepsilon}\le 1-S(\rho_{2,3})+S(\rho_3)+\frac{\mathcal{O}(\varepsilon^2)}{\varepsilon}.$$

Haciendo $\varepsilon\to 0$, obtenemos la subaditividad fuerte para la entropía de von Neumann. $\square$

# Apéndice A

# Apéndice

## A.1. El producto tensorial de espacios de Hilbert

La mayor parte de los resultados que presentamos en este apéndice se encuentran demostrados en textos sobre espacios de Hilbert. En general no incluimos demostraciones, éstas se pueden encontrar en por ejemplo [12] o [13] .

**Teorema A.1.1.** *Sean $\mathcal{V}, \mathcal{W}$ espacios vectoriales sobre un campo $\mathbb{K}$, tales que $\dim(\mathcal{V}) = n$ y $\dim(\mathcal{W}) = m$. Existe un espacio $\mathcal{V} \otimes \mathcal{W}$ de dimensión $nm$ y una función bilineal $\tau : \mathcal{V} \times \mathcal{W} \to \mathcal{V} \otimes \mathcal{W}$, tal que $(v, w) \mapsto \tau(v, w) = v \otimes w$ satisface las siguientes relaciones*

*(i) El producto de $v \in \mathcal{V}$ con $w \in \mathcal{W}$ se denota por $v \otimes w$ y satisface*

$$\begin{aligned} (v_1 + v_2) \otimes w &= v_1 w \otimes v_2 w && \forall\, v_1, v_2 \in \mathcal{V}, \forall w \in \mathcal{W} \\ v \otimes (w_1 + w_2) &= v \otimes w_1 + v \otimes w_2 && \forall\, w_1, w_2 \in \mathcal{W}, \forall\, v \in \mathcal{V} \\ c(v \otimes w) &= (c\, v) \otimes w = v \otimes (c\, w) && c \in \mathbb{K} \end{aligned}$$

*(ii) Todo elemento de $\mathcal{V} \otimes \mathcal{W}$ puede escribirse como*

$$\sum_i c_i v_i \otimes w_i,$$

*para algunos escalares $c_i$ y algunos $v_i \in \mathcal{V}$, $w_i \in \mathcal{W}$.*

*(iii) Si $\{v_i\}_{i=1}^n$ y $\{w_j\}_{j=1}^m$ son bases de $\mathcal{V}$ y $\mathcal{W}$ respectivamente, entonces $\{v_i \otimes w_j\}_{ij}$ forma una base para $\mathcal{V} \otimes \mathcal{W}$.*

*(iv) (Propiedad universal) Si $\mathcal{U}$ es un espacio vectorial sobre $\mathbb{K}$ y $\mathscr{L} : \mathcal{V} \times \mathcal{W} \to \mathcal{U}$ es una función bilineal. Entonces existe una única aplicación lineal $L : \mathcal{V} \otimes \mathcal{W} \to \mathcal{U}$ tal que*

$$\mathscr{L} = L \circ \tau.$$

*Es decir, $\mathscr{L}(v, w) = L(v \otimes w)$, para $v \in \mathcal{V}$ y $w \in \mathcal{W}$.*

Sean $A : \mathcal{V} \to \mathcal{V}$ y $B : \mathcal{W} \to \mathcal{W}$ operadores lineales. Definimos $\mathscr{L} : \mathcal{V} \times \mathcal{W} \to \mathcal{V} \otimes \mathcal{W}$ como $\mathscr{L}(v, w) = Av \otimes Bw$, es una función bilineal, entonces existe una única función lineal $L : \mathcal{V} \otimes \mathcal{W} \to \mathcal{V} \otimes \mathcal{W}$ tal que $L(v \otimes w) = \mathscr{L}(v, w) = Av \otimes Bw$. Denotamos a $L$ como $A \otimes B$ es decir, es el único operador lineal tal que para todo $v \in \mathcal{V}$ y $w \in \mathcal{W}$

$$(A \otimes B)(v \otimes w) = Av \otimes Bw.$$

La generalización para un operador $L$ que actúa sobre el espacio $\mathcal{H}_1 \otimes \mathcal{H}_2$, se puede consultar en [12].

Cuando $(\mathcal{H}_\infty, \langle \cdot, \cdot \rangle_1)$ y $(\mathcal{H}_\in, \langle \cdot, \cdot \rangle_2)$ 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 en cada espacio se define un producto interno en el espacio $\mathcal{H}_1 \otimes \mathcal{H}_2$ como sigue: Sean $\{v_1, \ldots, v_n\}$ y $\{w_1, \ldots, w_m\}$ 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 los elementos básicos, entonces se define el producto interno de $v \otimes w$ y $v' \otimes w'$ como

$$\begin{aligned} \langle v \otimes w, v' \otimes w' \rangle_\otimes &= \left\langle \sum_{ij} x_i y_j (v_i \otimes w_j), \sum_{rs} x'_r y'_s (v_r \otimes w_s) \right\rangle_\otimes \\ &= \sum_{i,j,r,s} \overline{x_i y_j} x'_r y'_s \langle v_i, v_r \rangle_1 \langle w_j, w_s \rangle_2. \end{aligned}$$

En particular $\langle \alpha(v \otimes w), \beta(v' \otimes w') \rangle_\otimes = \overline{\alpha}\beta \langle v, v' \rangle_1 \langle w, w' \rangle_2$. Se puede verificar que esta función satisface los axiomas de un producto interior.

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 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 de $\mathcal{H}_1 \otimes \mathcal{H}_2$. La construcción anterior se generaliza cuando aparecen mas de dos espacios Hilbert, veáse [12].

**Proposición A.1.2.** *Sean $A \in \mathcal{B}(\mathcal{H})$ y $B \in \mathcal{B}(\mathcal{K})$. Entonces*

$$\operatorname{Tr}(A \otimes B) = \operatorname{Tr} A \operatorname{Tr} B \tag{A.1}$$

*Demostración.* Sea $\{u_i \otimes v_j\}$ una base para el espacio $\mathcal{H} \otimes \mathcal{K}$. Entonces

$$\begin{aligned}\operatorname{Tr}(A \otimes B) &= \sum_{ij} \langle u_i \otimes v_j, (A \otimes B) u_i \otimes v_j \rangle = \sum_{ij} \langle u_i \otimes v_j, A\, u_i \otimes B\, v_j \rangle \\ &= \sum_{ij} \langle u_i, A\, u_i \rangle \langle v_j, B\, v_j \rangle = \operatorname{Tr}\, A \operatorname{Tr}\, B.\end{aligned}$$

□

## A.2. Operadores y matrices positivas

**Definición A.2.1.** *Sea $\mathcal{H}$ un espacio de Hilbert y $T : \mathcal{H} \to \mathcal{H}$ un operador lineal acotado. $T$ es positivo (o positivo semi-definido) si para todo $h \in \mathcal{H}$ se tiene que $\langle Th, h \rangle \geq 0$*

Se sigue de la definición que si un operador es positivo es autoadjunto: Sea $T \in \mathcal{B}(\mathcal{H})$ operador positivo; nótese que en general $\langle h, Th \rangle$ y $\langle Th, h \rangle$ son complejos conjugados, como cumplen la Definición A.2.1, entonces $\langle Th, h \rangle$ es real. Luego

$$\langle T^* h, h \rangle = \langle h, Th \rangle = \overline{\langle h, Th \rangle} = \langle Th, h \rangle \qquad \forall\, h \in \mathcal{H}\ .$$

Entonces, $(T^* - T)h = 0$ para toda $h \in \mathcal{H}$, por tanto $T^* = T$.

Si $T_1, T_2$ son operadores positivos, por la linealidad del producto interior $T_1 + T_2$ es también un operador positivo.

Sean $\mathcal{H}$ y $\mathcal{K}$ espacios de Hilbert de dimensión $n$ y $k$ respectivamente. Si fijamos bases en cada espacio entonces $\mathcal{B}(\mathcal{H})$ puede ser identificado por un álgebra de matrices $\mathbf{M}_n(\mathbb{C})$, análogamente para $\mathcal{B}(\mathcal{K})$ es $\mathbf{M}_k(\mathbb{C})$.

**Teorema A.2.2.** *Sea $T \in \mathcal{B}(\mathcal{H})$. Las siguientes condiciones son equivalentes.*

*(1) $T$ es positivo.*

*(2) $T = T^*$ y el espectro de $T$ está contenido en $\mathbb{R}^+$.*

*(3) $T$ es de la forma $A^* A$, para algún operador $A \in \mathcal{B}(\mathcal{H})$.*

Siguiendo la referencia [15] tenemos

**Lema A.2.3.** *La serie de potencias para $\sqrt{1-z}$ alrededor del cero, converge absolutamente para todos los números complejos $z$ tales que $|z| \leq 1$.*

*Demostración.* Sea $S = 1 + c_1 z + c_2 z^2 + \dots$ la serie de potencias de $f(z) = \sqrt{1-z}$ alrededor de cero, donde $c_k = \frac{1}{k!} f^{(k)}(0)$. Como $\sqrt{1-z}$ es analítica para $|z| < 1$, la serie converge absolutamente para $z$ tales que $|z| < 1$. Las derivadas de $f(z)$ en el origen son negativas, luego $c_k$ también para $k \geq 1$. Entonces

$$\sum_{k=0}^{N} |c_k| = 2 - \sum_{k=0}^{N} c_k = 2 - \lim_{x \to 1^-} \sum_{k=0}^{N} c_k x^k \leq 2 - \lim_{x \to 1^-} \sqrt{1-x} = 2.$$

Como esto es cierto para toda $N$, se sigue que $\sum_{k=0}^{\infty} |c_k| \leq 2$, lo cual implica que la serie converge absolutamente para $|z| = 1$. $\square$

**Lema A.2.4 (de la raíz cuadrada).** *Sea $A \in \mathcal{B}(\mathcal{H})$ y $A \geq 0$. Entonces, existe un único $B \in \mathcal{B}(\mathcal{H})$ con $B \geq 0$ y $B^2 = A$. Además, $B$ con cualquier operador que conmute con $A$.*

*Demostración.* Es suficiente considerar el caso $\|A\| \leq 1$. Dado que

$$\|I - A\| = \sup_{\|\varphi\|=1} |\langle (I - A)\varphi, \varphi \rangle| \leq 1,$$

el Lema anterior nos dice que la serie $1 + c_1(I - A) + c_2(I - A)^2 + \dots$ converge absolutamente a un operador $B$, en la norma de operadores. Por la convergencia absoluta de la serie podemos elevar al cuadrado y reacomodar los términos, esto prueba que $B^2 = A$. Más aún, como $0 \leq I - A \leq I$ tenemos que $0 \leq \langle \varphi, (I - A)^n \varphi \rangle \leq 1$ para toda $\varphi \in \mathcal{H}$ con $\|\varphi\| = 1$. Luego

$$\langle \varphi, B\varphi \rangle = 1 + \sum_{k=1}^{\infty} c_k \langle \varphi, (I - A)^k \varphi \rangle \geq 1 + \sum_{k=1}^{\infty} c_k = 0,$$

hemos usado que $c_k < 0$ y la estimación como en el Lema anterior. Así, $B \geq 0$. Como la serie de $B$ converge absolutamente, entonces conmuta con cualquier operador que conmute con $A$.

Supongamos que existe $B' \in \mathcal{B}(\mathcal{H})$, con $B' \geq 0$ y $(B')^2 = A$. Dado que

$$B'A = (B')^3 = AB',$$

$B'$ conmuta con $A$, así también con $B$. Se sigue que

$$(B - B')B(B - B') + (B - B')B'(B - B') = (B^2 - B'^2)(B - B') = 0.$$

Ya que ambos términos son positivos tenemos que deben ser cero, así la diferencia $(B - B')^3 = 0$. Como $B - B'$ es autoadjunto, por ser positivo, tenemos $\|(B - B')^4\| = \|(B - B')^4\| = 0$, por lo tanto $B - B' \equiv 0$. $\square$

Recordemos que $s(A) = (s_1(A), s_2(A), \ldots, s_n(A))$ denota al vector de valores singulares de $A$ ordenados de manera decreciente: $s_1(A) \geq s_2(A) \geq \cdots \geq s_n(A)$.

**Proposición A.2.5.** *Sea $A$ operador lineal en $\mathbb{C}^n$. Sean $\{s_i(A)\}_{i=1}^n$ los valores singulares de $A$ ordenados en forma decreciente, entonces*

$$\|A\| = \| \, |A| \, \| = s_1(A). \tag{A.2}$$

*Demostración.* Sea $x \in \mathbb{C}^n$ calculamos

$$\| \, |A|x \, \|^2 = \langle |A|x, |A|x \rangle = \langle |A|^2 x, x \rangle = \langle A^* Ax, x \rangle = \langle Ax, Ax \rangle = \|Ax\|^2;$$

tomando supremo sobre las $x \in \mathbb{C}^n$ tales que $\|x\| = 1$ tenemos $\|A\| = \| \, |A| \, \|$. Ahora, $|A|$ es operador autoadjunto, aplicamos el Teorema espectral en dimensión finita se obtiene que

$$|A| = \sum_{j=1}^{n} s_j(A) \langle \cdot, \phi_j \rangle \phi_j,$$

donde $\{\phi_j\}_j \subset \mathbb{C}^n$ es la base ortonormal que diagonaliza a $|A|$. Usando la identidad de Parseval se tiene que

$$\begin{aligned}
\|Ax\|^2 = \| \, |A|x \, \|^2 &= \left\| \sum_{i=1}^{n} s_i(A) \langle x, \phi_i \rangle \phi_i \right\|^2 = \sum_{i=1}^{n} |s_i(A)|^2 |\langle x, \phi_i \rangle|^2 \\
&\leq |s_1(A)|^2 \sum_{i=1}^{n} |\langle x, \phi_i \rangle|^2 = |s_1(A)|^2 \|x\|^2;
\end{aligned}$$

de donde $\|A\| \leq |s_1(A)|$. Por otro lado, sea $\phi_k$ vector propio de $|A|$, entonces

$$\|A\phi_k\| = \| \, |A|\phi_k\| = \|s_k(A)\phi_k\| = |s_k(A)| \, \|\phi_k\| = |s_k(A)|.$$

Así $|s_k(A)| \leq \sup \|Ax\| = \|A\|$ para toda $k = 1, 2, \ldots, n$. Por lo tanto $\|A\| = |s_1(A)|$. □

**Definición A.2.6.** *Sean $\mathcal{H}, \mathcal{K}$ espacios de Hilbert con producto interior $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ y $\langle \cdot, \cdot \rangle_{\mathcal{K}}$, respectivamente. El producto interior en la suma directa $\mathcal{H} \oplus \mathcal{K}$ está dado por*

$$\left\langle \begin{bmatrix} h \\ h' \end{bmatrix}, \begin{bmatrix} u \\ u' \end{bmatrix} \right\rangle_{\mathcal{H} \oplus \mathcal{K}} = \langle h, u \rangle_{\mathcal{H}} + \langle h', u' \rangle_{\mathcal{K}} \tag{A.3}$$

**Definición A.2.7.** *Sea $T$ un operador, decimos que es una **contracción** si $\|T\| \leq 1$.*

**Teorema A.2.8.** *Sean $A$, $B \in \mathcal{B}(\mathcal{H})$ positivos y sea $C$ operador en $\mathcal{H}$. El operador autoadjunto*

$$\begin{bmatrix} A & C^* \\ C & B \end{bmatrix} \tag{A.4}$$

*en $\mathcal{H} \oplus \mathcal{H}$ es positivo si y sólo si existe una contracción $K \in \mathcal{B}(\mathcal{H})$ tal que $C = B^{1/2} \, K \, A^{1/2}$. Cuando $A$ es invertible, entonces esta condición es equivalente a que*

$$B \, A^{-1} \, B^* \leq C. \tag{A.5}$$

## A.3. Transformaciones positivas

**Definición A.3.1.** *Sea $\alpha : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$, donde $\mathcal{H}, \mathcal{K}$ son espacios de Hilbert de dimensión finita. Decimos que $\alpha$ es positiva si manda operadores positivos en operadores positivos. Es decir si $A \in \mathcal{B}(\mathcal{H})$ es positivo entonces $\alpha(A) \in \mathcal{B}(\mathcal{H})$ es positivo.*

**Teorema A.3.2.** *Sea $\alpha : \mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_k(\mathbb{C})$ una aplicación lineal positiva y unital (ver definición 1.2.5). Sea $f : \mathbb{R} \to \mathbb{R}$ una función convexa. Entonces, para todo $A \in \mathbf{M}_n(\mathbb{C})$ autoadjunto*

$$\mathrm{Tr}\, \left(f(\alpha(A))\right) \leq \mathrm{Tr}\, \alpha(f(A)) \tag{A.6}$$

Consideremos la aplicación $\alpha : \mathbf{M}_n(\mathbb{C}) \oplus \mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_n(\mathbb{C})$, definida mediante

$$\alpha(A \oplus B) = tA + (1-t)B \qquad \text{para } 0 < t < 1 \tag{A.7}$$

Como tenemos una combinación lineal de operadores, $\alpha$ es lineal. La aplicación es unital pues: $\alpha(I \oplus I) = t\,I + (1-t)\,I = I$. Es positiva: Si $A \oplus B$ es operador positivo, se cumple que

$$0 \leq \langle (x,y), A \oplus B(x,y) \rangle = \langle x, Ax \rangle + \langle y, By \rangle \qquad \text{para todos } x, y \in \mathbb{C}^n.$$

Entonces $\langle x, \alpha(A \oplus B)x \rangle = \langle x, tA + (1-t)B\,x \rangle = \langle x, tAx \rangle + (1-t)\langle x, Bx \rangle = \langle t^{1/2}x, A(t^{1/2}x) \rangle + \langle (1-t)^{1/2}x, B((1-t)^{1/2}x) \rangle \geq 0$. Aplicando el Teorema A.3.2 obtenemos

$$\mathrm{Tr}\ f(tA + (1-t)B) \leq t\,\mathrm{Tr}\ f(A) + (1-t)\mathrm{Tr}\ F(B). \tag{A.8}$$

Sea $A \in \mathbf{M}_n(\mathbb{C})$ y sea $p(x) := \sum_j c_j\, x^j$ un polinomio. Resulta natural representar por $p(A)$ a la matriz $\sum_j c_j\, A^j$. El Cálculo Funcional puede ser extendido a otras funciones $f : \mathbb{C} \to \mathbb{C}$. Cuando $A = diag(\lambda_1, \lambda_2 \ldots, \lambda_n)$ y la función $f$ esté definida en los valores propios de $A$, entonces

$$f(A) = diag(f(\lambda_1), f(\lambda_2) \ldots, f(\lambda_n)).$$

Si asumimos que $A$ es diagonalizable, es decir $A = S\, diag(\lambda_1, \lambda_2 \ldots, \lambda_n)\, S^{-1}$. Entonces $f(A)$ se define como

$$f(A) = S\, diag(f(\lambda_1), f(\lambda_2) \ldots, f(\lambda_n))\, S^{-1}.$$

Recordemos que para una matriz autoadjunta existe una descomposición espectral. Sea $A = \sum_j \lambda_j |v_j\rangle\langle v_j|$ la descompisición espectral de $A$. ($\lambda_j$ son los distintos valores propios de $A$ y $|v_j\rangle\langle v_j|$ son las correspondientes proyecciones sobre los espacios propios, y el rango de $|v_j\rangle\langle v_j|$ es la multiplicidad de $\lambda_j$). Entonces

$$f(A) = \sum_j f(\lambda_j) |v_j\rangle\langle v_j|.$$

Usualmente a la función $f$ la consideramos continua o bien suave en un intervalo que contiene a los valores propios de $A$.

**Definición A.3.3.** *Decimos que una función $f$ es* ***convexa en operadores*** *si para matrices hermitianas $A$, $B \in \mathbf{M}_n$ y para toda $0 \leq \lambda \leq 1$,*

$$f\big((1-\lambda)A + \lambda B\big) \leq (1-\lambda)f(A) + \lambda f(B) \tag{A.9}$$

**Teorema A.3.4.** *Sean $A$, $B \in \mathbf{M}_n(\mathbb{C})$ autoadjuntos y $t \in \mathbb{R}$. Supongamos que $f : (a,b) \to \mathbb{R}$ es una función continua diferenciable definida en un intervalo y tal que los valores propios de $(A+tB)$ están contenidos en tal intervalo. Para $t - t_0$ pequeña se tiene que*

$$\frac{d}{dt}\mathrm{Tr}\ f(A+tB)\Big|_{t=t_0} = \mathrm{Tr}\ [B\, f'(A+t_0B)]\,. \tag{A.10}$$

**Lema A.3.5.** *Sea $A$ una matriz de bloques positiva de tamaño $n \times n$ con bloques de tamaño $k \times k$. Entonces $A$ es la suma de matrices de bloques $B$ de la forma $(B_{ij}) = X_i^* X_j$, para matrices $X_1, X_2, \ldots, X_n$ de tamaño $k \otimes k$.*

**Teorema A.3.6.** *Sea $\mathscr{E} : \mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_k(\mathbb{C})$ una aplicación lineal. Las siguientes condiciones son equivalentes.*

*(i) $\mathscr{E} \otimes I_n$ es aplicación positiva, donde $I_n : \mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_n(\mathbb{C})$ es la aplicación identidad.*

*(ii) La matriz de bloques $X$ definida como $X_{ij} = \mathscr{E}(|e_i\rangle\langle e_j|)$ para $1 \leq i,j \leq n$ es positiva.*

*(iii) Existen operadores $V_t : \mathbb{C}^n \to \mathbb{C}^k$, con $1 \leq t \leq k^2$, tales que*

$$\mathscr{E}(A) = \sum_t V_t\, A\, V_t^*. \tag{A.11}$$

*(iv) Para familias $\{A_i\} \subset \mathbf{M}_n(\mathbb{C})$ y $\{B_i\} \subset \mathbf{M}_k(\mathbb{C})$, se cumple*

$$\sum_{ij} B_i^* \mathscr{E}(A_i^* A_j) B_j \geq 0 \tag{A.12}$$

*Demostración.* $(i) \Rightarrow (ii)$ Supongamos que $\mathscr{E} \otimes I_n$ es positiva, donde tal aplicación actúa $(A,B) \mapsto \mathscr{E}(A) \otimes I_n(B) = \mathscr{E}(A) \otimes B$. Sea $X$ la matriz

$X_{ij} = \mathscr{E}(|e_i\rangle\langle e_j|)$, para $1 \leq i, j \leq n$. Sea $e_s \otimes e_t$ un tensor simple, evaluamos

$$\begin{aligned}
&\left[\sum_{ij} |e_i\rangle\langle e_j| \otimes |e_i\rangle\langle e_j|\right]^2 (e_s \otimes e_t) \\
&= \left[\sum_{ij} |e_i\rangle\langle e_j| \otimes |e_i\rangle\langle e_j|\right] \left[\sum_{ij} |e_i\rangle\langle e_j|e_s \otimes |e_i\rangle\langle e_j|e_t\right] \\
&= \left[\sum_{ij} |e_i\rangle\langle e_j| \otimes |e_i\rangle\langle e_j|\right] \left[\sum_{ij} \delta_{js}\delta_{jt} e_i \otimes e_i\right] \\
&= n\sum_{ij} |e_i\rangle\langle e_j| \otimes |e_i\rangle\langle e_j| \left[\sum_{j} \delta_{js}\delta_{jt} e_i \otimes e_i\right] \\
&= n\sum_{ij}\sum_{j} \delta_{js}\delta_{jt}\langle e_j, e_i\rangle e_i \otimes \langle e_j, e_i\rangle e_i = n\sum_{ij}\sum_{j} \delta_{js}\delta_{jt}\delta_{ji} e_i \otimes e_i \\
&= n\sum_{ij} \delta_{js}\delta_{jt} e_i \otimes e_i = \left(\sum_{ij} |e_i\rangle\langle e_j| \otimes |e_i\rangle\langle e_j|\right)(e_s \otimes e_t)
\end{aligned}$$

Tenemos la relación $\frac{1}{n}\left[\sum_{ij} |e_i\rangle\langle e_j| \otimes |e_i\rangle\langle e_j|\right]^2 = \sum_{ij} |e_i\rangle\langle e_j| \otimes |e_i\rangle\langle e_j|$, entonces tenemos una matriz positiva. Por lo tanto

$$0 \leq (\mathscr{E} \otimes I_n)\left(\sum_{ij} |e_i\rangle\langle e_j| \otimes |e_i\rangle\langle e_j|\right) = \sum_{ij} \mathscr{E}(|e_i\rangle\langle e_j|) \otimes |e_i\rangle\langle e_j| = X.$$

$(ii) \Rightarrow (iii)$ Supongamos que $X = \sum_{ij} X_{ij} \otimes |e_i\rangle\langle e_j|$ es positiva. Consideremos $\{e_l \otimes e_m\}_{l,m}$ una base para el producto tensorial. Fijemos $1 \leq i \leq k$ y definimos para toda $1 \leq l \leq n$ la proyección ortogonal $P_i(e_l \otimes e_m) = e_l \otimes e_i$. Sean $e_l \otimes e_m$, $e_{l'} \otimes e_{m'}$ tensores simples fijos, calculamos

$$\begin{aligned}
&\langle e_l \otimes e_m, P_r X P_s\, e_{l'} \otimes e_{m'}\rangle \\
&\quad = \langle P_r e_l \otimes e_m, X P_s\, e_{l'} \otimes e_{m'}\rangle = \langle e_l \otimes e_r, X e_{l'} \otimes e_s\rangle \\
&\quad = \langle e_l \otimes e_r, \sum_{ij} X_{ij} \otimes |e_i\rangle\langle e_j|\,(e_{l'} \otimes e_s)\rangle = \langle e_l \otimes e_r, \sum_{ij} X_{ij} e_{l'} \otimes \delta_{js} e_i\rangle \\
&\quad = \sum_{ij} \langle e_l \otimes e_r, X_{ij} e_{l'} \otimes \delta_{js} e_i\rangle = \sum_{ij} \langle e_l, X_{ij} e_{l'}\rangle\langle e_r, \delta_{js} e_i\rangle \\
&\quad = \sum_{i} \langle e_l, X_{is} e_{l'}\rangle\langle e_r, e_i\rangle = \langle e_l, X_{rs} e_{l'}\rangle = (X_{ll'})_{rs}.
\end{aligned}$$

Entonces se cumple $(X_{ij}) = P_i X P_j$. Ahora, como $X$ es positivo podemos escribir

$$X = \sum_{t=1}^{nk} \lambda_t |v_t\rangle\langle v_t| = \sum_{t=1}^{nk} |\lambda_t^{1/2} v_t\rangle\langle \lambda_t^{1/2} v_t| = \sum_{t=1}^{nk} |f_t\rangle\langle f_t|,$$

donde $v_t \in \mathbb{C}^{nk}$, además $f_t$ siguen siendo vectores propios. Por la completez de los proyectores tenemos que $|f_t\rangle = \sum_{i=1}^{k} P_i |f_t\rangle$, entonces podemos definir para cada $1 \leq t \leq nk$ las aplicaciones $V_t : \mathbb{C}^k \to \mathbf{Ran}(P_i) \simeq \mathbb{C}^n$ que a cada vector $e_i$ le asigna el vector $V_t e_i = P_i |f_t\rangle = |P_i f_t\rangle$. Entonces

$$\begin{aligned} X &= \sum_{t=1}^{nk} |f_t\rangle\langle f_t| = \sum_{t=1}^{nk} \left( \sum_{i=1}^{k} |P_i f_t\rangle \right) \left( \sum_{j=1}^{k} \langle P_j f_t| \right) \\ &= \sum_{t=1}^{nk} \sum_{ij}^{k} P_i |f_t\rangle\langle f_t| P_j^* = \sum_{ij}^{k} P_i \left( \sum_{t=1}^{nk} |f_t\rangle\langle f_t| \right) P_j^* \\ &= \sum_{ij}^{k} P_i \left( \sum_{t=1}^{nk} P_i |f_t\rangle\langle f_t| P_j^* \right) P_j^* = \sum_{ij}^{k} P_i \left( \sum_{t=1}^{nk} V_t |e_i\rangle\langle e_j| V_t^* \right) P_j^*. \end{aligned}$$

Entonces para $i \neq i'$ y $j \neq j'$tenemos que

$$\mathscr{E}(|e_i\rangle\langle e_j|) = P_{i'} X P_{j'} = \sum_{ij}^{k} P_{i'} P_i \left[ \sum_{t=1}^{nk} V_t |e_i\rangle\langle e_j| V_t^* \right] P_j^* P_{j'} = \sum_t V_t |e_i\rangle\langle e_j| V_t^*.$$

Como $\{|e_i\rangle\langle e_j|\}_{ij}$ es una base para las matrices de tamaño $n \times n$, extendiendo por linealidad se sigue la implicación.

$(iii) \Rightarrow (iv)$ Sea $X \in \mathbf{M}_n(\mathbb{C})$. Supongamos que $\mathscr{E}(A) = \sum_t V_t\, X\, V_t^*$. Sean $\{A_i\} \subset \mathbf{M}_n(\mathbb{C})$ y $\{B_i\} \subset \mathbf{M}_k(\mathbb{C})$. Consideramos

$$\begin{aligned} \sum_{ij} B_i^* \mathscr{E}(A_i^* A_j) B_j &= \sum_{ij} B_i^* \left( \sum_t V_t\, A_i^* A_j\, V_t^* \right) B_j \\ &= \sum_t \sum_{ij} B_i^* V_t\, A_i^* A_j\, V_t^* B_j \\ &= \sum_t \left( \sum_i B_i^* V_t\, A_i^* \right) \left( \sum_j A_j\, V_t^* B_j \right) \\ &= \sum_t \left( \sum_i (A_i V_t\, B_i)^* \right) \left( \sum_j A_j\, V_t^* B_j \right) \geq 0. \end{aligned}$$

$(iv) \Rightarrow (i)$ Nótese que $\mathscr{E} \otimes I_n : \mathbf{M}_n(\mathcal{B}(\mathcal{H})) \to \mathbf{M}_n(\mathcal{B}(\mathcal{K}))$. Supongamos que para $\{A_i\} \subset \mathbf{M}_k(\mathbb{C})$ y $\{B_i\} \subset \mathbf{M}_n(\mathbb{C})$ tales que $\sum_{ij} B_i^* \mathscr{E}(A_i^* A_j) B_j \geq 0$. Sea $A \in \mathbf{M}_n(\mathcal{B}(\mathcal{H}))$ un operador positivo. Por el Lema A.3.5 podemos escribir

$$A = \sum_{ij} A_i^* A_j \otimes |e_i\rangle\langle e_j|$$

Luego $X = \sum_{ij} \mathscr{E}(A_i^* A_j) \otimes |e_i\rangle\langle e_j|$. Sea $u \otimes v$, entonces

$$\begin{aligned}
&\left\langle u \otimes v, \sum_{ij} \mathscr{E}(A_i^* A_j) \otimes |e_i\rangle\langle e_j|(u \otimes v) \right\rangle \\
&= \sum_{ij} \langle u \otimes v, \mathscr{E}(A_i^* A_j) \otimes |e_i\rangle\langle e_j|(u \otimes v)\rangle \\
&= \sum_{ij} \langle u \otimes v, \mathscr{E}(A_i^* A_j) u \otimes |e_i\rangle\langle e_j|v\rangle = \sum_{ij} \langle u, \mathscr{E}(A_i^* A_j) u\rangle \langle v, |e_i\rangle\langle e_j|v\rangle \\
&= \sum_{ij} \langle u, \mathscr{E}(A_i^* A_j) u\rangle \langle v, e_i\rangle \langle e_j, v\rangle = \sum_{ij} \langle u, \langle v, e_i\rangle \mathscr{E}(A_i^* A_j) \langle e_j, v\rangle u\rangle \\
&= \sum_{ij} \langle u, \langle v, e_i\rangle\, I_k \mathscr{E}(A_i^* A_j) \langle e_j, v\rangle\, I_k u\rangle = \langle u, \sum_{ij} B_i^* \mathscr{E}(A_i^* A_j) B_j u\rangle \geq 0
\end{aligned}$$

Así $\mathscr{E} \otimes I_n$ es positiva. □

**Definición A.3.7.** *Sea $\mathscr{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ una aplicación entre espacios de Hilbert de dimensión finita. Decimos que $\mathscr{E}$ es Completamente Positiva si cumple con alguna, y por lo tanto todas, las condiciones del Teorema A.3.6. A la identidad* (A.11) *le llamamos Representación de Kraus.*

**Proposición A.3.8.** *Sean $\mathscr{E}_1$, $\mathscr{E}_2$ aplicaciones completamente positivas. Entonces $\mathscr{E}_1 + \mathscr{E}_2$ es completamente positiva.*

*Demostración.* Sea $A \in \mathbf{M}_n$. Por el Teorema A.3.6 tenemos

$$\mathscr{E}_1(A) = \sum_{s=1}^{n} V_s\, A\, V_s^*, \qquad \mathscr{E}_2(A) = \sum_{t=1}^{m} W_t\, A\, W_t^*.$$

Entonces

$$(\mathscr{E}_1 + \mathscr{E}_2)(A) = \sum_{s=1}^{n} V_s\, A\, V_s^* + \sum_{t=1}^{m} W_t\, A\, W_t^* = \sum_{j=1}^{n+m} U_j\, A\, U_j^*.$$

□

Nótese que un transformación completamente positiva actúa sobre operadores. En ramas como la Física a este tipo de aplicaciones se les llama super-operadores. En la siguiente proposición se muestra un ejemplo de una aplicación completamente positiva.

**Proposición A.3.9.** *Sea* $\mathscr{E} : \mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_k(\mathbb{C})$ *una aplicación lineal positiva tal que para* $A$, $B \in \mathbf{M}_n(\mathbb{C})$ *las imágenes* $\mathscr{E}(A)$ *y* $\mathscr{E}(B)$ *conmutan entre sí. Entonces* $\mathscr{E}$ *es un aplicación completamente positiva.*

*Demostración.* Sean $X \in \mathbf{M}_{mn}(\mathbb{C})$ tal que $X_{ij} \in \mathbf{M}_n(\mathbb{C})$ y $\Psi$ funcional positivo, entonces la matriz $[\Psi(X_{ij})]_{ij}$ es positiva. Sea $F \geq 0$, por el Teorema A.3.6 tenemos que la matriz de bloques $[\Psi(X_{ij})F]_{ij}$ es positiva por lo tanto $\Psi(A)F$ es completamente positiva. Ahora, sean $C$, $D \in \mathbf{Ran}(\mathscr{E})$ tales que $CD = DC$. Entonces por el Teorema de diagonalización simultánea tenemos que

$$\mathscr{E}(A) = \sum_i \Psi_i(A)|e_i\rangle\langle e_i|,$$

donde $\Psi_i$ son funcionales positivos. Como la suma de aplicaciones completamente positivas es completamente positiva, tenemos que $\mathscr{E}$ es completamente positiva. □

**Definición A.3.10.** *Sea* $\rho \in \mathbf{M}_n(\mathbb{C})$. *Decimos que* $\rho$ *es un operador de densidad si es positiva y* $Tr\,\rho = 1$. *También les llamaremos estados.*

**Definición A.3.11.** *Sean* $\rho_1$, $\rho_2$ *matrices de densidad, decimos que son ortogonales si cualquier vector propio de* $\rho_1$ *es ortogonal a cualquier vector propio de* $\rho_2$.

**Lema A.3.12.** *Sea* $\rho$ *una matriz de densidad, tal que*

$$\rho = \sum_{i=1}^{k} |x_i\rangle\langle x_i| = \sum_{j=1}^{k} |y_j\rangle\langle y_j|. \tag{A.13}$$

*Entonces, existe una matriz unitaria* $(U_{ij})$ *tal que* $\sum_{j=1}^{k} U_{ij}|x_j\rangle = |y_i\rangle$.

*Demostración.* Sea $\rho$ matriz de densidad. Para $n \leq k$ sea

$$\rho = \sum_{i=1}^{n} \lambda_i |z_i\rangle\langle z_i|,$$

su descomposición espectral, donde $n$ es el rango de $\rho$. Sean $|z_i\rangle := 0$ y $\lambda_i =: 0$ para $n < i < k$. Sea $\sqrt{\lambda_i}\,|z_i\rangle = y_i$, entonces $y_i \in Span\{|z_i\rangle\}$. Así $|y_i\rangle = \sum_{j=1}^{n}\langle z_j, y_i\rangle z_j$. Para $i = 1, \ldots, k$ y $j = 1, \ldots, n$ sea

$$(U_{ij}) = \frac{\langle z_j, y_i\rangle}{\sqrt{\lambda_j}}.$$

Es operador unitario:

$$\begin{aligned}\sum_{i=1}^{k} U_{it}U_{ir}^{*} =& \sum_{i=1}^{k}\frac{\langle z_t, y_i\rangle}{\sqrt{\lambda_t}}\frac{\langle y_i, z_r\rangle}{\sqrt{\lambda_r}} = \frac{1}{\sqrt{\lambda_t}\sqrt{\lambda_r}}\sum_{i=1}^{k}\langle z_r, \langle y_i, z_r\rangle y_i\rangle\\ =& \frac{1}{\sqrt{\lambda_t}\sqrt{\lambda_r}}\langle z_r, \sum_{i=1}^{k}|y_i\rangle\langle y_i|z_r\rangle = \frac{1}{\sqrt{\lambda_t}\sqrt{\lambda_r}}\langle z_r, \rho z_r\rangle = \delta_{tr}\end{aligned}$$

□

El Teorema A.3.6 nos da las condiciones para caracterizar una aplicación completamente positiva. Particularmente pedimos que acepte una representación de Kraus. ¿Pero tal representación es única?.

**Teorema A.3.13.** *Sea* $\mathscr{E} : \mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_k(\mathbb{C})$ *un aplicación lineal tal que acepta dos representaciones de Kraus. Digamos*

$$\mathscr{E}(A) = \sum_{t=1}^{k} V_t\,A\,V_t^{*} \qquad\qquad \mathscr{E}(A) = \sum_{t=1}^{k} W_t\,A\,W_t^{*}. \tag{A.14}$$

*Entonces existe una matriz unitaria* $(c_{tu})$ *tal que*

$$W_t = \sum_{u} c_{tu}V_u. \tag{A.15}$$

*Demostración.* Sea $\{x_i\}$ base para $\mathbb{C}^m$ y $\{y_j\}$ base para $\mathbb{C}^n$. Consideremos $v_t$ y $w_t$ los vectores en $\mathbb{C}^m \otimes \mathbb{C}^n$, escribimos

$$v_t := \sum_{ij} x_i \otimes V_t y_j \qquad\qquad w_t := \sum_{ij} x_i \otimes W_t y_j.$$

Ahora

$$\begin{aligned}|v_t\rangle\langle v_t| =& \Big|\sum_{ij} x_i \otimes V_t y_j\Big\rangle\Big\langle\sum_{i'j'} x_{i'} \otimes V_t y_{j'}\Big| = \sum_{ij\,i'j'} |x_i \otimes V_t y_j\rangle\langle x_{i'} \otimes V_t y_{j'}|\\ =& \sum_{ij\,i'j'} |x_i\rangle\langle x_{i'}| \otimes V_t|y_j\rangle\langle y_{i'}|V_t^{*},\end{aligned}$$

análogamente $|w_t\rangle\langle w_t| = \sum_{ij\,i'j'} |x_i\rangle\langle x_{i'}| \otimes W_t|y_j\rangle\langle y_{i'}|W_t^*$. Así, de nuestra hipótesis y lo anterior

$$\sum_t |v_t\rangle\langle v_t| = \sum_{ij\,i'j'} |x_i\rangle\langle x_{i'}| \otimes \mathscr{E}(A) = \sum_t |w_t\rangle\langle w_t|.$$

Entonces por el Lema A.3.12 existe una matriz unitaria $(c_{tu})$ tal que $w_t = \sum_u c_{tu}v_u$. Sea $x_k \otimes y_{k'}$, entonces

$$\begin{aligned}&\langle x_k \otimes y_{k'}, w_t\rangle = \langle x_k \otimes y_{k'}, \sum_{ij} x_i \otimes W_t y_j\rangle = \sum_{ij}\langle x_k, x_i\rangle\langle y_{k'}, W_t y_j\rangle \\ &= \sum_j \langle y_{k'}, W_t y_j\rangle.\end{aligned}$$

Por otra parte tenemos que

$$\begin{aligned}\langle x_k \otimes y_{k'}, w_t\rangle &= \langle x_k \otimes y_{k'}, \sum_u c_{tu}v_u\rangle = \sum_u c_{tu}\langle x_k \otimes y_{k'}, \sum_{ij} x_i \otimes V_u y_j\rangle \\ &= \sum_u c_{tu}\sum_{ij}\langle x_k, x_i\rangle\langle y_{k'}, V_u y_j\rangle = \sum_j \langle y_{k'}, \sum_u c_{tu} V_u y_j\rangle.\end{aligned}$$

Por lo tanto $W_t = \sum_u c_{tu}V_u$. □

## A.4. El operador de traza parcial

En ésta sección estudiaremos las propiedades de la aplicación de Traza Parcial, la cual veremos que es otro ejemplo de una transformación completamente positiva. En esta parte seguimos la exposición de [12].

**Definición A.4.1.** *Sean $\mathcal{H}_1$ y $\mathcal{H}_2$ espacios de Hilbert. Sea $\{e_i\}_i$ una base ortonormal de $\mathcal{H}_2$. Sea $\rho$ un operador sobre $\mathcal{H}_1 \otimes \mathcal{H}_2$. La traza parcial de $\rho$ es el operador $\mathrm{Tr}_2\rho \in \mathcal{B}(\mathcal{H}_1)$ definido para todos $u$, $v \in \mathcal{H}_1$ mediante*

$$\langle u, \mathrm{Tr}_2\rho v\rangle_{\mathcal{H}_1} = \sum_i \langle u \otimes e_i, \rho(v \otimes e_i)\rangle_{\mathcal{H}_1\otimes\mathcal{H}_2}. \tag{A.16}$$

*Denotaremos mediante $\rho_1 = \mathrm{Tr}_2\rho$, a la traza parcial del operador $\rho$. Análogamente, para la traza parcial $\rho_2 = \mathrm{Tr}_1\rho$ la calculamos tomando la base en $\mathcal{H}_1$.*

**Proposición A.4.2.** *Sea $\rho$ un operador sobre $\mathcal{H}_1 \otimes \mathcal{H}_2$ y sea $\rho_1 \in \mathcal{B}(\mathcal{H}_1)$. Entonces $\rho_1 = Tr_2\rho$ si y sólo si para toda $X \in \mathcal{B}(\mathcal{H}_1)$ se cumple que*

$$\mathrm{Tr}\left(\rho_1 X\right) = \mathrm{Tr}\,\rho(X \otimes I). \tag{A.17}$$

**Proposición A.4.3 (Propiedades de la traza parcial).** *Sean $\mathcal{H}_1$,$\mathcal{H}_2$ espacios de Hilbert de dimensión finita.*

*(i) Sean $\rho, \gamma \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$, $\lambda \in \mathbb{C}^n$. Entonces*

$$\mathrm{Tr}_2(\lambda\rho + \gamma) = \lambda \mathrm{Tr}_2\rho + \mathrm{Tr}_2\gamma. \tag{A.18}$$

*(ii) Sea $\rho \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$. Entonces*

$$\mathrm{Tr}\,\rho = \mathrm{Tr}\,(\rho_1) = Tr(\rho_2). \tag{A.19}$$

*(iii) Sean $\alpha \in \mathcal{B}(\mathcal{H}_1)$ y $\beta \in \mathcal{B}(\mathcal{H}_2)$. Entonces*

$$\mathrm{Tr}_2(\alpha \otimes \beta) = \alpha \mathrm{Tr}\,\beta \qquad\qquad \mathrm{Tr}_1(\alpha \otimes \beta) = \beta \mathrm{Tr}\,\alpha. \tag{A.20}$$

**Teorema A.4.4.** *Sea $\rho \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ positivo. Sea $f$ una función real tal que los valores propios de $\rho$ están contenidos en $\mathscr{D}(f)$. Entonces*

$$\mathrm{Tr}_2\, f(\rho) = f(\mathrm{Tr}_2\,\rho). \tag{A.21}$$

**Ejemplo A.4.5.** *La traza parcial es un operador completamente positivo.*

**Ejemplo A.4.6.** *La traza $\mathrm{Tr}\,:\, \mathbf{M}_k(\mathbb{C}) \to \mathbb{C}$ es una aplicación Completamente positiva si*

$$\mathrm{Tr} \otimes I_n : \mathcal{M}_k(\mathbb{C}) \otimes \mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_n(\mathbb{C}), \tag{A.22}$$

*es una aplicación positiva.*

*Demostración.* Sea $\rho \in \mathcal{B}(\mathcal{H}_1) \otimes \mathcal{B}(\mathcal{H}_2)$, supongamos que $\rho = A \otimes B$. Veamos que $\mathrm{Tr} \otimes I_n$ es una traza parcial. Es decir, $\rho_2 = \mathrm{Tr}_1\rho = (\mathrm{Tr} \otimes I_n) \cdot \rho$. Sea $\{e_i\}$ una base ortonormal de $\mathcal{H}_1$. Entonces por el inciso $(iii)$ de la Proposición A.4.3 tenemos que $\mathrm{Tr}_1(A \otimes B) = (\mathrm{Tr}\ A)\, B$.

Por otra parte tenemos que $(\mathrm{Tr} \otimes I_n)(A \otimes B) = \mathrm{Tr}\ A \otimes B$. Luego

$$\begin{aligned}\sum_i \langle e_i \otimes u, \mathrm{Tr}\ A \otimes B(e_i \otimes v)\rangle &= \sum_i \langle e_i \otimes u, \mathrm{Tr}\ Ae_i \otimes B\,v)\rangle \\ &= \sum_i \mathrm{Tr}\ A\langle e_i, e_i\rangle\langle u, B\,v\rangle = \langle u, (\mathrm{Tr}\ A)\, B\,v\rangle.\end{aligned}$$

Así por el Teorema A.3.6, la traza es una aplicación completamente positiva. □

Sean $\mathcal{H}_1$, $\mathcal{H}_2$, $\mathcal{H}_3$ Espacios de Hilbert de dimensión finita. Sea $X \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3)$. Consideremos $\{u_i\}_i$ una base ortonormal del espacio $\mathcal{H}_1$. Tenemos de la definición A.4.1 que $\operatorname{Tr}_1 X$ es el operador de traza parcial que actúa de $\mathcal{H}_2 \otimes \mathcal{H}_3$ en sí mismo. Sea $\{v_{j,k}\}_{j,k}$ una base para el espacio $\mathcal{H}_1 \otimes \mathcal{H}_2$, entonces $\operatorname{Tr}_{1,2} X$ denotará el operador de traza parcial de $X$ que actúa sobre el espacio de Hilbert $\mathcal{H}_3$. Análogamnete $\operatorname{Tr}_{2,3} X$ denota el operador de traza parcial de $X$ que actúa sobre el espacio de Hilbert $\mathcal{H}_1$.

**Proposición A.4.7.** *Sea $\rho \in \mathcal{B}(\mathcal{H}_2 \otimes \mathcal{H}_3)$. Entonces*

$$\operatorname{Tr}_3(I_1 \otimes \rho) = I_1 \otimes \operatorname{Tr}_3 \rho \tag{A.23}$$

**Proposición A.4.8.** *Sea $\rho \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ y $\varphi \in \mathcal{B}(\mathcal{H}_3)$ tal que $\operatorname{Tr}\varphi = 1$. Entonces*

$$(i)\ \operatorname{Tr}_{1,3}(\rho \otimes \varphi) = \operatorname{Tr}_1 \rho. \qquad (ii)\ \operatorname{Tr}_1(\rho \otimes \varphi) = \operatorname{Tr}_1 \rho \otimes \varphi. \tag{A.24}$$

## A.5. Entropía cuántica o de von Neumann

**Proposición A.5.1.** *Sea $\mathscr{E} : \mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_n(\mathbb{C})$ una aplicación positiva, unital y que preserva traza. Entonces $\mathscr{E}$ manda al operador de densidad $\rho$ en la matriz de densidad $\mathscr{E}(\rho)$.*

**Definición A.5.2.** *Sea $\rho$ un operador de densidad. Se define la Entropía de von Neumann como*

$$S(\rho) = -\operatorname{Tr}\ \rho \log \rho. \tag{A.25}$$

Como $\rho$ es positivo ponemos escribir su representación espectral

$$\rho = \sum_i \lambda_i |v_i\rangle\langle v_i|.$$

Ahora, para calcular la traza tomemos la base que diagonaliza a $\rho$. Entonces

$$\begin{aligned} S(\rho) &= -\operatorname{Tr}(\rho \log \rho) = -\sum_k \left\langle v_k, \rho \log \rho\, v_k \right\rangle \\ &= -\sum_k \left\langle \sum_i \lambda_i |v_i\rangle\langle v_i|v_k, \sum_i \log(\lambda_i)|v_i\rangle\langle v_i|, v_k \right\rangle = -\sum \lambda_k \log \lambda_k. \end{aligned}$$

**Proposición A.5.3.** *Sea $\mathscr{E}$ una aplicación positiva, unital y que preserva traza. Entonces para toda $\rho$ matriz de densidad*

$$S(\rho) \leq S(\mathscr{E}(\rho)). \tag{A.26}$$

**Proposición A.5.4.** *Sea $U \in \mathcal{B}(\mathcal{H})$ y sea $\rho$ un estado. Entonces*

$$S(U\rho U^*) = S(\rho). \tag{A.27}$$

**Definición A.5.5.** *Sean $\rho$, $\sigma$ estados tales que $Ker(\sigma) \subset Ker(\rho)$. La* ***entropía relativa*** *de $\rho$ respecto a $\sigma$ se define mediante*

$$S(\rho\|\sigma) = \operatorname{Tr}\rho\log\rho - \operatorname{Tr}\rho\log\sigma. \tag{A.28}$$

**Proposición A.5.6.** *Sea $U \in \mathcal{B}(\mathcal{H})$ unitario y sean $\rho$, $\sigma$ estados. Entonces*

$$S\Big(U\rho U^*\|U\sigma U^*\Big) = S(\rho\|\sigma). \tag{A.29}$$

**Teorema A.5.7.** *Sea $A \in \mathcal{B}(\mathcal{H})$ autoadjunto, sea $\rho$ un estado. Si $f$ es una función convexa real, con dominio que para toda $k$ contenga a $\langle\alpha_k, A\,\alpha_k\rangle$ y al conjunto de valores propios de $A$. Entonces*

$$f\big(\operatorname{Tr}\,A\rho\big) \leq \operatorname{Tr}\,\big(f(A)\rho\big). \tag{A.30}$$

# Bibliografía


[1] Bhatia, R., *Matrix Analysis*, Springer-Verlag, 1997.

[2] Bhatia, R., *Positive Definite Matrices*, Princeton University Press, 2007.

[3] Bhatia, R., *Linear Algebra to Quantum Cohomology: The story of Alfred Horn's Inequalities*, The American Mathematical Monthly, Vol.108, No. 4, (Apr. 2001), pp 289-318.

[4] Carlen, E.A. and Lieb, E., *A Minkowski type trace inequality and strong subadditivity of quantum entropy*, Advances in the Mathematical Science, AMS Transl. vol. 189, Series 2, (1999) pp 59-68.

[5] Carlen, E.A. and Lieb, E., *A Minkowski type trace inequality and strong subadditivity of quantum entropy II: convexity and concavity*, Lett. math. Phys. 2008 83, pp 107-126.

[6] Epstein, H., *Remarks on Two Theorems on E. Lieb*, Commun. Math. Phys. 31 (1973), pp 317-325.

[7] Gibilisco, P., Imparato, D. and Isola, T., *Schrödinger Equation, $L^p-$Duality and the Geometry of Wigner-Yanase-Dyson Information*, in Quantum Probability and Related Topics, Proceedings of the 28th Conference, pg. 157-164. 2008.

[8] Lieb, E. H. and Loss, M., *Analysis*, 2nd. Ed. Graduate text in Mathematics Vol. 14, American Mathematical Society.

[9] McCarthy, C. A., $\mathcal{C}_p$, Israel J. Math. 5 (1967), pp. 249-271.

[10] Nielsen, M. A., *Probability distributions consistent with a mixted state*, arXiv: quant-ph/9909020v1, Sept. 1999.

[11] Pantaleón, L., *Desigualdades para la Entropía de von Neumann*, Tesis de Maestría, Posgrado en Matemáticas, UAM-I, 2004.

[12] Pantaleón, L. y Quezada, R., *Una introducción a la Teoría cuántica de la información*, Departamento de Matemáticas UAM-I, 2008.

[13] Petz, D., *Quantum Information Theory and Quantum Statistics*, Springer, 2008.

[14] Prasolov, V. V.,*Problems and Theorems in Linear Algebra*, Translations of Mathematical Monographs, Vol. 134, american Mathematical Society, 2000.

[15] Reed, M., Simon, B., *Methods of Modern Mathematical Physics*, Academic Press, 1980.

[16] Ruskai, M. B., *Inequalities for Quantum Entropy*, J. Math. Phys. 43, No. 9, 4358-4375 (2002).

[17] Ruskai, M. B. and Jenčová, A., *A Unified Treatment of Convexity of Relative Entropy and Related Trace Functions, with Conditions for Equality*, arxiv:0903.2895v1 [quant-ph] 17 Mar 2009.