UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Departamento de Matemática

Tesis de Licenciatura

Desigualdades en espacios de operadores

René Elencwajg

Director: Gabriel Larotonda

Abril 2010

Agradecimientos

A Abigail, el mejor daño colateral de haber abandonado ingeniería.

A mis viejos y a mis hermanos, por soportar durante toda la carrera las oscilaciones de mi humor al compás de los exámenes.

Al resto de mi familia, por los motivos standard.

A la familia de Abigail, por haberme recibido como si fuese un buen partido.

A mis amigos, por bancarse la defensa de este trabajo en silencio (si no viniste no sos amigo, a lo sumo, un conocido).

A toda la gente de la facultad con la que compartí cursadas, cafés, partidos, etc., que hicieron que todo sea más leve.

A Gabriel, director de este proyecto de matemático que soy, por su dedicación absoluta, y por adelantado por la paciencia que espero no agotar en los próximos años.

A la gente del IAM, por su buena onda en la falsa defensa, en particular a Cristian, por la infinidad de errores que encontró, y a Eugenia, que no paró de sonreir y asentir durante la defensa oficial.

A Alejandra y a Silvia, jurados de esta tesis, por haberse tomado un tiempo para mí.

A los que puedan estar pensando que me olvidé de ellos, a quienes les reservé estos dos renglones especialmente.

Índice general

Resumen

Esta tesis está orientada hacia el estudio de desigualdades de operadores en espacios de Hilbert, más precisamente se presentan algunas nuevas demostraciones, publicadas por E. Effros en [10], de resultados ya conocidos, entre los cuales se destacan

- La convexidad conjunta de la entropía relativa

$$S(\rho||\sigma) = Tr(\rho \log \rho - \rho \log \sigma),$$

sobre operadores estrictamente positivos de la forma $I + N$, con N nuclear.

- La concavidad conjunta de la aplicación definida sobre los pares de operadores positivos

$$(A, B) \mapsto Tr(A^q X^* B^p X),$$

donde X es un operador nuclear y $0 < p, q$ verifican $p + q \leq 1$.

Si bien en [10] se trabaja en dimensión finita, las mismas ideas sirven para extender los resultados a espacios de dimensión infinita. Lo más valioso del enfoque de estas demostraciones, es la relativa sencillez respecto a otras, por ejemplo, las publicadas en [24], [25], [19]. Cronológicamente, los resultados e ideas necesarios para llegar a las pruebas de Effros, se podrían resumir de la siguiente manera:

- En 1906, J. L. W. V. Jensen muestra en [15] que toda función convexa (de punto medio) y continua f, definida en un intervalo $I \subset \mathbb{R}$, satisface

$$f(\sum \lambda_i t_i) \leq \sum \lambda_i f(t_i),$$

para cualquier combinación convexa $\{\lambda_i\}$ de puntos $\{t_i\}$ en I. Esto se puede reescribir, si $f(0) = 0$, como $f(a^*xa) \leq a^*f(x)a$ con x la matriz diagonal cuya coordenada ii es t_i y a la matriz cuya primera columna está formada por las raíces de los λ_i completada con ceros.

- En 1982, en [12], y en 2003, en [13], F. Hansen y G. K. Pedersen establecen una serie de propiedades equivalentes a la convexidad de operadores, extendiendo la desigualdad de Jensen.

- En 2005, en [21] y [22] P. Marechal introduce una generalización de la perspectiva de una función convexa f, definiendo, para h convexa,

$$f\Delta h(x, y) = f\left(\frac{x}{h(y)}\right) h(y),$$

que se generaliza a operadores mediante el cálculo funcional. Usando las equivalencias de Hansen y Pedersen se prueba que la perspectiva de Marechal es conjuntamente convexa para cierto tipo de operadores.

- Finalmente, en 2009, E. G. Effros [10] usa la convexidad conjunta de la perspectiva de Marechal para re-demostrar algunos resultados, entre ellos la convexidad conjunta de la entropía relativa (resultado original de Lieb y Ruskai publicado en [18], en 1973) y la concavidad conjunta para cierto tipo de operadores de la aplicación $F(A, B) = tr(A^p K^* B^q K)$, con $0 < p, q$ y $p + q \leq 1$ (resultado obtenido en 1973 por Lieb [17]).

La teoría de convexidad de matrices y de operadores, de interés por sí misma, se ha revitalizado en las últimas décadas con la aparición de la teoría cuántica de la información y la computación cuántica, que han proporcionado numerosos ejemplos y nuevos problemas. Por citar alguno, en [26], se muestra que la entropía

$$S(\rho) = tr(\rho \log \rho)$$

se puede interpretar, en cierto sentido, como la cantidad promedio de qubits necesarios para codificar con fidelidad mensajes emitidos por la fuente ρ. Una introducción a la teoría cuántica de la información y la computación cuántica puede encontrarse en [26], [3] y [1].

Esta tesis esta organizada del siguiente modo: En el primer capítulo, se repasan ciertos resultados básicos de teoría de operadores que nos servirán a lo largo de todo el trabajo. También extendemos las nociones de monotonía y convexidad a funciones definidas sobre el espacio de matrices Hermitianas. El segundo capítulo está dedicado a estudiar algunas propiedades de las funciones monótonas y convexas de matrices, culminando con la caracterización de Loewner. El tercer capítulo está dedicado a las C^*-álgebras. De particular interés para este trabajo resulta la transformada de Gelfand, que nos permitirá definir funciones continuas sobre operadores normales y acotados sobre espacios de Hilbert. El último capítulo se concentra en las demostraciones recientes de Effros a las que nos referimos anteriormente. Repasamos allí la desigualdad de Jensen para operadores y extendemos algunos resultados al contexto infinito dimensional, esencialmente a clases de operadores compactos.

Capítulo 1

Funciones monótonas y convexas de matrices.

Comenzaremos este capítulo con un repaso de definiciones y resultados que serán utilizados en reiteradas ocasiones en este trabajo. Luego extenderemos el concepto de monotonía y de convexidad a funciones definidas sobre el espacio de matrices Hermitianas. Sorprendentemente no toda función monótona o convexa en el sentido clásico lo es para matrices. Mostraremos algunos ejemplos y probaremos que tanto el espacio de funciones monótonas de matrices como el de funciones convexas de matrices son cerrados por combinaciónes lineales con coeficientes positivos, y por límite puntual. Seguiremos a grandes rasgos el capítulo 5 del libro de Bhatia [4].

1.1. Resultados básicos.

Notación 1.1.1. Notaremos $M_n(\mathbb{C})$ al conjunto de los operadores $\mathbb{C}$-lineales sobre $\mathbb{C}^n$, y lo identificaremos con el espacio de matrices $\mathbb{C}^{n\times n}$.

Proposición 1.1.2. *Sea $\mathcal{F}$ una familia conmutativa de operadores lineales diagonalizables en $\mathbb{C}^n$. Entonces existe una base de V en la cual todo operador de $\mathcal{F}$ está representado por una matriz diagonal.*

Demostración. Hacemos inducción en n. Si $n = 1$, es trivial. Supongamos que la proposición es cierta para todo $n < N$, veamos que también vale para N. Sea $A \in \mathcal{F}$ tal que A no es un múltiplo de la identidad. Sean $\lambda_1, ..., \lambda_k$ los autovalores distintos de A, y $W_i = Ker(A - \lambda_i I)$ $i = 1, ..., k$. Para un índice fijo cualquiera, digamos i_0, se tiene que W_{i_0} es invariante por cada operador que conmuta con A. Sea $\mathcal{F}_{i_0}$

la familia de operadores lineales en W_{i_0} que se obtiene restringiendo los operadores de $\mathcal{F}$ a W_{i_0}. Cada operador de $\mathcal{F}_{i_0}$ es diagonalizable, pues su polinomio minimal divide al polinomio minimal del correspondiente operador en $\mathcal{F}$. Como $dim(W_{i_0}) < dim(V)$, los operadores en $\mathcal{F}_{i_0}$ se diagonalizan simultáneamente en una base $\mathcal{B}_{i_0}$. Luego $\mathcal{B}$ =($\mathcal{B}_1$,...,$\mathcal{B}_k$) es la base de $\mathbb{C}^N$ en la cual se diagonalizan simultáneamente los operadores de $\mathcal{F}$. □

Definición 1.1.3. *Dada una matriz A, definimos la* matriz adjunta de A *como*

$$A^* = \overline{A^t}.$$

La siguiente es una propiedad fundamental de las matrices adjuntas, de demostración trivial.

Proposición 1.1.4. *Sea $A \in M_n(\mathbb{C})$. Si para todo $x, y \in \mathbb{C}^n$ se tiene*

$$\langle x, Ay\rangle = \langle Bx, y\rangle,$$

entonces $B = A^*$.

Definición 1.1.5. *Decimos que una matriz $A \in M_n(\mathbb{C})$ es*

- normal*, si $AA^* = A^*A$.*
- Hermitiana o autoadjunta*, si $A = A^*$.*
- unitaria*, si $A^*A = I$.*

Definición 1.1.6. *Al conjunto de autovalores de una matriz A lo llamaremos* espectro de *A.*

Lema 1.1.7. *Sea $T : \mathbb{C}^n \to \mathbb{C}^n$ un operador lineal. Si W es un subespacio invariante por T, entonces el complemento ortogonal de W es invariante por T^*.*

Demostración. Sean $x \in W^\perp$, $y \in W$, entonces $\langle y, T^*x\rangle = \langle Ty, x\rangle = 0$, pues $Ty \in W$. □

Lema 1.1.8. *Para todo operador $T : \mathbb{C}^n \to \mathbb{C}^n$, existe una base ortonormal (notaremos b.o.n.) en la cual T se representa por una matriz triangular superior.*

Demostración. Hacemos inducción en n. Si $n = 1$ el resultado es trivialmente cierto. Supongamos que es cierto para $n-1$. Sea e_1 un autovector de T^* de norma 1. Sea W el complemento ortogonal del subespacio generado por e_1, y sea S la restricción de T a W. Como, por el Lema 1.1.7, W es un espacio invariante por T, resulta que S es un operador lineal sobre W. Por hipótesis inductiva, existe $\{e_2, ..., e_n\}$ b.o.n. de W en la cual S es triangular superior. Luego $\{e_1, ..., e_n\}$ es una b.o.n. de V en la cual T se representa por una matriz triangular superior. □

Lema 1.1.9. *Sea T una transformación lineal que en la b.o.n. $\{e_1, ..., e_n\}$ se representa por una matriz $N = (N_{ij})$ triangular superior y normal. Entonces N es diagonal.*

Demostración. Como N es triangular superior $Ne_1 = N_{11}e_1$. Luego, como los autovalores de N^* son los conjugados de los de N, se tiene $N^*e_1 = \overline{N_{11}}e_1$. Por otro lado $N^*e_1 = \sum_j (N^*)_{j1}e_j = \sum_j \overline{N_{1j}}e_j$. Por lo tanto, $N_{1j} = 0$, para todo $j > 1$. En particular $N_{12} = 0$, y como N es triangular superior se tiene que $Ne_2 = N_{22}e_2$. Así, $N^*e_2 = \overline{N_{22}}e_2$ y $N_{2j} = 0$ para todo $j \neq 2$. Continuando de esta manera se llega a que N es diagonal. □

Proposición 1.1.10. *Toda matriz normal es diagonalizable. Si la matriz es Hermitiana, sus autovalores son reales.*

Demostración. Es inmediato por los Lemas 1.1.8 y 1.1.9 que las matrices normales son diagonalizables. Si la matriz es Hermitiana, es diagonalizable, y los elementos de la diagonal de toda matriz Hermitiana diagonal son claramente reales. □

Definición 1.1.11. *Decimos que una matriz $P \in M_n\mathbb{C}$ es* positiva (estrictamente)*, si es Hermitiana y todos sus autovalores son no negativos (positivos). Escribiremos $P \geq 0$ ($P > 0$) para indicar que P es una matriz positiva (estrictamente).*

Notación 1.1.12. Notaremos $A \leq B$ para indicar que A y B son matrices Hermitianas tales que $B - A$ es positiva ($B - A \geq 0$). De la siguiente proposición se deduce fácilmente que la relación $\leq$ es un orden parcial en las matrices Hermitianas.

Proposición 1.1.13. *Sea $P = (p_{ij}) \in M_n(\mathbb{C})$, son equivalentes*

1. $P > 0$.

2. $\langle Pz, z\rangle > 0 \; \forall z \neq 0 \in \mathbb{C}^n$.

Demostración. $1 \Rightarrow 2$) Como $P > 0$, resulta $P = UDU^*$ con U matriz unitaria y $D = (d_{ij})$ matriz diagonal con $d_{ii} > 0\ \forall i$. Luego,

$$\begin{aligned} \langle Pz, z\rangle &= \langle UDU^*z, z\rangle = \langle DU^*z, U^*z\rangle \\ &= \sum_{i=1}^{n} d_{ii}|(U^*z)_i|^2 > 0. \end{aligned}$$

$2 \Rightarrow 1$) Veamos primero que P es Hermitiana.

$$\langle P(x+y), x+y\rangle = \langle Px, x\rangle + \langle Py, y\rangle + \langle Px, y\rangle + \langle Py, x\rangle,$$

como $\langle P(x+y), x+y\rangle$, $\langle Px, x\rangle$ y $\langle Py, y\rangle$ son reales, resulta que $\langle Px, y\rangle + \langle Py, x\rangle$ también es un número real. El mismo razonamiento con $x - iy$ en lugar de $x + y$ nos dice que $-i\langle Px, y\rangle + i\langle Py, x\rangle$ es real. Entonces resulta

$$\langle Px, y\rangle + \langle Py, x\rangle = \overline{\langle Px, y\rangle} + \overline{\langle Py, x\rangle} \tag{1.1}$$

$$-i\langle Px, y\rangle + i\langle Py, x\rangle = i\overline{\langle Px, y\rangle} - i\overline{\langle Py, x\rangle}. \tag{1.2}$$

Multiplicando (1.2) por i y sumando a (1.1) resulta

$$2\langle Px, y\rangle = 2\overline{\langle Py, x\rangle},$$

y como $\overline{\langle Py, x\rangle} = \langle x, Py\rangle$, la matriz P resulta Hermitiana. Veamos que sus autovalores son positivos. Como P es Hermitiana, existen U unitaria y D diagonal tal que $P = UDU^*$. Por hipótesis $\langle Pz, z\rangle > 0\ \forall z \in \mathbb{C}^n$, en particular, si $z = Ue_i$, resulta

$$0 < \langle PUe_i, Ue_i\rangle = \langle UDU^*Ue_i, Ue_i\rangle = \langle De_i, e_i\rangle = d_{ii}.$$

□

Definición 1.1.14. *Sea f una función compleja definida en un conjunto $I \subseteq \mathbb{C}$. Sea A una matriz normal con autovalores en I, tal que $A = UDU^*$, con U matriz unitaria y $D = diag(\lambda_1, ..., \lambda_n)$, es decir la matriz diagonal*

$$\begin{pmatrix} \lambda_1 & ... & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & ... & ... & \lambda_n \end{pmatrix}.$$

Definimos $f(A) = Udiag(f(\lambda_1), ..., f(\lambda_n))U^*$.

Observación 1.1.15. En particular, dada una función real definida en un intervalo $I \subseteq \mathbb{R}$, queda definida f sobre el espacio de matrices Hermitianas con espectro contenido en I. Una tal f aplica matrices Hermitianas en matrices Hermitianas.

Proposición 1.1.16. *Dada una matriz Hermitiana A y una función continua f, entonces $f(A)$ conmuta con cualquier matriz que conmute con A.*

Demostración. Es inmediato por la diagonalización simultánea (Proposición 1.1.2). □

Notación 1.1.17. Dada una matriz A, A^*A es siempre positiva. Notaremos $|A|$ a la única matriz positiva que elevada al cuadrado da A^*A.

Lema 1.1.18 (Descomposición Polar)**.** *Toda matriz A se puede escribir como $A = UP$ con U unitaria y P positiva. En la descomposición polar P resulta único ($P = |A|$), y si A es inversible, U también es único.*

Demostración. Unicidad: sea $A = UP$ una descomposición polar, luego $A^* = PU^*$. Así $A^*A = PU^*UP = P^2$ lo que implica que $P = |A|$. Si A es inversible, P también lo es (pues $P = U^*A$) luego queda unívocamente determinado $U = AP^{-1}$.

Existencia: Si A es inversible, sólo resta probar que U es unitaria.

$$\begin{aligned} UU^* &= AP^{-1}P^{-1}A^* \\ &= A(P^{-1})^2A^* \\ &= A(P^2)^{-1}A^* \\ &= A(A^*A)^{-1}A^* \\ &= AA^{-1}(A^*)^{-1}A^* = I. \end{aligned}$$

En el caso general, se define U sobre el rango de P como $UP\alpha = A\alpha$, y sobre $ran(P)^\perp$ como $U\alpha = U_0\alpha$, donde $U_0 : ran(P)^\perp \to ran(A)^\perp$ es un isomorfismo. U está bien definida sobre $ran(P)$, ya que $P\alpha = P\beta$ si y sólo si $A\alpha = A\beta$, pues

$$\begin{aligned} \|A(\beta-\alpha)\|^2 &= \langle A(\beta-\alpha), A(\beta-\alpha)\rangle \\ &= \langle A^*A(\beta-\alpha), \beta-\alpha\rangle \\ &= \langle P^2(\beta-\alpha), \beta-\alpha\rangle \\ &= \langle P(\beta-\alpha), P(\beta-\alpha)\rangle \\ &= \|P(\beta-\alpha)\|^2 . \end{aligned}$$

De esto se deduce que $Ker(A) = Ker(P)$,

$$dim(ran(A)) = dim(ran(P)) \text{ y } dim(ran(A)^\perp) = dim(ran(P)^\perp),$$

por lo que el isomorfismo U_0 existe.

Lo único que resta probar es que U es unitaria: sea $\alpha = \beta + \gamma$, con $\beta \in ran(P)$ y $\gamma \in ran(P)^{\perp}$

$$\begin{aligned} \langle U\alpha, U\alpha\rangle &= \langle A\beta + U_0\gamma, A\beta + U_0\gamma\rangle \\ &= \langle A\beta, A\beta\rangle + \langle U_0\gamma, U_0\gamma\rangle \\ &= \langle P\beta, P\beta\rangle + \langle \gamma, \gamma\rangle \\ &= \langle \alpha, \alpha\rangle. \end{aligned}$$

□

1.2. Definiciones, ejemplos y propiedades.

Definición 1.2.1. *Decimos que una función real f es* monótona de matrices de orden n *en un intervalo $I \subseteq \mathbb{R}$ si preserva el orden parcial, es decir, si $f(A) \geq f(B)$ cada vez que $A \geq B$, donde $A, B \in M_n(\mathbb{C})$. Decimos que f es* monótona de matrices *en I si es monótona de matrices de orden n en I para todo $n \in \mathbb{N}$.*

Notación 1.2.2. Sea $X \in M_n(\mathbb{C})$ Hermitiana, notamos $\lambda^{\downarrow}(X) = (\lambda_1^{\downarrow}, ..., \lambda_n^{\downarrow})$ al vector formado por los autovalores de X (contados con multiplicidad) ordenados en forma decreciente.

Teorema 1.2.3. *(H. Weyl) Sean $A, B \in M_n(\mathbb{C})$ Hermitianas, entonces*

- $\lambda_j^{\downarrow}(A+B) \leq \lambda_i^{\downarrow}(A) + \lambda_{j-i+1}^{\downarrow}(B)$ *si* $i \leq j$.
- $\lambda_j^{\downarrow}(A+B) \geq \lambda_i^{\downarrow}(A) + \lambda_{j-i+n}^{\downarrow}(B)$ *si* $i \geq j$.

Demostración. Sean u_j, v_j, w_j los autovectores de A, B y $A+B$ respectivamente, asociados a autovalores ordenados en forma decreciente. Sea $i \leq j$. Consideremos los subespacios $W = span\{w_1, ..., w_j\}$, $U = span\{u_i, ..., u_n\}$ y $V = span\{v_{j-i+1}, ..., v_n\}$. Como la suma de sus dimensiones es mayor que $2n$, tienen intersección no trivial. Sea x un vector de norma unitaria en la intersección, entonces

$$\lambda_j^{\downarrow}(A+B) \leq \langle x, (A+B)x\rangle = \langle x, Ax\rangle + \langle x, Bx\rangle \leq \lambda_i^{\downarrow}(A) + \lambda_{j-i+1}^{\downarrow}(B),$$

donde en la primera desigualdad usamos que $x \in W$ y en la otra que $x \in U \cap V$.

La desigualdad para $i \geq j$ se obtiene intercambiando A y B por $-A$ y $-B$ respectivamente. □

Corolario 1.2.4. *Dadas A, B matrices Hermitianas con espectro en un intervalo I, cualquier combinación convexa de A y B también es Hermitiana con espectro en I.*

Demostración. Es inmediato, tomando $i = j$ en el teorema anterior: sea $c \in [0,1]$,

$$c\lambda_j^{\downarrow}(A) + (1-c)\lambda_n^{\downarrow}(B) \leq \lambda_j^{\downarrow}(cA + (1-c)B) \leq c\lambda_j^{\downarrow}(A) + (1-c)\lambda_1^{\downarrow}(B).$$

□

Definición 1.2.5. *Dada $A \in M_n(\mathbb{C})$, definimos la* norma de A *como*

$$\|A\| = \sup \{Ax : x \in \mathbb{C}^n, \ \|x\| \leq 1\}.$$

Proposición 1.2.6. *Si $A \in M_n(\mathbb{C})$ es normal, entonces*

$$\|A\| = \text{máx}\{|\lambda_1^{\downarrow}(A)|, |\lambda_n^{\downarrow}(A)|\}.$$

Demostración. Es trivial, trabajando en una b.o.n. en la que A es diagonal. □

Corolario 1.2.7. *Para todo $j = 1, 2, ..., n$ se tiene*

$$\lambda_j^{\downarrow}(A) - \|B\| \leq \lambda_j^{\downarrow}(A+B) \leq \lambda_j^{\downarrow}(A) + \|B\|.$$

Demostración. Usando que $\|B\| = \text{máx}\left\{|\lambda_1^{\downarrow}(B)|, |\lambda_n^{\downarrow}(B)|\right\}$, y las desigualdades de Weyl (con $i = j$) resulta

$$\lambda_j^{\downarrow}(A) - \|B\| \leq \lambda_j^{\downarrow}(A) + \lambda_n^{\downarrow}(B) \leq \lambda_j^{\downarrow}(A+B) \leq \lambda_j^{\downarrow}(A) + \lambda_1^{\downarrow}(B) \leq \lambda_j^{\downarrow}(A) + \|B\|.$$

□

Definición 1.2.8. *Decimos que una función real f es* convexa de matrices de orden n *en un intervalo I si para todo $c \in [0,1]$ resulta*

$$f(cA + (1-c)B) \leq cf(A) + (1-c)f(B)$$

$\forall\ A, B \in M_n(\mathbb{C})$, Hermitianas con espectro contenido en I. Decimos que f es convexa de matrices *en I, si es convexa de matrices de orden n en I para todo $n \in \mathbb{N}$.*

Observación 1.2.9. Puede ocurrir que una función sea monótona en un intervalo I, pero no sea monótona de matrices en el mismo intervalo. Lo mismo puede ocurrir con la convexidad. Ilustremos esto con dos ejemplos.

Ejemplo 1.2.10. La función $f(t) = t^2$ no es monótona de matrices en $[0, \infty)$. Consideremos las matrices

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \text{ y } B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Es fácil ver que $B \geq A$ pero $B^2 - A^2$ no es positiva.

Ejemplo 1.2.11. La función $f(t) = t^3$ no es convexa de matrices en $[0, \infty)$. Tomando

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \text{ y } B = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix},$$

resulta

$$\frac{f(A) + f(B)}{2} - f\left(\frac{A + B}{2}\right) = \begin{pmatrix} 6 & 1 \\ 1 & 0 \end{pmatrix},$$

que no es positiva.

A continuación mostraremos algunos ejemplos de funciones monótonas de matrices. Para probar que lo son necesitaremos los siguientes lemas.

Lema 1.2.12. *Si $B \geq A$, entonces $X^*BX \geq X^*AX$ para toda matriz X.*

Demostración. Para todo vector $u \in \mathbb{C}^n$ se tiene

$$\langle u, X^*BXu \rangle = \langle Xu, BXu \rangle \geq \langle Xu, AXu \rangle = \langle u, X^*AXu \rangle.$$

□

Lema 1.2.13. *Si $B \geq A \geq 0$, y B es inversible, entonces $1 \geq \left\| A^{\frac{1}{2}} B^{-\frac{1}{2}} \right\|$.*

Demostración. Por el Lema 1.2.12, se tiene

$$I \geq B^{-\frac{1}{2}} A B^{-\frac{1}{2}} = (A^{\frac{1}{2}} B^{-\frac{1}{2}})^* (A^{\frac{1}{2}} B^{-\frac{1}{2}}),$$

por lo tanto

$$1 \geq \left\| A^{\frac{1}{2}} B^{-\frac{1}{2}} \right\|.$$

□

Proposición 1.2.14. *La función $f(t) = t^{\frac{1}{2}}$ es monótona de matrices en $[0, \infty)$.*

Demostración. Sea $B \geq A \geq 0$. Supongamos que B es inversible. Luego,

$$1 \geq \left\| A^{\frac{1}{2}} B^{-\frac{1}{2}} \right\| \geq \lambda_1^{\downarrow}(A^{\frac{1}{2}} B^{-\frac{1}{2}}) = \lambda_1^{\downarrow}(B^{-\frac{1}{4}} A^{\frac{1}{2}} B^{-\frac{1}{4}}).$$

Como $B^{-\frac{1}{4}} A^{\frac{1}{2}} B^{-\frac{1}{4}}$ es positiva (por el Lema 1.2.12), resulta $I \geq B^{-\frac{1}{4}} A^{\frac{1}{2}} B^{-\frac{1}{4}}$, y (otra vez por el Lema 1.2.12) se tiene $B^{\frac{1}{2}} \geq A^{\frac{1}{2}}$.

Para probar el caso general (o sea B no necesariamente inversible), consideramos $B + \epsilon I$ con $\epsilon > 0$. Como $B + \epsilon I$ es estrictamente positivo, $(B + \epsilon I)^{\frac{1}{2}} \geq A^{\frac{1}{2}}$. Haciendo tender ϵ a 0 obtenemos $B^{\frac{1}{2}} \geq A^{\frac{1}{2}}$. □

Lema 1.2.15. *Sean $A, B \in M_n(\mathbb{C})$, tales que $A \geq B > 0$ y $AB = BA$. Entonces $A^{-1} \leq B^{-1}$.*

Demostración. Es inmediato a partir de la diagonalización simultánea (Proposición 1.1.2). □

Proposición 1.2.16. *La función $f(t) = -\frac{1}{t}$ es monótona de matrices en $(0, \infty)$.*

Demostración. Sea $B \geq A > 0$. Luego, por el Lema 1.2.12, $I \geq B^{-\frac{1}{2}} A B^{-\frac{1}{2}}$. Por el Lema anterior nos queda $I \leq B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}}$. Finalmente usando el Lema 1.2.12 otra vez obtenemos $B^{-1} \leq A^{-1}$. □

Lema 1.2.17. *Sea f_n una sucesión de funciones reales tales que $f_n \to f$ puntualmente. Entonces $f_n(X) \to f(X)$ para toda matriz Hermitiana X.*

Demostración. Si D es una matriz diagonal, entonces $f_n(D) \to f(D)$ claramente. Si X es Hermitiana, $X = UDU^*$ con U unitaria y D diagonal, luego

$$f_n(X) = U f_n(D) U^* \to U f(D) U^* = f(X).$$

□

Proposición 1.2.18. *Sea f_n una sucesión de funciones convexas de matrices en un intervalo I, tal que $f_n \to f$ puntualmente en I. Entonces f es convexa de matrices en I.*

Demostración. Ante todo notemos que $f_n(A) \to f(A)$ (Lema 1.2.17). También notemos que el límite de matrices positivas es positivo (≥ 0). Entonces se tiene la conclusión tomando límite en

$$c f_n(A) + (1 - c) f_n(B) \geq f_n(cA + (1 - c)B).$$

□

Proposición 1.2.19. *Sea f_n una sucesión de funciones monótonas de matrices en un intervalo I, tal que $f_n \to f$ puntualmente en I. Entonces f es monótona de matrices en I.*

Demostración. Sean A, B matrices tales que $A \leq B$. Tomando límite en $f_n(A) \leq f_n(B)$, por el Lema 1.2.17, queda demostrada la proposición. $\square$

Proposición 1.2.20. *$f(t) = t^s$ es monótona de matrices en $I = [0, \infty)$ para $0 \leq s \leq 1$.*

Demostración. Sea s un racional diádico, i.e. $s = \frac{m}{2^n}$, con $n \in \mathbb{N}$ y $1 \leq m \leq 2^n$. Hacemos inducción en n.

Si $n = 1$, el resultado vale por la Proposición 1.2.14. Supongamos que es cierto para todos los racionales diádicos $\frac{m}{2^j}$, con $1 \leq j \leq n-1$. Sea $B \geq A$ y $s = \frac{m}{2^n}$.

Si $m \leq 2^{n-1}$, luego, por la hipótesis inductiva, $B^{\frac{m}{2^{n-1}}} \geq A^{\frac{m}{2^{n-1}}}$, y por la Proposición 1.2.14, $B^{\frac{m}{2^n}} \geq A^{\frac{m}{2^n}}$.

Si $m > 2^{n-1}$, sea $B \geq A \geq 0$. Luego por el Lema 1.2.16 $A^{-1} \geq B^{-1}$, y por el Lema 1.2.12 tenemos

$$B^{\frac{m}{2^n}} A^{-1} B^{\frac{m}{2^n}} \geq B^{\frac{m}{2^n}} B^{-1} B^{\frac{m}{2^n}} = B^{\frac{m}{2^{n-1}}-1}.$$

Del mismo modo tenemos

$$\begin{aligned} A^{-\frac{1}{2}} B^{\frac{m}{2^n}} A^{-1} B^{\frac{m}{2^n}} A^{-\frac{1}{2}} &\geq A^{-\frac{1}{2}} B^{\frac{m}{2^{n-1}}-1} A^{-\frac{1}{2}} \\ &\geq A^{-\frac{1}{2}} A^{\frac{m}{2^{n-1}}-1} A^{-\frac{1}{2}}, \end{aligned} \tag{1.3}$$

donde en (1.3) usamos la hipótesis inductiva. Esta desigualdad se puede reescribir como

$$(A^{-\frac{1}{2}} B^{\frac{m}{2^n}} A^{-\frac{1}{2}})^2 \geq A^{\frac{m}{2^{n-1}}-2}.$$

Luego, por la monotonía de la raíz cuadrada (Proposición 1.2.14) se tiene

$$A^{-\frac{1}{2}} B^{\frac{m}{2^n}} A^{-\frac{1}{2}} \geq A^{\frac{m}{2^n}-1},$$

luego $B^{\frac{m}{2^n}} \geq A^{\frac{m}{2^n}}$ (por el Lema 1.2.12, conjugando con $A^{\frac{1}{2}}$).

Hemos probado la proposición para los diádicos $s \in [0, 1]$. Por la densidad de los diádicos y la Proposición 1.2.19, vale la monotonía para todo $0 \leq s \leq 1$, y por la densidad de las matrices positivas estrictas en las positivas se tiene la proposición. $\square$

Proposición 1.2.21. *Toda combinación lineal positiva de funciones monótonas de matrices es una función monótona de matrices.*

Demostración. Sean $\{f_1, ..., f_n\}$ funciones monótonas de matrices y $\{\lambda_1, ..., \lambda_n\}$ números positivos. Entonces, dadas matrices $A \leq B$, resulta

$$\langle (\sum \lambda_i f_i(A) - \sum \lambda_i f_i(B))X, X\rangle = \sum \lambda_i \langle (f_i(A) - f_i(B))X, X\rangle \geq 0,$$

luego $\sum \lambda_i f_i(A) \geq \sum \lambda_i f_i(B)$. □

Observación 1.2.22. En particular, toda combinación convexa de funciones monótonas de matrices en un intervalo I también lo es.

De la misma manera que para las funciones monótonas se prueba la siguiente proposición.

Proposición 1.2.23. *Toda combinación lineal positiva de funciones convexas de matrices también lo es.*

Capítulo 2

Teorema de Loewner.

En este capítulo estudiaremos la derivabilidad de las funciones convexas, y la diferenciabilidad de funciones definidas sobre el espacio de matrices Hermitianas. De manera análoga al caso clásico, encontraremos relaciones entre la monotonía de una función f y la positividad de la matriz $f^{[1]}$ (la primera diferencia dividida de f, ver Definición 2.3.2), que a su vez está relacionada con el diferencial de f. Veremos también que la convexidad de f se relaciona con la monotonía de $f^{[1]}$. Hallaremos formas de obtener funciones monótonas de matrices a partir de funciones convexas de matrices y viceversa. Finalmente mostraremos una caracterización (obtenida por Loewner en 1934, publicada originalmente en [20]) de las funciones monótonas y convexas de matrices.

2.1. Regularización de funciones.

Definición 2.1.1. *Sea ϕ una función real C^∞ tal que*

1. $\phi(x) \geq 0 \ \forall x \in \mathbb{R}$.
2. $\phi(x) = \phi(-x) \ \forall x \in \mathbb{R}$.
3. $\int \phi(x)dx = 1$.
4. $sop(\phi) = [-1, 1]$.

Para todo $\epsilon > 0$, sea $\phi_\epsilon(x) = \frac{1}{\epsilon}\phi(\frac{x}{\epsilon})$. Es fácil ver que ϕ_ϵ verifica las condiciones 1, 2 y 3, y además $sop(\phi_\epsilon) = [-\epsilon, \epsilon]$. Las funciones ϕ_ϵ se denominan aproximaciones suaves a la identidad.

Definición 2.1.2. *Sea f una función localmente integrable. Llamamos* regularización de orden ϵ de f *a la función*

$$f_\epsilon(x) = (f * \phi_\epsilon)(x) = \int f(x-y)\phi_\epsilon(y)dy = \int f(x-\epsilon t)\phi(t)dt.$$

Proposición 2.1.3. *Sea f una función localmente integrable, entonces*

1. $f_\epsilon \in C^\infty$.

2. *Si el soporte de f está contenido en un compacto K, el soporte de f_ϵ esta contenido en un $\epsilon-$entorno de K.*

3. *Si f es continua en x_0, entonces* $\lim\limits_{\epsilon\to 0} f_\epsilon(x_0) = f(x_0)$.

4. *Si f tiene una discontinuidad del primer tipo (salto finito) en x_0, entonces* $\lim\limits_{\epsilon\to 0} f_\epsilon(x_0) = \frac{1}{2}(\lim\limits_{x\to x_0^+} f(x) + \lim\limits_{x\to x_0^-} f(x))$.

5. *Si f es derivable, entonces* $(f_\epsilon)' = (f')_\epsilon$.

Demostración. 1. Veamos que $f_\epsilon^{(n)}(x) = f * \phi_\epsilon^{(n)}(x)$:

$$\begin{aligned} \frac{f_\epsilon(x+h) - f_\epsilon(x)}{h} &= \int f(t)\left(\frac{\phi_\epsilon(x-t+h) - \phi_\epsilon(x-t)}{h}\right)dt \\ &= \int f(t)\phi'_\epsilon(x-t+c)dt, \qquad (2.1) \end{aligned}$$

para algún $c \in (0,h)$. Tomando $h \to 0$, resulta que la proposición es válida para $n = 1$. Este mismo argumento aplicado sucesivamente prueba el caso general.

2. Es inmediato por la definición de f_ϵ.

3. Dado $\epsilon_0 > 0$, sea $\delta > 0$ tal que $|f(x) - f(x_0)| \le \epsilon_0$ para todo $x \in (x_0-\delta, x_0+\delta)$. Entonces para todo $0 < \epsilon \le \delta$ se tiene

$$f(x_0) - \epsilon_0 = \int (f(x_0) - \epsilon_0)\phi_\epsilon(x_0 - y)dy \le f_\epsilon(x_0),$$

$$f(x_0) + \epsilon_0 = \int (f(x_0) + \epsilon_0)\phi_\epsilon(x_0 - y)dy \ge f_\epsilon(x_0).$$

4. Sin pérdida de generalidad, supongamos $x_0 = 0$. Sea $g(x) = f(x)$ si $x > 0$, $g(x) = f(-x)$, si $x < 0$, y $g(0) = \lim_{x\to 0^+} f(x)$. Por el ítem 3. $\lim_{\epsilon\to 0} g_\epsilon(0) = g(0)$. Definiendo de manera análoga h pero de modo tal que $h(x) = f(x)$ si $x < 0$, resulta

$$\begin{aligned}
\lim_{x\to 0^+} f(x) + \lim_{x\to 0^-} f(x) &= \lim_{\epsilon\to 0} g_\epsilon(0) + h_\epsilon(0) \\
&= \lim_{\epsilon\to 0} \int (g+h)(y)\phi_\epsilon(-y)dy \\
&= \lim_{\epsilon\to 0} \int (f(y) + f(-y))\phi_\epsilon(-y)dy \\
&= \lim_{\epsilon\to 0} 2 \int f(y)\phi_\epsilon(-y)dy \qquad (2.2) \\
&= 2\lim_{\epsilon\to 0} f_\epsilon(0),
\end{aligned}$$

donde (2.2) vale pues $\int f(-y)\phi_\epsilon(-y)dy = \int f(-y)\phi_\epsilon(y)dy = \int f(y)\phi_\epsilon(-y)dy$.

5. Demostración análoga a 1.

□

Proposición 2.1.4. *Sea f una función monótona de matrices en (a,b), entonces f_ϵ lo es en $(a+\epsilon, b-\epsilon)$.*

Demostración.

$$\begin{aligned}
f_\epsilon(x) &= \int_{-\epsilon}^{\epsilon} f(x-y)\phi_\epsilon(y)dy \\
&= \lim_{\|\pi\|\to 0} \sum_{i=1}^{n} f(x-\xi_i)\phi_\epsilon(\xi_i)(y_i - y_{i-1}),
\end{aligned}$$

donde π es la partición $\{y_0, ..., y_n\}$, $\|\pi\| = \max\{|y_i - y_{i-1}|\}$, y $\xi_i \in (y_{i-1}, y_i)$. En particular la igualdad vale tomando los ξ_i de modo tal que $\sum \phi_\epsilon(\xi_i)(y_i - y_{i-1}) = 1$. Luego f_ϵ es el límite puntual de una combinación lineal convexa de las funciones $f(x-\xi_i)$, que son monótonas de matrices en $(a+\epsilon, b-\epsilon)$. Entonces, por la Observación 1.2.22 y la Proposición 1.2.19, f_ϵ resulta monótona de matrices. □

2.2. Derivación de funciones convexas.

Proposición 2.2.1. *Sea f una función real convexa definida en un intervalo I. Entonces f es Lipschitz en todo intervalo cerrado contenido en el interior de I. En particular f es continua en I°.*

Demostración. Sean $[a,b] \subset I^\circ$, $F(x,y) = \frac{f(y)-f(x)}{y-x}$ con $x < y$, $x, y \in [a,b]$. Es fácil ver que F es creciente en cada coordenada, luego

$$\lim_{x \to a^+} F(a,x) \leq \frac{f(y)-f(x)}{y-x} \leq \lim_{y \to b^-} F(y,b). \tag{2.3}$$

Como $g(x) = F(a,x)$ es creciente, basta ver que es acotada inferiormente para concluir que $\lim_{x \to a^+} F(a,x)$ existe y es finito. Como $[a,b] \subset I^\circ$, $\exists\, c < a$ tal que $[c,b] \subset I^\circ$. Luego $F(c,a) \leq F(a,x)$ para todo $x \in (a,b)$. De manera análoga se prueba que $\lim_{y \to b^-} F(y,b)$ existe y es finito. Así, de (2.3), se deduce que existen $m, M \in \mathbb{R}$ tales que para todo $x, y \in [a,b]$, $x < y$

$$m \leq \frac{f(y)-f(x)}{y-x} \leq M,$$

de donde la proposición se deduce inmediatamente. $\square$

Corolario 2.2.2. *Sea f una función real convexa en un intervalo I. Entonces para todo $x_0 \in I^\circ$, existen las derivadas laterales de f.*

Demostración. Un razonamiento análogo al de la demostración de la Proposición 2.2.1 prueba que $F(x, x_0)$ es creciente y acotada superiormente, por lo que $f'_-(x_0) := \lim_{x \to x_0^-} \frac{f(x_0)-f(x)}{x_0 - x}$ existe y es finito. Análogamente se prueba que existe la otra derivada lateral $f'_+(x_0)$. $\square$

Corolario 2.2.3. *Sea f una función real convexa en un intervalo I. Entonces f es derivable excepto en un conjunto numerable E, y f' es continua en $I^\circ - E$.*

Demostración. Es fácil ver que

$$\lim_{x \nearrow x_0} f'_\pm(x) = f'_-(x_0) \text{ y } \lim_{x \searrow x_0} f'_\pm(x) = f'_+(x_0). \tag{2.4}$$

Entonces,

$$f'_+(x_0) \geq f'_-(x_0) = \lim_{x \nearrow x_0} f'_+(x), \tag{2.5}$$

pero como f'_+ es monótona, vale la igualdad en (2.5) en casi todo punto y por lo tanto f es derivable en casi todo punto. Sea E el conjunto donde f no es derivable. Como f' es monótona (en $I_0 - E$), resulta que f' sólo tiene discontinuidades de tipo salto finito. Esto podría ocurrir sólo en los puntos de discontinuidad de las derivadas laterales, pero por (2.4), resulta que en esos puntos las derivadas laterales no coinciden y por lo tanto los puntos de discontinuidad de f' estarían en E. $\square$

2.3. Propiedades de las funciones monótonas y convexas de matrices.

Definición 2.3.1. *Dadas* $A = (a_{ij}), B = (b_{ij}) \in M_n(\mathbb{C})$*, definimos el* producto de Schur *de* A *y* B *como*

$$A \circ B = (a_{ij} b_{ij}).$$

El producto de Schur, debe su nombre al trabajo [27], donde I. Schur prueba la Proposición 2.3.15, entre otras cosas. Este producto también es conocido como *producto de Hadamard*, debido a que en [11] J. Hadamard prueba un resultado sobre el producto coeficiente a coeficiente de dos funciones analíticas. Si bien en ese trabajo no hay nada sobre matrices, se cree que Von Neumann acuñó el término por la influencia de la obra de Hadamard sobre los analistas.

Definición 2.3.2. *Dada una función* $f \in C^1(I)$*, con* $I = (-1, 1)$*, definimos la* primera diferencia dividida *de* f *en* (λ, μ) $(\lambda, \mu \in I)$ *como*

$$f^{[1]}(\lambda, \mu) = \frac{f(\lambda) - f(\mu)}{\lambda - \mu}, \ \text{si } \lambda \neq \mu$$

$$f^{[1]}(\lambda, \lambda) = f'(\lambda).$$

Si $\Lambda = diag(\lambda_1, ..., \lambda_n)$*, con* $\lambda_i \in I \ \forall i$*, notamos* $f^{[1]}(\Lambda) = (f^{[1]}(\lambda_i, \lambda_j))_{i,j}$*. Si* A *es Hermitiana,* $A = U\Lambda U^*$ *definimos* $f^{[1]}(A) = U f^{[1]}(\Lambda) U^*$*.*

Definición 2.3.3. *Una función* f*, definida en las matrices Hermitianas a valores en* $\mathbb{C}$*, es* diferenciable Fréchet *en* A*, si existe una transformación lineal* $Df(A)$ *sobre el espacio de matrices Hermitianas tal que*

$$\|f(A + H) - f(A) - Df(A)(H)\| = o(\|H\|).$$

Observación 2.3.4. Si f es diferenciable en A, resulta

$$Df(A)(H) = \frac{d}{dt}|_{t=0} f(A + tH). \tag{2.6}$$

Teorema 2.3.5 (Teorema del Valor Medio, versión 1)**.** *Sean* $I = [a, b] \subset \mathbb{R}$*,* Y *un espacio de Banach,* $f : I \to Y$ *derivable tal que* $\|f'(t)\| \leq M$ *para todo* $t \in I$*, entonces*

$$\|f(b) - f(a)\| \leq M\|b - a\|.$$

Demostración. Dado $\epsilon > 0$, sea

$$A = \{\xi \in I : \|f(\zeta) - f(a)\| \leq M(\zeta - a) + \epsilon(\zeta - a)\ \forall a \leq \zeta < \xi\}$$

Queremos ver que $A = [a, b]$, independientemente del ϵ dado. Claramente $a \in A$, y A es un intervalo cerrado incluido en $[a, b]$. Luego, $A = [a, \gamma]$. Veamos que $\gamma = b$.

Supongamos $\gamma < b$. Como f es derivable en γ, se tiene que

$$\lim_{t \to 0} \frac{f(\gamma + t) - f(\gamma)}{t} = f'(\gamma)$$

luego, dado ϵ, para todo $t > 0$ suficientemente chico se tiene

$$\|f(\gamma + t) - f(\gamma) - f'(\gamma)t\| \leq \epsilon t.$$

Así resulta

$$\|f(\gamma + t) - f(\gamma)\| \leq \|f'(\gamma)\|t + \epsilon t \leq Mt + \epsilon t. \tag{2.7}$$

Entonces, por (2.7) y porque $\gamma \in A$ obtenemos

$$\begin{aligned} \|f(\gamma + t) - f(a)\| &\leq \|f(\gamma + t) - f(\gamma)\| + \|f(\gamma) - f(a)\| \\ &\leq Mt + \epsilon t + M(\gamma - a) + \epsilon(\gamma - a) \\ &= M(\gamma + t - a) + \epsilon(\gamma + t - a), \end{aligned}$$

por lo que $\gamma + t \in A$, lo cual implica que $[a, \gamma + t] \subseteq A$, contradiciendo la suposición $A = [a, \gamma]$. Entonces $\gamma = b$. □

Teorema 2.3.6 (Teorema del Valor Medio, versión 2)**.** *Sean X, Y espacios de Banach, $U \subset X$ convexo, y $f : U \to Y$ diferenciable. Dados $a, b \in U$, se tiene*

$$\|f(b) - f(a)\| \leq \|b - a\| \sup_{u \in \mathcal{L}} \|Df(u)\|$$

donde $\mathcal{L}$ es el segmento que une a y b.

Demostración. Sea $g : [0, 1] \to Y$, $g(t) = f(a + t(b - a))$, entonces

$$g'(t) = Df(a + t(b - a))(b - a).$$

Por el Teorema 2.3.5 resulta

$$\|f(b) - f(a)\| \leq \sup_{0 \leq t \leq 1} \|Df(a + t(b - a))(b - a)\| \leq \sup_{u \in \mathcal{L}} \|Df(u)\| \|(b - a)\|.$$

□

Lema 2.3.7. *Sea f un polinomio. Para toda matriz diagonal $\Lambda = diag(\lambda_1, ..., \lambda_s)$ y toda matriz Hermitiana $H = (h_{ij})$, se tiene*

$$Df(\Lambda)(H) = f^{[1]}(\Lambda) \circ H. \tag{2.8}$$

Demostración. Ambos lados de la igualdad (2.8) son lineales en f. Luego, basta probar la igualdad para los polinomios $f(t) = t^n$ con $n \in \mathbb{N}$. Por la igualdad (2.6) resulta

$$\begin{aligned} Df(\Lambda)(H) &= \frac{d}{dt}|_{t=0}(\Lambda + tH) \\ &= \lim_{h\to 0} \frac{(\Lambda + hH)^n - \Lambda^n}{h} \\ &= \sum_{k=1}^{n} \Lambda^{k-1} H \Lambda^{n-k} \\ &= \left(\sum_{k=1}^{n} \lambda_i^{k-1} \lambda_j^{n-k} h_{ij} \right)_{ij}, \end{aligned}$$

y esta última matriz es justamente $f^{[1]}(\Lambda) \circ H$. $\square$

Corolario 2.3.8. *Sean $A = U\Lambda U^*$, con U unitaria y Λ diagonal, y f un polinomio. Entonces, para toda H Hermitiana se tiene,*

$$Df(A)(H) = U[f^{[1]}(\Lambda) \circ (U^* H U)]U^*.$$

Demostración. Como $f(U\Lambda U^* + tH) = Uf(\Lambda + tU^* HU)U^*$ resulta

$$\frac{d}{dt}|_{t=0} f(U\Lambda U^* + tH) = U[\frac{d}{dt}|_{t=0} f(\Lambda + tU^* HU)]U^*,$$

y aplicando (2.8) se obtiene el corolario. $\square$

Definición 2.3.9. *Decimos que una norma $\|.\|$ sobre el espacio de matrices cuadradas de dimensión n es* unitariamente invariante *si para toda matriz A y todo par de matrices unitarias U, V, se tiene que*

$$\|A\| = \|UAV\|.$$

Ejemplo 2.3.10. Un ejemplo sencillo de norma unitariamente invariante es la familia de *normas p de Schatten*, definidas del siguiente modo:

$$\|A\|_p = \left(\sum_{i=1}^{n} s_j(A)^p \right)^{\frac{1}{p}}, \text{ si } 1 \leq p < \infty,$$

$$\text{y } \|A\|_\infty = s_1(A) = \|A\|$$

donde $s(A) = (s_1(A), ..., s_n(A))$ es el vector formado por los autovalores de $|A| = (A^*A)^{\frac{1}{2}}$ ordenados en forma decreciente. A los autovalores de $|A|$ se los denomina *valores singulares de* A. Como los autovalores no dependen de la base es fácil ver que las normas p de Schatten son unitariamente invariantes.

Observación 2.3.11. $\|A\|_2 = Tr(A^*A)^{\frac{1}{2}}$, luego $\|A\|_2 = \left(\sum_{i,j} |a_{ij}|^2\right)^{\frac{1}{2}}$.

Teorema 2.3.12. *Sean, $I = (-1, 1)$, $f \in C^1(I)$ y A una matriz Hermitiana con espectro contenido en I. Entonces, para toda H Hermitiana,*

$$Df(A)(H) = U[f^{[1]}(\Lambda) \circ (U^*HU)]U^*,$$

con $A = U^*\Lambda U$, U unitaria y Λ diagonal.

Demostración. Si f es un polinomio la igualdad vale por el lema anterior. Usaremos la densidad de los polinomios en $C^1(I)$ para probar el caso general.

Notemos $\mathcal{D}f(A)(H) = U[f^{[1]}(\Lambda) \circ (U^*HU)]U^*$. Para cada $f \in C^1(I)$, $\mathcal{D}f(A)$ es un operador lineal en las matrices Hermitianas. Se tiene

$$\|\mathcal{D}f(A)(H)\|_2 = \|f^{[1]}(\Lambda) \circ (U^*HU)\|_2$$

(pues $\|.\|_2$ es una norma unitariamente invariante). Todas las coordenadas de la matriz $f^{[1]}(\Lambda)$ están acotadas en módulo por

$$\text{máx}\{|f'(t)| : |t| \leq \|A\|\}$$

(por el teorema del valor medio). Luego

$$\|\mathcal{D}f(A)(H)\|_2 \leq \text{máx}\{|f'(t)| : |t| \leq \|A\|\}\|H\|_2.$$

Sea H Hermitiana de norma suficientemente pequeña para que el espectro de $A+H$ esté contenido en I. Sea $[a, b] \subseteq I$ tal que los autovalores de A y de $A + H$ están incluidos en $[a, b]$. Sea (f_n) una sucesión de polinomios tal que $f_n \to f$ y $f_n' \to f'$ uniformemente en $[a, b]$. Sea $\mathcal{L}$ el segmento que une A y $A + H$. Por el teorema del valor medio para la derivada de Fréchet, se tiene

$$\begin{aligned} \|f_m(A + H) - f_n(A + H) - (f_m(A) - f_n(A))\| \\ \leq \|H\| \sup_{X \in \mathcal{L}} \|Df_m(X) - Df_n(X)\| \\ = \|H\| \sup_{X \in \mathcal{L}} \|\mathcal{D}f_m(X) - \mathcal{D}f_n(X)\|, \end{aligned} \tag{2.9}$$

donde la igualdad (2.9) vale pues $D = \mathcal{D}$ en polinomios.

Dado $\epsilon > 0$, existe n_0 tal que si $m, n \geq n_0$, entonces

$$\sup_{X \in \mathcal{L}} \|\mathcal{D}f_m(X) - \mathcal{D}f_n(X)\| \leq \frac{\epsilon}{3} \tag{2.10}$$

$$\text{y } \|\mathcal{D}f_n(A) - \mathcal{D}f(A)\| \leq \frac{\epsilon}{3}. \tag{2.11}$$

Tomando $m \to \infty$, por (2.9) y (2.10) resulta

$$\|f(A+H) - f(A) - (f_n(A+H) - f_n(A))\| \leq \frac{\epsilon}{3}\|H\|. \tag{2.12}$$

Si $\|H\|$ es suficientemente chico, por la definición de derivada de Fréchet resulta

$$\|f_n(A+H) - f_n(A) - \mathcal{D}f_n(A)(H)\| \leq \frac{\epsilon}{3}. \tag{2.13}$$

Entonces tenemos

$$\begin{aligned} \|f(A+H) - f(A) - \mathcal{D}f(A)(H)\| \;\leq\; & \|f(A+H) - f(A) - (f_n(A+H) - f_n(A))\| \\ & + \|f_n(A+H) - f_n(A) - \mathcal{D}f_n(A)(H)\| \\ & + \|(\mathcal{D}f(A) - \mathcal{D}f_n(A))(H)\|, \end{aligned}$$

y por las ecuaciones (2.11),(2.12) y (2.13), concluimos que si $\|H\|$ es suficientemente chico, se tiene

$$\|f(A+H) - f(A) - \mathcal{D}f(A)(H)\| \leq \epsilon\|H\|,$$

y esto implica que $Df(A) = \mathcal{D}f(A)$. $\square$

Observación 2.3.13. Sea $t \mapsto A(t)$ una aplicación C^1 del intervalo $[0,1]$ al espacio de matrices Hermitianas con espectro contenido en $I = (-1,1)$. Sea $f \in C^1(I)$ y sea $F(t) = f(A(t))$. Entonces, por la regla de la cadena, resulta $DF(t) = Df(A(t))(A'(t))$. Además, por el teorema anterior, se tiene

$$F(1) - F(0) = \int_0^1 f^{[1]}(A(t)) \circ A'(t)dt,$$

donde para cada t, el producto de Schur se considera en una base donde $A(t)$ es diagonal.

Lema 2.3.14. *Sea $A \in M_n(\mathbb{C})$, son equivalentes:*

1. $A \geq 0$.

2. Existen $w_i, ..., w_r \in \mathbb{C}^n$ *tales que* $A = \sum_{i=1}^r w_i w_i^*$.

Demostración. $1 \Rightarrow 2)$ Sea $\{v_1, ..., v_n\}$ una b.o.n. de autovectores de A asociados a los autovalores (no necesariamente distintos) $\lambda_1, ..., \lambda_n$. Luego, si $z \in \mathbb{C}^n$,

$$Az = A\left(\sum_{i=1}^n \langle z, v_i\rangle v_i\right) = \sum_{i=1}^n \langle z, v_i\rangle A v_i = \sum_{i=1}^n \langle z, v_i\rangle \lambda_i v_i.$$

Tomando $w_i = \lambda_i^{\frac{1}{2}} v_i$ se tiene que $Az = \sum_{i=1}^n w_i w_i^* z$.

$2 \Rightarrow 1)$ Es trivial pues $w_i w_i^*$ es positiva para todo i y la suma de matrices positivas también lo es. $\square$

Proposición 2.3.15. *Sean* $A, B \in M_n(\mathbb{C})$ *positivas. Entonces* $A \circ B \geq 0$. *Si* $A, B > 0$, *entonces* $A \circ B > 0$.

Demostración. Sean $w_1, ..., w_r \in \mathbb{C}^n$ (vectores columna) tales que $\sum w_i w_i^* = A$ (existen por el Lema 2.3.14). Como el producto de Schur es distributivo (trivial), basta ver que $vv^* \circ B \geq 0$ para todo $v \in \mathbb{C}^n$. Pero

$$vv^* \circ B = (v_i \overline{v_j} B_{ij})_{ij} = diag(v) B diag(v)^* \geq 0.$$

Si $A, B > 0$, existe $a \in \mathbb{R}, a > 0$ tal que $A \geq aI$, $B \geq aI$. Por la distributividad del producto Schur y, aplicando dos veces el caso demostrado, tenemos

$$A \circ B \geq A \circ aI \geq aI \circ aI = a^2 I > 0.$$

$\square$

Teorema 2.3.16. *Sean* $I = (-1, 1)$, $f \in C^1(I)$. *Entonces* f *es monótona de matrices en* I *si y sólo si, para toda matriz Hermitiana* A *con todos sus autovalores en* I, *la matriz* $f^{[1]}(A)$ *es positiva.*

Demostración. Sea f una función monótona de matrices y A una matriz Hermitiana con espectro incluido en I. Sea H la matriz $H = (h_{ij})$ tal que $h_{ij} = 1\ \forall i, j$. La matriz H es positiva, pues dado un vector $x = (x_i)$ resulta $\langle Hx, x\rangle = (\sum x_i)^2$. Luego $A + tH \geq A$ si $t \geq 0$ y, por lo tanto, $f(A + tH) - f(A) \geq 0$ para $t \geq 0$ (suficientemente pequeño para que el espectro de $A + tH$ esté contenido en I). Así, resulta $Df(A)(H) \geq 0$ y, por el Teorema 2.3.12, $f^{[1]}(A) \circ H \geq 0$. Pero por la elección particular de H, $f^{[1]}(A) \circ H = f^{[1]}(A)$, por lo que $f^{[1]}(A)$ resulta positiva.

Veamos la implicación inversa. Sean $A \leq B$ matrices Hermitianas con espectros contenidos en I. Sea $A(t) = (1-t)A + tB$, $t \in [0, 1]$. Entonces $A(t)$ también tiene su espectro en I (por el Corolario 1.2.4). Luego, por hipótesis, $f^{[1]}(A(t)) \geq 0$ para todo t. Como el producto de Schur de matrices positivas es positivo, $f^{[1]}(A(t)) \circ A'(t) \geq 0$ para todo t. Luego, por la Observación 2.3.13, $f(B) - f(A) = F(1) - F(0) \geq 0$. □

Lema 2.3.17. *Si K es una contracción, i.e. $\|K\| \leq 1$, entonces*

$$K^*(I - KK^*)^{\frac{1}{2}} = (I - K^*K)^{\frac{1}{2}} K^*.$$

Demostración. Usando la descomposición polar $K = UP$,

$$\begin{aligned} K^*(I - KK^*)^{\frac{1}{2}} &= PU^*(I - UPPU^*)^{\frac{1}{2}} = PU^*U(I - P^2)^{\frac{1}{2}}U^* \\ &= (I - P^2)^{\frac{1}{2}}PU^* = (I - PU^*UP)^{\frac{1}{2}}PU^* = (I - K^*K)^{\frac{1}{2}}K^*. \end{aligned}$$

□

Una consecuencia inmediata es el siguiente lema.

Lema 2.3.18. *Si K es una contracción en $\mathcal{H}$, entonces las matrices*
$$U = \begin{pmatrix} K & (I - KK^*)^{\frac{1}{2}} \\ (I - K^*K)^{\frac{1}{2}} & -K^* \end{pmatrix},$$
$$V = \begin{pmatrix} K & -(I - KK^*)^{\frac{1}{2}} \\ (I - K^*K)^{\frac{1}{2}} & K^* \end{pmatrix}$$ *son unitarias en $\mathbb{C}^n \oplus \mathbb{C}^n$.*

Demostración. Simplemente calcular U^*U y V^*V usando el lema anterior.

□

Observación 2.3.19. En particular, tomando $K = \lambda^{\frac{1}{2}} I$ con $0 \leq \lambda \leq 1$ resulta que la matriz
$$W = \begin{pmatrix} \lambda^{\frac{1}{2}} I & -(1-\lambda)^{\frac{1}{2}} I \\ (1-\lambda)^{\frac{1}{2}} I & \lambda^{\frac{1}{2}} I \end{pmatrix}$$
es unitaria.

El siguiente es un resultado de F. Hansen y G. Pedersen, publicado en [12].

Teorema 2.3.20 (Desigualdad de Jensen-Hansen-Pedersen, versión 1)**.** *Sea I un intervalo tal que $0 \in I$, y $f : I \to \mathbb{R}$. Son equivalentes:*

1. *f es convexa de matrices en I y $f(0) \leq 0$.*

2. *$f(K^*AK) \leq K^*f(A)K$, $\forall K$ contracción y $\forall A$ Hermitiana con espectro en I.*

3. *$f(K_1^*AK_1 + K_2^*AK_2) \leq K_1^*f(A)K_1 + K_2^*f(A)K_2$ para todo par de matrices K_1, K_2 tal que $K_1^*K_1 + K_2^*K_2 \leq I$ y toda matriz Hermitiana A con espectro en I.*

4. *$f(PAP) \leq Pf(A)P$ para toda proyección P y A matriz Hermitiana con espectro en I.*

Demostración. $(1 \Rightarrow 2)$.

Sean $T = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$, U, V los definidos en el Lema 2.3.18, sea $L = (I - KK^*)^{\frac{1}{2}}$. Luego,

$$U^*TU = \begin{pmatrix} K^*AK & K^*AL \\ LAK & LAL \end{pmatrix}, \quad V^*TV = \begin{pmatrix} K^*AK & -K^*AL \\ -LAK & LAL \end{pmatrix}.$$

Entonces,

$$\begin{pmatrix} K^*AK & 0 \\ 0 & LAL \end{pmatrix} = \frac{U^*TU + V^*TV}{2}.$$

Así,

$$\begin{aligned} \begin{pmatrix} f(K^*AK) & 0 \\ 0 & f(LAL) \end{pmatrix} &= f\left(\frac{U^*TU + V^*TV}{2}\right) \leq \frac{f(U^*TU) + f(V^*TV)}{2} \\ &= \frac{U^*f(T)U + V^*f(T)V}{2} \\ &= \frac{1}{2}\left\{U^*\begin{pmatrix} f(A) & 0 \\ 0 & f(0) \end{pmatrix}U + V^*\begin{pmatrix} f(A) & 0 \\ 0 & f(0) \end{pmatrix}V\right\} \\ &\leq \frac{1}{2}\left\{U^*\begin{pmatrix} f(A) & 0 \\ 0 & 0 \end{pmatrix}U + V^*\begin{pmatrix} f(A) & 0 \\ 0 & 0 \end{pmatrix}V\right\} \\ &= \begin{pmatrix} K^*f(A)K & 0 \\ 0 & Lf(A)L \end{pmatrix}. \end{aligned}$$

$(2 \Rightarrow 3)$.

Sea $T = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, $K = \begin{pmatrix} K_1 & 0 \\ K_2 & 0 \end{pmatrix}$, luego K es una contracción.

Como

$$K^*TK = \begin{pmatrix} K_1^*AK_1 + K_2^*BK_2 & 0 \\ 0 & 0 \end{pmatrix},$$

resulta

$$\begin{pmatrix} f(K_1^*AK_1 + K_2^*BK_2) & 0 \\ 0 & f(0) \end{pmatrix} = f(K^*TK) \leq K^*f(T)K$$
$$= \begin{pmatrix} K_1^*f(A)K_1 + K_2^*f(B)K_2 & 0 \\ 0 & 0 \end{pmatrix}.$$

$(3 \Rightarrow 4)$.
Trivial tomando en $K_1 = P, K_2 = 0$ en el ítem 3.
$(4 \Rightarrow 1)$.
Sean A, B matrices Hermitianas con espectro en I, $0 \leq \lambda \leq 1$, $T = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, $P = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$, y W la matriz unitaria definida en la Observación 2.3.19.
Entonces
$$PW^*TWP = \begin{pmatrix} \lambda A + (1 - \lambda B) & 0 \\ 0 & 0 \end{pmatrix}.$$
Luego,
$$\begin{pmatrix} f(\lambda A + (1 - \lambda B)) & 0 \\ 0 & f(0) \end{pmatrix} = f(PW^*TWP)$$
$$\leq Pf(W^*TW)P = PW^*f(T)WP$$
$$= \begin{pmatrix} \lambda f(A) + (1 - \lambda)f(B) & 0 \\ 0 & 0 \end{pmatrix},$$
donde la última desigualdad vale por el ítem 4, lo que prueba que f es convexa de matrices y $f(0) \leq 0$.

$\square$

Lema 2.3.21. *Sea I un intervalo que contiene al 0 y $f : I \to \mathbb{R}$ tal que $f(0) \leq 0$. Entonces para toda matriz Hermitiana A con espectro en I y todo proyector P se tiene*
$$f(PAP) \leq Pf(PAP) = Pf(PAP)P.$$

Demostración. Sea $V = Ran(P) \subseteq \mathbb{C}^n$, entonces respecto a una b.o.n. tal que los primeros vectores son b.o.n. de V y el resto b.o.n. de $V^{\perp}$, PAP se representa por bloques como $\begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix}$.
Luego
$$f(PAP) = \begin{pmatrix} f(B) & 0 \\ 0 & f(0) \end{pmatrix} \leq \begin{pmatrix} f(B) & 0 \\ 0 & 0 \end{pmatrix} = Pf(PAP).$$

La igualdad del lema es trivial considerando las matrices por bloques. $\square$

Lema 2.3.22. *Sea $f : [0, \infty) \to \mathbb{R}$ continua. Entonces para toda matriz positiva A, y toda proyección P se tiene*

$$f(A^{\frac{1}{2}}PA^{\frac{1}{2}})A^{\frac{1}{2}}P = A^{\frac{1}{2}}Pf(PAP).$$

Demostración. Primero lo probamos para $f(t) = t^n$, por inducción en n. Si $n = 1$ es trivial. Supongamos que vale $\forall n \leq n_0$, veamos que vale para $n_0 + 1$:

$$(A^{\frac{1}{2}}PA^{\frac{1}{2}})^{n_0+1}A^{\frac{1}{2}}P = (A^{\frac{1}{2}}PA^{\frac{1}{2}})A^{\frac{1}{2}}P(PAP)^{n_0} = A^{\frac{1}{2}}P(PAP)^{n_0+1}$$

donde en la primera igualdad usamos la hipótesis inductiva.

El caso general se deduce del Lema 1.2.17, usando la densidad de los polinomios en las funciones continuas sobre compactos.

$\square$

Teorema 2.3.23. *Sea $f : [0, \alpha) \to \mathbb{R}$ continua. Son equivalentes:*

1. *f es convexa de matrices en $[0, \alpha)$ y $f(0) \leq 0$.*
2. *$g(t) = \frac{f(t)}{t}$ es monótona de matrices en $(0, \alpha)$.*

Demostración. $(1 \Rightarrow 2)$.

Sea $0 < A \leq B$, luego $0 < A^{\frac{1}{2}} \leq B^{\frac{1}{2}}$ (por el Lema 1.2.14) y $B^{-\frac{1}{2}}A^{\frac{1}{2}}$ es una contracción (por el Lema 1.2.13). Usando la Proposición 2.3.20,

$$f(A) = f(A^{\frac{1}{2}}B^{-\frac{1}{2}}BB^{-\frac{1}{2}}A^{\frac{1}{2}}) \leq A^{\frac{1}{2}}B^{-\frac{1}{2}}f(B)B^{-\frac{1}{2}}A^{\frac{1}{2}}$$

y por el Lema 1.2.12,

$$A^{-\frac{1}{2}}f(A)A^{-\frac{1}{2}} \leq B^{-\frac{1}{2}}f(B)B^{-\frac{1}{2}}$$

así queda $A^{-1}f(A) \leq B^{-1}f(B)$.

$(2 \Rightarrow 1)$.

Sea $g(t) = \frac{f(t)}{t}$ monótona en $(0, \alpha)$, veamos que f satisface la condición 4 del Teorema 2.3.20 (lo cual implica que $f(0) \leq 0$ y que f es convexa de matrices).

Sea P una proyección y A un operador positivo con espectro en $(0, \alpha)$. Entonces existe $\epsilon_0 > 0$ tal que $(1+\epsilon)A$ también tiene su espectro en $(0, \alpha)$ para todo $0 < \epsilon \leq \epsilon_0$. Como $P + \epsilon I \leq (1+\epsilon)I$, tenemos que $A^{\frac{1}{2}}(P+\epsilon)IA^{\frac{1}{2}} \leq (1+\epsilon)A$ (por el Lema 1.2.12). Usando la monotonía de g obtenemos

$$A^{-\frac{1}{2}}(P + \epsilon I)^{-1}A^{-\frac{1}{2}}f(A^{\frac{1}{2}}(P + \epsilon I)A^{\frac{1}{2}}) \leq (1 + \epsilon)^{-1}A^{-1}f((1 + \epsilon)A).$$

Por el Lema 1.2.12, conjugando con $(P+\epsilon I)A^{\frac{1}{2}}$ obtenemos

$$A^{-\frac{1}{2}}f(A^{\frac{1}{2}}(P+\epsilon I)A^{\frac{1}{2}})A^{\frac{1}{2}}(P+\epsilon I) \leq (1+\epsilon)^{-1}(P+\epsilon I)f((1+\epsilon)A)(P+\epsilon I).$$

Tomando $\epsilon \to 0$ queda

$$A^{-\frac{1}{2}}f(A^{\frac{1}{2}}PA^{\frac{1}{2}})A^{\frac{1}{2}}P \leq Pf(A)P.$$

Usando el Lema 2.3.22, resulta $Pf(PAP) \leq Pf(A)P$ y, por el Lema 2.3.21, resulta $f(PAP) \leq Pf(A)P$, que es justamente el ítem 4.

□

Proposición 2.3.24. *$f(t) = -t^s$ es convexa de matrices en $[0,\infty)$ para $0 < s < 1$.*

Demostración. Como para $0 < r < 1$ se tiene que $g(t) = t^r$ y $h(t) = -\frac{1}{t}$ son monótonas de matrices en $(0,\infty)$ (Proposiciones 1.2.20 y 1.2.16), se tiene que $h \circ g(t) = -\frac{1}{t^r}$ es monótona de matrices en $(0,\infty)$. Luego por la Proposición 2.3.23, $-t^{1-r}$ es convexa de matrices. Cambiando s por $1-r$, se tiene la conclusión. □

Proposición 2.3.25. *$f(t) = t\log t$ es convexa de matrices en $(0,\infty)$.*

Demostración. Dado $1 > r > 0$ sea $g_r(t) = \frac{t(t^r-1)}{r}$, con $g_r(0) = 0$, es fácil ver que g_r es continua en $[0,\infty)$. Veamos que g_r es convexa de matrices. Por la Proposición 2.3.23 basta ver que la función $h(x) = \frac{t^r-1}{r}$ es monótona de matrices, pero esto es cierto por la Proposición 1.2.20, luego g_r es convexa de matrices. Ahora, usando que $\lim\limits_{r\to 0} g_r(t) = t\log t$, y que límite de convexas de matrices también lo es (Proposición 1.2.18) queda demostrada la proposición. □

Lema 2.3.26. *Si f es continua y monótona de matrices en (-1,1), entonces $g_\lambda(t) = (t+\lambda)f(t)$ es convexa de matrices $\forall\lambda \in [0,1]$.*

Demostración. Supongamos primero que f es continua y monótona de matrices en $[-1,1]$. Entonces la función $f(t-1)$ es monótona de matrices en $[0,2)$. Sea $g(t) = tf(t-1)$. Resulta $g(0) = 0$ y $\frac{g(t)}{t}$ monótona de matrices en $(0,2)$. Luego, por el Teorema 2.3.23, $g(t)$ es convexa de matrices en $[0,2)$. Entonces la función $h_1(t) = g(t+1) = (t+1)f(t)$ es convexa de matrices en $[-1,1)$. El mismo argumento aplicado a $-f(-t)$ (que es monótona de matrices en $[-1,1]$ porque f lo es), permite concluir que $h_2(t) = -(t+1)f(-t)$ es convexa de matrices en $[-1,1)$. Como cambiar t por $-t$ preserva la convexidad (trivial), la función $h_3(t) = h_2(-t) = (t-1)f(t)$ es convexa de matrices en $(-1,1]$. Para $\lambda \in [-1,1]$, $g_\lambda(t) = \frac{1+\lambda}{2}h_1(t) + \frac{1-\lambda}{2}h_3(t)$ es una combinación convexa de funciones convexas de matrices, por lo tanto g_λ es convexa de matrices en $(-1,1)$.

Veamos el caso general, o sea f continua y monótona en $(-1,1)$. Para $\epsilon > 0$, sea $f_\epsilon(t) = f((1-\epsilon)t)$. Entonces f_ϵ es continua y monótona de matrices en $[-1,1]$. Por el caso anterior, resulta $(t+\lambda)f_\epsilon(t)$ convexa de matrices en $(-1,1)$. Pero

$$\lim_{\epsilon\to 0}(t+\lambda)f_\epsilon(t) = (t+\lambda)f(t),$$

y como el límite puntual de funciones convexas de matrices también lo es, el teorema queda demostrado. □

Teorema 2.3.27. *Toda función f monótona de matrices en $I = (-1,1)$ es continuamente diferenciable.*

Demostración. Sea $0 < \epsilon < 1$, y sean ϕ_ϵ aproximaciones a la identidad. Sea $f_\epsilon = f * \phi_\epsilon$. $f_\epsilon \in C^\infty(-1+\epsilon, 1-\epsilon)$, y es monótona de matrices (por Proposición 2.1.4). Sea $\tilde{f}(t) = \lim_{\epsilon\to 0} f_\epsilon(t)$. Entonces

$$\tilde{f}(t) = \frac{1}{2}(\lim_{x\to t^+} f(x) + \lim_{x\to t^-} f(x)).$$

Sea $g_\epsilon(t) = (t+1)f_\epsilon(t)$. Por el Lema 2.3.26, g_ϵ es convexa de matrices. Entonces $\tilde{g}(t) = \lim_{\epsilon\to 0} g_\epsilon(t)$ es convexa de matrices. Pero toda función convexa (en un intervalo abierto) es continua, luego $\tilde{g}$ lo es. Como $\tilde{g}(t) = (t+1)\tilde{f}(t)$ y $t+1 > 0$ si $t \in I$, resulta que $\tilde{f}$ es continua. Luego $f(t) = \tilde{f}(t)$ y, por lo tanto f es continua.

Sea $g(t) = (t+1)f(t)$. Entonces $g(t)$ es una función convexa en I (por el Lema 2.3.26). Luego g es derivable a derecha e izquierda (ver Corolario 2.2.2) y las derivadas laterales satisfacen

$$g'_-(t) \le g'_+(t), \ \lim_{s\to t^+} g'_\pm(s) = g'_+(t), \ \lim_{s\to t^-} g'_\pm(s) = g'_-(t). \tag{2.14}$$

Como $g'_\pm(t) = f(t) + (t+1)f'_\pm(t)$ y $t+1 > 0$, las derivadas laterales de f también satisfacen las relaciones (2.14).

Sea $A = \begin{pmatrix} s & 0 \\ 0 & t \end{pmatrix}$, con $-1 < s,t < 1$. Para ϵ suficientemente chico $s,t \in (-1+\epsilon, 1-\epsilon)$. Como f_ϵ es monótona de matrices en este intervalo, por el Teorema 2.3.16, la matriz $f_\epsilon^{[1]}(A)$ es positiva, por lo tanto su determinante también es positivo, lo que implica que

$$\left(\frac{f_\epsilon(s) - f_\epsilon(t)}{s-t}\right)^2 \le f'_\epsilon(s)f'_\epsilon(t). \tag{2.15}$$

Tomando $\epsilon \to 0$, como $f_\epsilon \to f$ uniformemente sobre compactos, $f_\epsilon(s)-f_\epsilon(t)$ converge a $f(s)-f(t)$. Además $f'_\epsilon(t)$ tiende a $\frac{1}{2}[f'_+(t)+f'_-(t)]$. Luego, tomando $\epsilon \to 0$ en (2.15),

$$\left(\frac{f(s)-f(t)}{s-t}\right)^2 \leq \frac{1}{4}[f'_+(s)+f'_-(s)][f'_+(t)+f'_-(t)].$$

Tomando $s \to t^+$, y usando que las derivadas de f verifican condiciones del tipo (2.14), resulta

$$[f'_+(t)]^2 \leq \frac{1}{4}[f'_+(t)+f'_+(t)][f'_+(t)+f'_-(t)],$$

lo que implica que $f'_+(t) = f'_-(t)$. Entonces f es derivable. El hecho de que las ecuaciones (2.14) sean cumplidas por f también, implica la continuidad de f'. $\square$

Definición 2.3.28. *Dada $f \in C^2(-1,1)$, definimos la* segunda diferencia dividida de f *como*

$$f^{[2]}(\lambda_1,\lambda_2,\lambda_3) = \frac{f^{[1]}(\lambda_1,\lambda_2)-f^{[1]}(\lambda_1,\lambda_3)}{\lambda_2-\lambda_3},$$

si λ_1, λ_2 y λ_3 son distintos, y extendiendo por continuidad al resto de los casos.

Lema 2.3.29. *Sea $f(t) = t^m$ con $m = 2,3,...$. Entonces*

$$f^{[2]}(\lambda_1,\lambda_2,\lambda_3) = \sum_{i+j+k=m-2} \lambda_1^i\lambda_2^j\lambda_3^k.$$

Demostración. Basta probarlo para $\lambda_r \neq \lambda_s,\ r \neq s$, pues el resto de los casos se deducen inmediatamente por continuidad.

$$\begin{aligned}
f^{[2]}(\lambda_1,\lambda_2,\lambda_3) &= \frac{f^{[1]}(\lambda_1,\lambda_2)-f^{[1]}(\lambda_1,\lambda_3)}{\lambda_2-\lambda_3} \\
&= \frac{\frac{\lambda_1^m-\lambda_2^m}{\lambda_1-\lambda_2}-\frac{\lambda_1^m-\lambda_3^m}{\lambda_1-\lambda_3}}{\lambda_2-\lambda_3} \\
&= \frac{\sum_{k=0}^{m-1}\lambda_1^k\lambda_2^{m-1-k}-\sum_{k=0}^{m-1}\lambda_1^k\lambda_3^{m-1-k}}{\lambda_2-\lambda_3} \\
&= \frac{\sum_{k=0}^{m-1}\lambda_1^k(\lambda_2^{m-1-k}-\lambda_3^{m-1-k})}{\lambda_2-\lambda_3} \\
&= \sum_{k=0}^{m-2}\lambda_1^k\left(\sum_{j=0}^{m-2-k}\lambda_2^j\lambda_3^{m-2-k-j}\right) \\
&= \sum_{i+j+k=m-2}\lambda_1^i\lambda_2^j\lambda_3^k.
\end{aligned}$$

$\square$

Lema 2.3.30. *Sea $f \in C^2$, $A \in M_n(\mathbb{C})$ diagonal, $A = \sum_{i=1}^n \lambda_i P_i$, con P_i las proyecciones a los ejes coordenados. Entonces para toda $H \in M_n(\mathbb{C})$ se tiene*

$$\frac{d^2}{dt^2}|_{t=0} f(A+tH) = 2\sum_{i,j,k} f^{[2]}(\lambda_i,\lambda_j,\lambda_k) P_i H P_j H P_k. \tag{2.16}$$

Demostración. Primero lo probamos para $f(t) = t^m$, luego probaremos el caso general.

$$\begin{aligned}
\frac{d}{dt}|_{t=t_0} f(A+tH) &= \lim_{s\to 0} \frac{(A+(t_0+s)H)^m - (A+t_0H)^m}{s} \\
&= \lim_{s\to 0} \frac{((A+t_0H)+sH)^m - (A+t_0H)^m}{s} \\
&= \lim_{s\to 0} \frac{[\sum_{k=1}^m (A+t_0H)^{k-1} sH (A+t_0H)^{m-k}] + s^2 Q}{s} \qquad (2.17)\\
&= \sum_{k=1}^m (A+t_0H)^{k-1} H (A+t_0H)^{m-k},
\end{aligned}$$

donde en (2.17) Q es un polinomio. Luego,

$$\begin{aligned}
\frac{d^2}{dt^2}|_{t=0}(A+tH)^m &= \lim_{s\to 0} \frac{\sum_{k=1}^m (A+sH)^{k-1} H (A+sH)^{m-k} - \sum_{k=1}^m A^{k-1} H A^{m-k}}{s} \\
&= \sum_{k=1}^m A^{k-1} H [\sum_{j=1}^{m-k} A^{j-1} H A^{m-k-j}] + [\sum_{j=1}^{k-1} A^{j-1} H A^{k-1-j}] H A^{m-k} \\
&= 2 \sum_{i+j+k=m-2} A^i H A^j H A^k \\
&= 2 \sum_{i+j+k=m-2} \sum_{1\le p,q,r\le n} \lambda_p^i \lambda_q^j \lambda_r^k P_p H P_q H P_r \\
&= 2 \sum_{1\le p,q,r\le n} f^{[2]}(\lambda_p,\lambda_q,\lambda_r) P_p H P_q H P_r,
\end{aligned}$$

donde la última igualdad vale por el Lema 2.3.29.

Para el caso general, sea (f_n) una sucesión de polinomios tal que

$$f_n \to f, f_n' \to f' \text{ y } f_n'' \to f''$$

uniformemente en un intervalo que contenga al espectro de $A+sH$, $-1 \le s \le 1$, y tal que

$$f_n^{[1]}(A) = f^{[1]}(A)\ \forall n. \tag{2.18}$$

Entonces,

$$
\begin{aligned}
& \|\frac{d}{dt^2}|_{t=0} f(A+tH) - 2\sum_{i,j,k} f^{[2]}(\lambda_i,\lambda_j,\lambda_k) P_i H P_j H P_k\| \\
\leq\ & \|\frac{d}{dt^2}|_{t=0}(f-f_n)(A+tH)\| && (2.19)\\
+\ & \|f_n(A+tH) - 2\sum_{i,j,k} f_n^{[2]}(\lambda_i,\lambda_j,\lambda_k) P_i H P_j H P_k\| && (2.20)\\
+\ & \|2\sum_{i,j,k}(f_n-f)^{[2]}(\lambda_i,\lambda_j,\lambda_k) P_i H P_j H P_k\|. && (2.21)
\end{aligned}
$$

El sumando (2.20) es igual a cero, por lo probado en la primera parte, y es fácil ver que el (2.21) tiende a cero. Veamos que (2.19) también tiende a cero. Sean $g(t) = f(A+tH)$ y $g_n(t) = f_n(A+tH)$

$$
\begin{aligned}
\|\frac{d^2}{dt^2}|_{t=0}(f-f_n)(A+tH)\|_2 &= \left\| \lim_{s\to 0} \frac{(g-g_n)'(s) - (g-g_n)'(0)}{s} \right\|_2 \\
&= \left\| \lim_{s\to 0} \frac{D(f-f_n)(A+sH)(H) - D(f-f_n)(A)(H)}{s} \right\|_2 \\
&= \left\| \lim_{s\to 0} \frac{(f-f_n)^{[1]}(A+sH)\circ H - (f-f_n)^{[1]}(A)\circ H}{s} \right\|_2 \\
&= \left\| \lim_{s\to 0} \frac{(f-f_n)^{[1]}(A+sH)\circ H}{s} \right\|_2 \\
&\leq \left\| \lim_{s\to 0} \frac{(f-f_n)^{[1]}(A+sH)}{s} \right\|_2 \|H\|_2
\end{aligned}
$$

Veamos que

$$
\lim_{n\to\infty} \left\| \lim_{s\to 0} \frac{(f-f_n)^{[1]}(A+sH)}{s} \right\|_2 = 0.
$$

Para eso acotaremos cada coordenada de $\frac{(f-f_n)^{[1]}(A+sH)}{s}$. Sean $\lambda_1, ..., \lambda_n$ los autovalores de A, ordenados en forma decreciente. Dados i, j, hay dos situaciones posibles: $\lambda_i = \lambda_j$ o $\lambda_i \neq \lambda_j$.

Si $\lambda_i \neq \lambda_j$: Sean $\lambda_1(s), ..., \lambda_n(s)$ los autovalores de $A + sH$, ordenados en forma decreciente. Por el Corolario 1.2.7, si s es suficientemente chico,

$$
|\lambda_k(s) - \lambda_k| \leq s\|H\| \ \forall k, \tag{2.22}
$$

y además

$$
|\lambda_i(s) - \lambda_j(s)| \geq \frac{\lambda_i - \lambda_j}{2}. \tag{2.23}
$$

Sin pérdida de generalidad, podemos suponer que

$$\|f_n - f\| \leq \frac{1}{n} \text{ y } \|(f_n - f)'\| \leq \frac{1}{n},$$

luego

$$\left|\frac{(f_n - f)(a) - (f_n - f)(b)}{a-b}\right| = |(f_n - f)'(c)| \leq \frac{1}{n}.$$

Resulta entonces que $(f_n - f)$ es Lipschitz con constante $\frac{1}{n}$. Luego,

$$\begin{aligned}
\left|\frac{(f-f_n)(\lambda_i(s)) - (f-f_n)(\lambda_j(s))}{(\lambda_i(s)-\lambda_j(s))s}\right| &\leq \left|\frac{(f-f_n)(\lambda_i(s)) - (f-f_n)(\lambda_i)}{(\lambda_i(s)-\lambda_j(s))s}\right| && (2.24)\\
&+ \left|\frac{(f-f_n)(\lambda_i) - (f-f_n)(\lambda_j)}{(\lambda_i(s)-\lambda_j(s))s}\right| && (2.25)\\
&+ \left|\frac{(f-f_n)(\lambda_j) - (f-f_n)(\lambda_j(s))}{(\lambda_i(s)-\lambda_j(s))s}\right| && (2.26)\\
&\leq \left|\frac{\frac{1}{n}s\|H\|}{\frac{\lambda_i-\lambda_j}{2}s}\right| \xrightarrow{n} 0
\end{aligned}$$

donde en (2.24) y (2.26) acotamos los numeradores por Lipschitz (usando (2.22)) y los denominadores usando (2.23), mientras que (2.25) es cero por (2.18).

Caso $\lambda_i = \lambda_j$:

Si además $\lambda_i(s) \neq \lambda_j(s)$ se tiene

$$\frac{(f-f_n)(\lambda_i(s)) - (f-f_n)(\lambda_j(s))}{(\lambda_i(s)-\lambda_j(s))s} = \frac{(f-f_n)'(\xi(s))}{s} \tag{2.27}$$

con $\xi(s) \in (\lambda_i(s), \lambda_j(s))$. Como $|\lambda_i(s) - \lambda_i| \leq s\|H\|$, $|\lambda_j(s) - \lambda_j| \leq s\|H\|$, y $\lambda_i = \lambda_j$, resulta

$$|\xi(s) - \lambda_i| \leq s\|H\|.$$

Entonces

$$\left|\frac{(f-f_n)'(\xi(s))}{s}\right| \leq \left|\frac{(f-f_n)'(\xi(s))}{\xi(s)-\lambda_i}\right| \|H\| \tag{2.28}$$

(si $\xi(s) = \lambda_i$, resulta $(f-f_n)'(\xi(s)) = 0$ y tenemos la cota que queremos). Como tomamos f_n de modo que $(f-f_n)'$ se anula en los autovalores de A (ver (2.18)) resulta que el último término de la desigualdad (2.28), suponiendo sin pérdida de generalidad $\|H\| = 1$, es igual a

$$\left|\frac{(f-f_n)'(\xi(s)) - (f-f_n)'(\lambda_i)}{\xi(s)-\lambda_i}\right| = |(f-f_n)''(\mu(s))| \xrightarrow{s} |(f-f_n)''(\lambda_i)| \xrightarrow{n} 0,$$

con $\mu(s) \in (\xi(s), \lambda_i)$.

Si $\lambda_i(s) = \lambda_j(s)$, el coeficiente ij de $\frac{(f-f_n)^{[1]}(A+sH)}{s}$ es $\frac{(f-f_n)'(\lambda_i(s))}{s}$, y la situación se resuelve como el caso $\lambda_i(s) \neq \lambda_j(s)$ a partir de (2.27). $\square$

Teorema 2.3.31. *Si $f \in C^2(I)$ es convexa de matrices, entonces para todo $\mu \in I = (-1,1)$, la función $g(\lambda) = f^{[1]}(\mu, \lambda)$ es monótona de matrices.*

Demostración. Como $f \in C^2(I)$, $g \in C^1$. Por el Teorema 2.3.16, basta ver que para todo n, $g^{[1]}(diag(\lambda_1, ..., \lambda_n))$ es positiva para toda $\lambda_1, ..., \lambda_n \in I$. Fijado n, sean $\lambda_1, ..., \lambda_{n+1} \in I$. Sea $A = diag(\lambda_1, ..., \lambda_{n+1})$. Como f es convexa de matrices y dos veces derivable, fijado z la función $h(t) = \langle f(A+tH)z, z\rangle$ es convexa y dos veces derivable, luego $h''(t) \geq 0$ y como $h''(t) = \langle \frac{d^2}{dt^2}|_{t=0} f(A+tH)z, z\rangle$ resulta que para toda H Hermitiana, la matriz $\frac{d^2}{dt^2}|_{t=0} f(A+tH)$ es positiva, y la ecuación (2.16) nos da una manera de expresarla. Sea

$$H = \begin{pmatrix} 0 & 0 & ... & ... & \overline{\xi_1} \\ 0 & 0 & ... & ... & \overline{\xi_2} \\ . & . & ... & ... & . \\ \xi_1 & \xi_2 & ... & \xi_n & 0 \end{pmatrix},$$

con $\xi_1, ..., \xi_n$ números complejos cualesquiera. Sea $x = (1, 1, ..., 1, 0)$ vector de $n+1$ coordenadas. Entonces

$$\langle x, P_i H P_j H P_k x\rangle = \xi_k \overline{\xi_i} \delta_{j,n+1} \text{ para } 1 \leq i, j, k \leq n+1. \tag{2.29}$$

Por la positividad de $\frac{d}{dt}|_{t=0} f(A+tH)$ y las ecuaciones (2.16) y (2.29), se tiene

$$\begin{aligned} 0 &\leq \sum_{1 \leq i,j,k \leq n+1} f^{[2]}(\lambda_i, \lambda_j, \lambda_k) \langle x, P_i H P_j H P_k x\rangle \\ &= \sum_{1 \leq i,k \leq n} f^{[2]}(\lambda_i, \lambda_j, \lambda_k) \xi_k \overline{\xi_i}. \end{aligned}$$

Pero, tomando $\mu = \lambda_{n+1}$ en la definición de g, resulta

$$\begin{aligned} f^{[2]}(\lambda_i, \lambda_{n+1}, \lambda_k) &= \frac{f^{[1]}(\lambda_{n+1}, \lambda_i) - f^{[1]}(\lambda_{n+1}, \lambda_k)}{\lambda_i - \lambda_k} \\ &= g^{[1]}(\lambda_i, \lambda_k) \end{aligned} \tag{2.30}$$

(si $\lambda_i, \lambda_k, \lambda_{n+1}$ son distintos, si no, la igualdad vale por continuidad). Luego,

$$0 \leq \sum_{1 \leq i,k \leq n} g^{[1]}(\lambda_i, \lambda_k) \xi_k \overline{\xi_i} = \langle \xi, g^{[1]}(\lambda)\xi\rangle$$

con $\xi = (\xi_1, ..., \xi_n), \lambda = (\lambda_1, ..., \lambda_n)$. Luego, como esto vale para cualquier ξ, resulta que $g^{[1]}(diag(\lambda))$ es positiva. $\square$

Corolario 2.3.32. *Si $f \in C^2(I)$, $f(0) = 0$, y f es convexa de matrices, entonces la función $g(t) = \frac{f(t)}{t}$ es monótona de matrices.*

Demostración. Por el teorema anterior, la función $f^{[1]}(0,t)$ es monótona de matrices. Pero $f^{[1]}(0,t) = g(t)$. □

Corolario 2.3.33. *Si f es monótona de matrices en I y $f(0) = 0$, entonces la función $g(t) = \frac{t+\lambda}{t} f(t)$ (definida en 0 por continuidad) es monótona de matrices para $|\lambda| \leq 1$.*

Demostración. Supongamos primero que $f \in C^2(I)$. Por el Lema 2.3.26, la función $g_\lambda(t) = (t+\lambda)f(t)$ es convexa de matrices y, por el corolario anterior, resulta $g(t)$ monótona de matrices. Para el caso general, consideremos las regularizaciones f_ϵ; $f_\epsilon(t) - f_\epsilon(0)$ cumple las hipótesis del caso ya probado, luego $\frac{t+\lambda}{t}(f_\epsilon(t) - f_\epsilon(0))$ es monótona de matrices. Como el límite puntual de funciones monótonas de matrices también lo es (Lema 1.2.18) y

$$\lim_{\epsilon \to 0} \frac{t+\lambda}{t}(f_\epsilon(t) - f_\epsilon(0)) = \frac{t+\lambda}{t} f(t)$$

el corolario queda probado. □

Corolario 2.3.34. *Si f es monótona de matrices en $I = (-1,1)$, entonces f es dos veces derivable en 0.*

Demostración. Por el Corolario 2.3.33, la función $g(t) = (1 + \frac{1}{t})f(t)$ es monótona de matrices, y por el Teorema 2.3.27, $f \in C^1(I)$. Entonces la función $h(t) = \frac{1}{t}f(t)$, $h(0) = f'(0)$ también tiene derivada continua, lo que implica que f es dos veces derivable en 0. □

2.4. Caracterización de Loewner.

Notación 2.4.1. Llamaremos K a la colección de todas las funciones monótonas de matrices en el intervalo $(-1,1)$ tales que $f(0) = 0$ y $f'(0) = 1$. El conjunto K es claramente convexo.

Lema 2.4.2. *Si $f \in K$, entonces*

- $f(t) \leq \frac{t}{1-t}$ *si* $0 \leq t < 1$,
- $f(t) \geq \frac{t}{1+t}$ *si* $-1 < t < 0$,

- $|f''(0)| \leq 2$

Demostración. Sea $A = \begin{pmatrix} t & 0 \\ 0 & 0 \end{pmatrix}$. Por el Teorema 2.3.16, la matriz

$$f^{[1]}(A) = \begin{pmatrix} f'(t) & \frac{f(t)}{t} \\ \frac{f(t)}{t} & 1 \end{pmatrix}$$

es positiva. Luego,

$$\frac{f(t)^2}{t^2} \leq f'(t). \tag{2.31}$$

Sean $g_{\pm}(t) = (t \pm 1)f(t)$. Por el Lema 2.3.26, las funciones g_+ y g_- son convexas de matrices, luego, sus respectivas derivadas son funciones crecientes.
Como $g'_{\pm}(t) = f(t) + (t \pm 1)f'(t)$ y $g'_{\pm}(0) = \pm 1$, resulta

$$f(t) + (t-1)f'(t) \geq -1 \text{ si } t > 0, \tag{2.32}$$

$$f(t) + (t+1)f'(t) \leq 1 \text{ si } t > 0. \tag{2.33}$$

De las desigualdades (2.31) y (2.32) obtenemos, para $t > 0$

$$f(t) + 1 \geq \frac{(1-t)f(t)^2}{t^2}. \tag{2.34}$$

Supongamos que para algún $t \in (0,1)$ se tiene $f(t) > \frac{t}{1-t}$, por lo tanto $f(t)^2 > \frac{t}{1-t}f(t)$. Entonces, por (2.34), $f(t) + 1 > \frac{f(t)}{t}$, luego $f(t) < \frac{t}{1-t}$, que contradice nuestra suposición. Luego $f(t) \leq \frac{t}{1-t}$ para $0 \leq t < 1$.

La desigualdad $f(t) \geq \frac{t}{1+t}$ se prueba de manera análoga usando (2.33) en lugar de (2.32).

Probemos la última desigualdad del lema.

$$\begin{aligned}
\lim_{t\to 0} \frac{(1+t^{-1})f(t) - f'(0)}{t} &= \lim_{t\to 0} \frac{(t+1)f(t) - tf'(0)}{t^2} \\
&= \lim_{t\to 0} \frac{f(t) + (t+1)f'(t) - f'(0)}{2t} \\
&= \lim_{t\to 0} \frac{f(t) - f(0)}{t}\frac{1}{2} + \frac{f'(t) - f'(0)}{t}\frac{1}{2} + \frac{f'(t)}{2} \\
&= f'(0) + \frac{1}{2}f''(0).
\end{aligned}$$

Si $t \to 0^+$, por la primera desigualdad del lema, obtenemos

$$\lim_{t\to 0^+} \frac{(1+t^{-1})f(t) - f'(0)}{t} \quad \leq \quad \lim_{t\to 0^+} \frac{(1+t^{-1})\frac{t}{1-t} - f'(0)}{t} = 2,$$

si $t \to 0^-$, por la segunda desigualdad del lema, obtenemos

$$\lim_{t\to 0^-} \frac{(1+t^{-1})f(t) - f'(0)}{t} \quad \geq \quad \lim_{t\to 0^-} \frac{(1+t^{-1})\frac{t}{1+t} - f'(0)}{t} = 0,$$

por lo tanto

$$0 \leq f'(0) + \frac{1}{2}f''(0) \leq 2,$$

y como $f'(0) = 1$, resulta $|f''(0)| \leq 2$. □

Proposición 2.4.3. *El conjunto K es compacto en la topología de la convergencia puntual.*

Demostración. Sea $\{f_i\}$ una red en K. El conjunto $\{f_i(t)\}$ es acotado para cada t (por las cotas impuestas por el lema anterior y la monotonía de las funciones f_i). Luego, por el teorema de Tychonoff, existe una subred de $\{f_i\}$ que converge puntualmente a una función acotada f. Veamos que $f \in K$. Claramente $f(0) = 0$, y ademas f es monótona de matrices (por la Proposición 1.2.19). Por el Corolario 2.3.33, las funciones $(1 + \frac{1}{t})f_i(t)$ son monótonas de matrices en $(-1, 1)$. Como para todo i, $\lim_{t\to 0}(1 + \frac{1}{t})f_i(t) = f_i'(0) = 1$ (usando L´Hospital), resulta $(1 + \frac{1}{t})f_i(t) \geq 1$ si $t \geq 0$, y $(1 + \frac{1}{t})f_i(t) \leq 1$ si $t \leq 0$ (por la monotonía de las funciones $(1 + \frac{1}{t})f_i(t)$). Luego, $(1 + \frac{1}{t})f(t) \geq 1$ si $t > 0$, y $(1 + \frac{1}{t})f(t) \leq 1$ si $t < 0$. Como f es continuamente diferenciable (ver Teorema 2.3.27), resulta $f'(0) = 1$ (de nuevo, usando L´Hospital). □

Proposición 2.4.4. *Todos los puntos extremales de K son de la forma*

$$f(t) = \frac{t}{1-\alpha t}, \quad \text{con } \alpha = \frac{1}{2}f''(0).$$

Demostración. Sea $f \in K$. Para cada $\lambda \in (-1, 1)$, sea $g_\lambda(t) = (1 + \frac{\lambda}{t})f(t) - \lambda$. Por el Corolario 2.3.33, g_λ es monótona de matrices. Definiendo $g_\lambda(0) = 0$, resulta que g_λ es continua en $(-1, 1)$ (se puede probar fácilmente usando L'Hospital, que $f(0) = 0$ y que $f'(0) = 1$). Además, $g_\lambda'(0) = 1 + \frac{1}{2}\lambda f''(0)$ (también es fácil de probar con L'Hospital). Entonces la función

$$h_\lambda(t) = \frac{1}{1 + \frac{1}{2}\lambda f''(0)}[(1 + \frac{\lambda}{t})f(t) - \lambda] \tag{2.35}$$

está en K (extendiendo la definición a $t = 0$ por continuidad). Como $|f''(0)| \leq 2$, resulta $|\frac{1}{2}\lambda f''(0)| < 1$. Podemos escribir a f como

$$f = \frac{1}{2}(1 + \frac{1}{2}\lambda f''(0))h_\lambda + \frac{1}{2}(1 - \frac{1}{2}\lambda f''(0))h_{-\lambda}.$$

Luego, si f es extremal en K, debe ser $f = h_\lambda$. Luego, remplazando y despejando en (2.35), se tiene

$$(1 + \frac{1}{2}\lambda f''(0))f(t) = (1 + \frac{\lambda}{t})f(t) - \lambda$$

y despejando obtenemos

$$f(t) = \frac{t}{1 - \frac{1}{2}f''(0)t}.$$

□

Teorema 2.4.5. *Para toda $f \in K$ existe una única medida de probabilidad μ en $[-1, 1]$ tal que*

$$f(t) = \int_{-1}^{1} \frac{t}{1 - \lambda t} d\mu(\lambda). \tag{2.36}$$

Demostración. Para $-1 \leq \lambda \leq 1$, definimos $h_\lambda(t) = \frac{t}{1-\lambda t}$. Por la Proposición 2.4.4 , los puntos extremales de K están incluidos en $\{h_\lambda\}$. Como K es convexo y compacto, resulta que K es igual a la cápsula convexa cerrada de sus puntos extremales (por el teorema de Krein-Milman). Las combinaciones convexas finitas de elementos de $\{h_\lambda\}$ se pueden escribir como $\int h_\lambda d\nu(\lambda)$, con ν una medida de probabilidad en $[-1, 1]$ con soporte finito. Como f está en la clausura de estas combinaciones, existe una red ν_i de medidas de probabilidad con soporte finito en $[-1, 1]$ tal que $f_i(t) = \int h_\lambda(t)d\nu_i(\lambda)$ converge a $f(t)$. Como el espacio de medidas de probabilidad esta contenido en el dual de $L^\infty([-1, 1])$ (vía $\int_{-1}^{1} f d\mu$), es claramente cerrado con la topología ω^*, y $\|\mu\| = 1$ para toda medida de probabilidad (vista como elemento del dual de L^∞), resulta, por el teorema de Banach-Alaoglu, que el espacio de medidas de probabilidad es ω^*-compacto. Luego, existe una subred de $\int h_\lambda d\nu_i(\lambda)$ que converge a $\int h_\lambda d\mu(\lambda)$. Entonces $f(t) = \int h_\lambda(t)d\mu(\lambda) = \int \frac{t}{1-\lambda t} d\mu(\lambda)$.

Veamos que la medida es única. Supongamos que existan dos medidas μ_1, μ_2 para los cuales vale la representación dada por (2.36). Usando que $\frac{t}{1-\lambda t} = \sum_{n=0}^{\infty} t^{n+1}\lambda^n$, donde la serie converge uniformemente si $-1 \leq \lambda \leq 1$ y $|t| < 1$ fijo, obtenemos que

$$\sum_{n=0}^{\infty} t^{n+1} \int_{-1}^{1} \lambda^n d\mu_1(\lambda) = \sum_{n=0}^{\infty} t^{n+1} \int_{-1}^{1} \lambda^n d\mu_2(\lambda)$$

para todo $|t| < 1$. Entonces resulta

$$\int_{-1}^{1} \lambda^n d\mu_1(\lambda) = \int_{-1}^{1} \lambda^n d\mu_2(\lambda)$$

para $n \geq 0$, y esto implica que $\mu_1 = \mu_2$ (por la densidad de los polinomios en las funciones continuas y el hecho de que μ_1 y μ_2 son medidas finitas).

$\square$

Corolario 2.4.6. *Sea f una función monótona de matrices no constante en el intervalo $(-1,1)$. Entonces existe una única medida de probabilidad μ en $[-1,1]$ tal que*

$$f(t) = f(0) + f'(0) \int_{-1}^{1} \frac{t}{1-\lambda t} d\mu(\lambda).$$

Demostración. Como f es monótona y no constante, por la ecuación (2.31) resulta $f'(0) \neq 0$. Como $\frac{f(t)-f(0)}{f'(0)} \in K$, del teorema anterior se deduce el corolario. $\square$

Corolario 2.4.7. *Sea f una función no lineal, convexa de matrices en $(-1,1)$. Entonces existe una única medida de probabilidad μ en $[-1,1]$ tal que*

$$f(t) = f(0) + f'(0)t + \int_{-1}^{1} \frac{t^2}{1-\lambda t} d\mu(\lambda). \tag{2.37}$$

Demostración. Supongamos, sin perdida de generalidad, que $f(0) = 0$ y $f'(0) = 0$. Sea $g(t) = \frac{f(t)}{t}$. Entonces, por el Corolario 2.3.32, g es monótona de matrices, $g(0) = 0$ y $g'(0) = \frac{1}{2}f''(0)$ (usando L´Hospital). Entonces g tiene una representación según la fórmula (2.36), de lo que se deduce inmediatamente que f tiene una representación de la forma (2.37). $\square$

Observación 2.4.8. De la representación (2.36) se deduce que toda función monótona de matrices en $(-1,1)$ es infinitamente diferenciable. Análogamente, de (2.37) se sigue que toda función convexa de matrices en $(-1,1)$ es infinitamente diferenciable.

Observación 2.4.9. Notemos que una función f es monótona (convexa) de matrices en (a,b) si y sólo si la función $f(\frac{(b-a)t}{2} + \frac{b+a}{2})$ es monótona (convexa) de matrices en $(-1,1)$.

Capítulo 3

C^*-álgebras.

En este capítulo estudiaremos a las C^*-álgebras. Veremos que toda C^*-álgebra se representa como una subálgebra cerrada de $\mathcal{B}(\mathcal{H})$, con $\mathcal{H}$ un espacio de Hilbert, y gracias a la transformada de Gelfand, podremos definir funciones continuas sobre los elementos normales, lo cual nos permitirá, en el último capítulo, extender algunos resultados, válidos para matrices, a álgebras de operadores.

3.1. Definiciones y propiedades básicas.

Definición 3.1.1. *Un álgebra de Banach es un álgebra compleja normada completa $\mathcal{U}$ que satisface*

$$\|AB\| \leq \|A\| \, \|B\| \ \ \forall A, B \in \mathcal{U}.$$

Definición 3.1.2. *Una *-álgebra de Banach es un álgebra de Banach $\mathcal{U}$ dotada de una operación (llamada* involución*) $* : \mathcal{U} \to \mathcal{U}$ que es un anti-isomorfismo, i.e.*

$$(A+B)^* = A^* + B^*$$

$$(\lambda A)^* = \overline{\lambda} A^*$$

$$A^{**} = A$$

$$(AB)^* = B^* A^*$$

para todo $A, B \in \mathcal{U}$ y $\lambda \in \mathbb{C}$.

Definición 3.1.3. *Una C^*-álgebra es una *-álgebra de Banach tal que $\forall A \in \mathcal{U}$*

$$\|A^* A\| = \|A\|^2 .$$

Proposición 3.1.4. *Si A es un elemento de una C^*-álgebra, entonces $\|A\| = \|A^*\|$.*

Demostración. $\|A\|^2 = \|A^*A\| \leq \|A^*\|\|A\|$, luego $\|A\| \leq \|A^*\|$. Cambiando A por A^* se obtiene la desigualdad $\|A^*\| \leq \|A\|$. □

Proposición 3.1.5. *Si A es un elemento de una C^*-álgebra, entonces $\|A^*A\| = \|AA^*\|$.*

Demostración. $\|A^*A\| = \|A\|^2 = \|A^*\|^2 = \|A^{**}A^*\| = \|AA^*\|$. □

Ejemplo 3.1.6. El álgebra de todos los operadores acotados en un espacio de Hilbert $\mathcal{H}$, notado $\mathcal{B}(H)$, es una C^*-álgebra con la adjunción usual como involución.

Ejemplo 3.1.7. Dado un operador T en un espacio de Hilbert $\mathcal{H}$, definimos $C^*(T)$, como la C^*-subálgebra de $\mathcal{B}(\mathcal{H})$ generada por T, i.e. la clausura uniforme de todos los polinomios en T, T^* e I, con I el operador identidad.

Ejemplo 3.1.8. Sea X un espacio localmente compacto, Hausdorff. Sea $C_0(X)$ el conjunto de todas las funciones continuas de X en $\mathbb{C}$ que tienden a cero en el infinito (i.e. $\forall \epsilon > 0\ \exists K$ compacto tal que $|f(x)| < \epsilon\ \forall x \notin K$). Este conjunto es una C^*-álgebra con la conjugación compleja por involución.

Ejemplo 3.1.9. El conjunto $M_n(\mathbb{C})$ con las operaciones habituales, la norma de operadores y * la adjunción habitual, es una C^*-álgebra.

Proposición 3.1.10. *Toda C^*-álgebra sin unidad $\mathcal{U}$ está contenida en una C^*-álgebra con unidad $\overline{\mathcal{U}}$, de manera tal que $\mathcal{U}$ es un ideal maximal (de codimensión 1) en $\overline{\mathcal{U}}$.*

Demostración. Sea $\overline{\mathcal{U}} = \mathcal{U} \oplus \mathbb{C}$. Definimos las siguientes operaciones en $\mathcal{U}$:

$$(A, \lambda)(B, \mu) = (AB + \lambda B + \mu A, \lambda\mu)$$

$$(A, \lambda)^* = (A^*, \overline{\lambda})$$

$$\|(A, \lambda)\| = \sup_{\|B\| \leq 1} \|AB + \lambda B\|$$

Estas operaciones hacen de $\overline{\mathcal{U}}$ una *-álgebra. Esta norma es una norma de álgebra de Banach, pues es la norma inducida por $\mathcal{B}(\mathcal{U})$ de los operadores acotados $\{L_A + \lambda I\}$ donde $L_A(B) = AB$. La unidad es claramente $(0, 1)$, y $\mathcal{U}$ es maximal (de codimensión 1) por construcción. La inclusión de $\mathcal{U}$ en $\overline{\mathcal{U}}$ es isométrica, pues

$$\|A\| = \left\| A \frac{A^*}{\|A\|} \right\| \leq \|(A, 0)\| = \sup_{\|B\| \leq 1} \|AB\| \leq \|A\|\,.$$

Por último, veamos que $\|(A,\lambda)^*(A,\lambda)\| = \|(A,\lambda)\|^2$:

$$\begin{aligned}
\|(A,\lambda)\|^2 &= \sup_{\|B\|\le 1} \|AB+\lambda B\|^2 = \sup_{\|B\|\le 1} \|(AB+\lambda B)^*(AB+\lambda B)\| \\
&= \sup_{\|B\|\le 1} \left\|B^*A^*AB + \lambda B^*A^*B + \overline{\lambda}B^*AB + |\lambda|^2 B^*B\right\| \\
&\le \sup_{\|B\|\le 1} \left\|A^*AB + \lambda A^*B + \overline{\lambda}AB + |\lambda|^2 B\right\| \\
&= \left\|(A^*A + \lambda A^* + \overline{\lambda}A, |\lambda|^2)\right\| \\
&= \|(A,\lambda)^*(A,\lambda)\| \le \|(A,\lambda)^*\| \, \|(A,\lambda)\| . \qquad (3.1)
\end{aligned}$$

Luego $\|(A,\lambda)\| \le \|(A,\lambda)^*\|$ y, por simetría, vale la igualdad, entonces de (3.1) resulta $\|(A,\lambda)^*(A,\lambda)\| = \|(A,\lambda)\|^2$. □

Definición 3.1.11. *El* espectro *de un elemento A de un álgebra de Banach con unidad $\mathcal{U}$ es el conjunto*

$$\sigma(A) = \{\lambda \in \mathbb{C} : \lambda I - A \textit{ no es inversible } \} .$$

El complemento del espectro es el resolvente*, donde está bien definida*

$$R_A(\lambda) = (\lambda I - A)^{-1},$$

la función resolvente.

Teorema 3.1.12. *Sea $A \in \mathcal{U}$, con $\mathcal{U}$ un álgebra de Banach con unidad, entonces el espectro de A es compacto y no vacío. Además la función resolvente es analítica en $\mathbb{C} - \sigma(A)$.*

Demostración. Si $|\lambda| > \|A\|$, entonces $\|\lambda^{-n}A^n\| \le (|\lambda|^{-1}\|A\|)^n$, que decrece geométricamente, luego la serie

$$\sum_{n\ge 0} \lambda^{-n-1}A^n$$

converge en norma. El límite es $(\lambda I - A)^{-1}$, pues

$$(\lambda I - A)\sum_{n=0}^{k} \lambda^{-n-1}A^n = I - \lambda^{-k-2}A^{k+1},$$

que converge claramente a I. Más aún, esto prueba que $R_A(\lambda)$ es analítica, y tiene desarrollo en serie de Laurent en el infinito. Además

$$\lim_{|\lambda|\to\infty} \|R_A(\lambda)\| \le \lim_{|\lambda|\to\infty} |\lambda|^{-1}(1-|\lambda|^{-1}\|A\|)^{-1} = 0.$$

Del mismo modo, si $\lambda_0 I - A$ es inversible y $|\lambda - \lambda_0| < \|(\lambda_0 I - A)^{-1}\|^{-1}$, entonces

$$(\lambda I - A)^{-1} = \sum_{n\geq 0} (\lambda - \lambda_0)^n (\lambda_0 I - A)^{-n-1}$$

es la serie de Taylor de R_A centrada en λ_0. Así, la función resolvente resulta ser analítica en el complemento del espectro. En particular, $f(R_A(\lambda))$ es una función escalar analítica para toda funcional lineal continua f definida sobre $\mathcal{U}$. El resolvente es entonces un conjunto abierto que contiene a todos los complejos de módulo mayor que $\|A\|$, y por lo tanto $\sigma(A)$ es compacto. Si el espectro de A fuera vacío, entonces $R_A(\lambda)$ sería una función entera y acotada. Por el teorema de Liouville resultaría el absurdo $R_A \equiv 0$. En efecto, para toda funcional f, $f(R_A(\lambda))$ es una función escalar entera que tiende a 0 cuando $|\lambda| \to \infty$, por lo tanto es la función constante 0. Por el teorema de Hahn-Banach, lo anterior implica que $R_A = 0$. □

Definición 3.1.13. *Análogamente a las definiciones para operadores lineales, tenemos las siguientes definiciones para elementos de C^*-algebras. Sea $\mathcal{U}$ una C^*-álgebra, decimos que $A \in \mathcal{U}$ es*

- autoadjunto *o* Hermitiano *si $A^* = A$.*
- normal *si $AA^* = A^*A$.*
- unitario *si $A^*A = I = AA^*$.*
- positivo *si es autoadjunto y $\sigma(A) \subseteq [0, \infty)$.*

Lema 3.1.14. *Si p es un polinomio y A es un elemento de un álgebra de Banach con unidad $\mathcal{U}$, entonces*

$$\sigma(p(A)) = p(\sigma(A)).$$

Demostración. Dado $\alpha \in \mathbb{C}$, $p - \alpha$ se factoriza

$$p(z) - \alpha = \prod_{i=1}^{n} (z - \beta_i).$$

Entonces se sigue que

$$p(A) - \alpha I = c \prod_{i=1}^{n} (A - \beta_i I).$$

Como todos los factores conmutan, $p(A) - \alpha I$ es inversible si y sólo si $A - \beta_i I$ es inversible para todo i. Esto implica que $\alpha \in \sigma(p(A))$ si y sólo si $\exists\, \beta_i \in \sigma(A)$ tal que $p(\beta_i) = \alpha$. □

Definición 3.1.15. *El* radio espectral *de A en un álgebra de Banach con unidad es*

$$spr(A) = sup\{|\lambda| : \lambda \in \sigma(A)\}.$$

Proposición 3.1.16. *Dada un álgebra de Banach $\mathcal{U}$, se tiene*

$$spr(A) = \lim_{n\to\infty} \|A^n\|^{\frac{1}{n}}.$$

Demostración. Miremos la serie de Laurent en el infinito de la función resolvente,

$$(\lambda I - A)^{-1} = \sum_{n\geq 0} \lambda^{-n-1} A^n.$$

Como R_A es analítica en $\{\lambda : |\lambda| > spr(A)\}$, la serie converge absoluta y uniformemente en $|\lambda| \geq r > spr(A)$. En particular, los coeficientes de la serie de Taylor $r^{-n-1}\|A^n\|$ convergen a 0 para $r > spr(A)$, lo que implica

$$\limsup_{n\to\infty} \|A^n\|^{\frac{1}{n}} \leq spr(A).$$

Por otro lado, $\exists \alpha \in \sigma(A)$ tal que $|\alpha| = spr(A)$ (pues el espectro de A es compacto). Por el Lema 3.1.14, $\alpha^n \in \sigma(A^n)$, entonces

$$spr(A) = |\alpha^n|^{\frac{1}{n}} \leq \|A^n\|^{\frac{1}{n}},\ \forall n \geq 1.$$

Por lo tanto,

$$\limsup_{n\to\infty} \|A^n\|^{\frac{1}{n}} \leq spr(A) \leq \liminf_{n\to\infty} \|A^n\|^{\frac{1}{n}}.$$

□

Lema 3.1.17. *Sea $\mathcal{U}$ un álgebra de Banach con unidad, y $A \in \mathcal{U}$. Si $\|A\| < 1$, entonces $I - A$ es inversible y su inversa es $\sum_{n=0}^{\infty} A^n$.*

Demostración. $\sum_{n=0}^{\infty} A^n$ es un elemento de $\mathcal{U}$ pues la serie $\sum_{n=0}^{\infty} \|A\|^n$ es convergente y un simple cálculo prueba que $A\sum_{n=0}^{\infty} A^n = \sum_{n=0}^{\infty} A^n A = I$. □

Proposición 3.1.18. *La única álgebra de Banach con unidad, simple (i.e. sin ideales propios no triviales) y conmutativa es $\mathbb{C}$.*

Demostración. Sea $\mathcal{U}$ un álgebra de Banach con unidad conmutativa que contiene un elemento A, que no es escalar. Sea $\alpha \in \sigma(A)$, y sea $\mathcal{J} = \overline{(A - \alpha I)\mathcal{U}}$. $\mathcal{J}$ es un ideal cerrado, y ningún elemento de la forma $(A - \alpha I)B$ es inversible (si no $(A - \alpha I)$ resultaría inversible), por lo tanto $\|(A - \alpha I)B - I\| \geq 1$ (por el Lema 3.1.17). Entonces I no pertenece a $\mathcal{J}$, por lo que $\mathcal{J}$ es un ideal propio. Así, cuando $\mathcal{U}$ es simple, todo elemento de $\mathcal{U}$ debe ser escalar, y por lo tanto $\mathcal{U} = \mathbb{C}$. □

3.2. Funcionales lineales multiplicativas.

Definición 3.2.1. *Una* funcional lineal multiplicativa *en un álgebra de Banach $\mathcal{U}$ es un homomorfismo no nulo de $\mathcal{U}$ en $\mathbb{C}$. El conjunto $\mathcal{M}_{\mathcal{U}}$ de todas las funcionales lineales multiplicativas en $\mathcal{U}$ es el* espacio ideal maximal *de $\mathcal{U}$.*

Proposición 3.2.2. *Sea $\mathcal{U}$ un álgebra de Banach abeliana con unidad, $\phi \in \mathcal{M}_{\mathcal{U}}$, entonces ϕ es continua de norma 1. La aplicación que manda cada funcional a su núcleo es una biyección entre el conjunto de funcionales lineales multiplicativas y el conjunto de ideales maximales de $\mathcal{U}$.*

Demostración. Sea $\phi \in \mathcal{M}_{\mathcal{U}}$. Supongamos $\|\phi\| > 1$, luego existe $A \in \mathcal{U}$ tal que $\|A\| < 1 = \phi(A)$. Sea $B = \sum_{n=1}^{\infty} A^n$. Entonces $A + AB = B$ y

$$\phi(B) = \phi(A + AB) = \phi(A) + \phi(A)\phi(B) = 1 + \phi(B),$$

lo cual es absurdo. Luego $\|\phi\| \leq 1$, pero como $\phi(I) = 1$, vale la igualdad.

De lo anterior se deduce que $M = ker(\phi)$ es un ideal cerrado de codimensión 1 en $\mathcal{U}$, por lo tanto es maximal, y como ϕ queda definido por M y por el hecho de que $\phi(I) = 1$, la correspondencia es inyectiva.

Sea M un ideal maximal de $\mathcal{U}$, luego $dist(M, I) = 1$ pues la bola de radio 1 y centro I está formada por elementos inversibles. Luego la clausura de M tampoco contiene a I y, como M es maximal, resulta que M es cerrado. Así $\mathcal{U}/M$ es un álgebra de Banach con unidad, simple y abeliana, y luego es isomorfo a $\mathbb{C}$ (ver Proposición 3.2.2). Luego la aplicación cociente es un homomorfismo continuo entre $\mathcal{U}$ y $\mathbb{C}$ con núcleo M. □

Corolario 3.2.3. *Sea $\mathcal{U}$ un álgebra de Banach abeliana sin unidad y $\overline{\mathcal{U}}$ la construida en la Proposición 3.1.10. Entonces hay un homeomorfismo entre $\mathcal{M}_{\mathcal{U}}$ y $\mathcal{M}_{\overline{\mathcal{U}}} - \{\phi_0\}$, con ϕ_0 la funcional lineal que se anula en $\mathcal{U}$.*

Demostración. $\mathcal{U}$ es un ideal maximal de $\overline{\mathcal{U}}$, que se corresponde con ϕ_0. Luego cualquier otro ideal maximal de $\overline{\mathcal{U}}$ tiene intersección no nula con $\mathcal{U}$, y esta intersección es un ideal de codimensión 1 en $\mathcal{U}$, que se corresponde con un elemento de $\mathcal{M}_{\mathcal{U}}$.

Por otro lado, si ϕ es un elemento de $\mathcal{M}_{\mathcal{U}}$, $\tilde{\phi}(A, \lambda) = \phi(A) + \lambda$ es la única extensión posible a $\overline{\mathcal{U}}$, por lo que la biyección queda determinada.

Veamos que esta correspondencia resulta homeomorfa, considerando en ambos espacios la topología débil-*: por el teorema de Banach-Alaoglu, $\mathcal{M}_{\mathcal{U}}$ es compacto y Hausdorff (pues es un cerrado incluido en la bola unitaria), y además es claro, por la construcción, que si $\phi_\alpha \to \phi$ entonces $\tilde{\phi_\alpha} \to \tilde{\phi}$. Como las funciones biyectivas, continuas, con dominio compacto son homeomorfismos, queda probado el corolario. □

Observación 3.2.4. Del corolario anterior se deduce que $\mathcal{M}_{\overline{\mathcal{U}}}$ es la compactificación de un punto de $\mathcal{M}_\mathcal{U}$.

3.3. Transformada de Gelfand.

Definición 3.3.1. *Definimos la* transformada de Gelfand Γ *de un álgebra de Banach conmutativa* $\mathcal{U}$ *en* $C_0(\mathcal{M}_\mathcal{U})$ *como* $\Gamma(A) = \widehat{A}$ *donde* $\widehat{A}(\phi) = \phi(A)$ *para todo* $A \in \mathcal{U}$.

Teorema 3.3.2. *La transformada de Gelfand es un morfismo de álgebras contractivo, y* $\Gamma(\mathcal{U})$ *separa puntos en* $\mathcal{M}_\mathcal{U}$.

Demostración. Las funciones $\widehat{A}$ son continuas pues si $\phi_\alpha \to \phi$ (en la topología débil *, que es la que consideramos siempre en $\mathcal{M}_\mathcal{U}$), entonces

$$\widehat{A}(\phi_\alpha) = \phi_\alpha(A) \to \phi(A) = \widehat{A}(\phi).$$

Si $\mathcal{U}$ tiene unidad, $\mathcal{M}_\mathcal{U}$ es compacto, por lo que en este caso $C(\mathcal{M}_\mathcal{U}) = C_0(\mathcal{M}_\mathcal{U})$. Si $\mathcal{U}$ no tiene unidad, $\mathcal{M}_\mathcal{U} = \mathcal{M}_{\overline{\mathcal{U}}} - \{\phi_0\}$, y como $\widehat{A}(\phi_0) = \phi_0(A) = 0$, resulta que $\widehat{A} \in C_0(\mathcal{M}_\mathcal{U})$.

La transformada de Gelfand es contractiva, pues

$$\|\Gamma\| = \sup_{\|A\|\leq 1} \|\Gamma(A)\| = \sup_{\|A\|\leq 1} \sup_{\|\phi\|\leq 1} \|\Gamma(A)(\phi)\| = \sup_{\|A\|\leq 1} \|\phi(A)\| = \|\phi\| \leq 1.$$

Es claro que $\Gamma(\mathcal{U})$ separa puntos, dados dos funcionales distintas en $\mathcal{M}_\mathcal{U}$, ϕ_1, ϕ_2, existe A un elemento de $\mathcal{U}$ tal que $\phi_1(A) \neq \phi_2(A)$, luego $\widehat{A}(\phi_1) \neq \widehat{A}(\phi_2)$. □

Corolario 3.3.3. *En un álgebra de Banach conmutativa con unidad* $\mathcal{U}$*,* A *es inversible si y sólo si* $\widehat{A}$ *lo es, y esto sucede cuando* $\widehat{A}$ *no se anula en* $\mathcal{M}_\mathcal{U}$*. Además*

$$\sigma(A) = \sigma(\widehat{A}) = \{\phi(A) : \ \phi \in \mathcal{M}_\mathcal{U}\}$$

y $\|\widehat{A}\|_\infty = spr(A)$.

Demostración. Si A es inversible, como Γ es morfismo, $\Gamma(A^{-1}) = \Gamma(A)^{-1}$. Si A no es inversible, el ideal $\mathcal{J} = \overline{A\mathcal{U}}$ es un ideal propio, y por lo tanto está contenido en un ideal maximal M. Sea ϕ la funcional multiplicativa asociada a M. Así $\widehat{A}(\phi) = 0$ y por lo tanto $\widehat{A}$ no es inversible. De esto se deduce inmediatamente que $\sigma(A) = \sigma(\widehat{A})$, y además es igual al rango de $\widehat{A}$. Tomando supremo en estos conjuntos, como $\|\phi\| \leq 1$ $\forall \phi \in \mathcal{M}_\mathcal{U}$, resulta $spr(A) = spr(\widehat{A}) = \|\widehat{A}\|_\infty$. □

Teorema 3.3.4. *Sea $\mathcal{U}$ una C^*-álgebra abeliana. Entonces la transformada de Gelfand es un *-isomorfismo isométrico de $\mathcal{U}$ en $C_0(\mathcal{M}_{\mathcal{U}})$.*

Demostración. Supongamos primero que $\mathcal{U}$ tiene unidad. Sea ϕ una funcional lineal multiplicativa en $\mathcal{U}$. Veamos que $\phi(A^*) = \overline{\phi(A)}$. Supongamos que A es autoadjunto. Definimos, para $t \in \mathbb{R}$, los operadores

$$U_t = e^{itA} = \sum_{n=0}^{\infty} \frac{(itA)^n}{n!}.$$

Luego

$$U_t^* = \sum_{n=0}^{\infty} \frac{(\overline{it}A)^n}{n!} = \sum_{n=0}^{\infty} \frac{(-itA)^n}{n!} = e^{-itA} = U_t^{-1},$$

por lo que los operadores U_t son unitarios. Así, $\|U_t\|^2 = \|U_t^* U_t\| = \|I\| = 1$.

Entonces,

$$1 \geq |\phi(U_t)| = \left| \sum_{n=0}^{\infty} \frac{(it\phi(A))^n}{n!} \right| = |e^{it\phi(A)}| = e^{-tIm(\phi(A))}.$$

Como esta desigualdad vale para todo t, resulta que $\phi(A)$ es real.

Supongamos ahora que X es un elemento cualquiera de $\mathcal{U}$, $X = A + iB$, con $A = \frac{X+X^*}{2}$, $B = \frac{X-X^*}{2i}$ (A, y B son respectivamente la parte real e imaginaria de X y son claramente autoadjuntos). Se tiene

$$\phi(X^*) = \phi(A - iB) = \phi(A) - i\phi(B) = \overline{\phi(A) + i\phi(B)} = \overline{\phi(X)}.$$

Así, Γ verifica $\widehat{A^*} = \widehat{A}^*$, por lo que la transformada de Gelfand es un *-homomorfismo.

Si A es autoadjunto, $\|A\|^2 = \|A^*A\| = \|A^2\|$, luego, por la Proposición 3.1.16, tenemos

$$\left\|\widehat{A}\right\|_\infty = spr(A) = \lim_{n\to\infty} (\left\|A^{2^n}\right\|)^{\frac{1}{2^n}} = \|A\|.$$

En el caso general, si X no es necesariamente autoadjunto, tenemos

$$\|X\|^2 = \|X^*X\| = \left\|\widehat{X^*X}\right\|_\infty = \left\|\widehat{X}^*\widehat{X}\right\|_\infty = \left\|\widehat{X}\right\|_\infty^2,$$

resulta entonces que la transformada de Gelfand es isométrica.

El conjunto $\Gamma(\mathcal{U})$ separa puntos en $\mathcal{M}_{\mathcal{U}}$ claramente, pues dadas dos funcionales distintas ϕ_1, ϕ_2, tomando un elemento de $\mathcal{U}$ tal que $\phi_1(A) \neq \phi_2(A)$, resulta $\widehat{A}(\phi_1) \neq \widehat{A}(\phi_2)$.

Resumiendo, la imagen de $\mathcal{U}$ por la transformada de Gelfand es una subálgebra de $C(\mathcal{M}_{\mathcal{U}})$, cerrada en norma, autoadjunta, con unidad, y que separa puntos. Por el teorema de Stone-Weierstrass, Γ es suryectiva, y por lo tanto es un *-isomorfismo.

Veamos ahora el caso en que $\mathcal{U}$ no tiene unidad. Por la Proposición 3.1.10, $\mathcal{U}$ está contenida en una C^*-álgebra con unidad $\overline{\mathcal{U}}$, de la cual es un ideal maximal de codimensión 1, $\overline{\mathcal{U}}$ también resulta abeliana. Más aún, por el Corolario 3.2.3, $\mathcal{M}_{\mathcal{U}}$ es canónicamente homeomorfo a $\mathcal{M}_{\overline{\mathcal{U}}} - \{\phi_0\}$, con ϕ_0 la funcional lineal multiplicativa de $\mathcal{M}_{\overline{\mathcal{U}}}$ cuyo núcleo es $\mathcal{U}$, $\phi_0((A,\lambda)) = \lambda$. Por lo visto en el caso con unidad, resulta que la transformada de Gelfand es un *-isomorfismo isométrico de $\overline{\mathcal{U}}$ en $C(\mathcal{M}_{\overline{\mathcal{U}}})$, y la imagen de $\mathcal{U}$ es el ideal formado por todas las funciones que se anulan en ϕ_0, que se identifica naturalmente con $C_0(\mathcal{M}_{\mathcal{U}})$. □

Corolario 3.3.5. *Si N es un elemento normal de una C^*-álgebra con unidad $\mathcal{U}$, entonces $C^*(N)$ es isométricamente isomorfo a $C(\sigma(N))$, las funciones continuas en el espectro de N, vía la aplicación que manda a N a la función identidad. La imagen de la C^*-álgebra generada por N y N^* a través de esta aplicación es $C_0(\sigma(N)-\{0\})$*

Demostración. Como N es normal, $C^*(N)$ resulta abeliano. Luego, por el Teorema 3.3.4, $C^*(N)$ es isométricamente *-isomorfo a $C_0(X)$, donde $X = \mathcal{M}_{C^*(N)}$. Veamos que X es homeomorfo a $\sigma(N)$. Toda funcional ϕ en X queda determinada por $\phi(N)$, pues $\phi(p(N,N^*)) = p(\phi(N), \overline{\phi(N)})$, para todo polinomio p. Luego la aplicación de X en $\mathbb{C}$ que evalúa a la funcional en N es un homeomorfismo de X en $\widehat{N}(X) = \{\phi(N) : \phi \in X\}$, pero $\widehat{N}(X) = \sigma(N)$ (por Corolario 3.3.3).

La aplicación que queda definida de $C^*(N)$ en $C(\sigma(N))$ manda a N en la identidad.

En el caso en que N no es inversible, la subálgebra generada por N y N^* (pero no por I) se corresponde con el ideal de funciones que se anulan en 0. □

Corolario 3.3.6 (Cálculo funcional continuo de un elemento normal)**.** *Si N es un elemento normal de una C^*-álgebra con unidad, y f una función continua sobre $\sigma(N)$, definimos el operador $f(N)$ como $\Gamma^{-1}(f)$. Esta aplicación es un *-isomorfismo isométrico de $C(\sigma(N))$ en $C^*(N)$. Si $0 \in \sigma(N)$ y $f(0) = 0$, entonces $f(N)$ pertenece al álgebra (sin unidad) generada por N y N^*. Más aún, $\sigma(f(N)) = f(\sigma(N))$, y si g es continua en $f(\sigma(N))$, entonces $g(f(N)) = (gof)(N)$.*

Demostración. Por el Corolario 3.3.5, se tiene que la aplicación es un *-isomorfismo isométrico. La igualdades

$$\sigma(f(N)) = \sigma(\widehat{f(N)}) = \sigma(f) = f(\sigma(N))$$

son consecuencia inmediata del Corolario 3.3.3 y de la definición de $f(N)$. Si p es un polinomio en z y $\overline{z}$, es inmediato, a partir del hecho de que el calculo funcional es un morfismo, que $p(f(N)) = (p \circ f)(N)$. El caso general (g continua) se deduce aproximando g por polinomios en z y $\overline{z}$. $\square$

Corolario 3.3.7. *Sea $\mathcal{U}$ una C^*-álgebra, $N \in \mathcal{U}$ normal, entonces $\|N\| = spr(N)$. Si $A \in \mathcal{U}$ es autoadjunto, $\sigma(A) \subseteq \mathbb{R}$.*

Demostración. Por el Corolario 3.3.5, tenemos

$$\|N\| = \|\widehat{N}\|_{\sigma(N)} = \|Id\|_{\sigma(N)} = spr(N).$$

Si A es autoadjunto, entonces $\overline{\widehat{A}} = \widehat{A^*} = \widehat{A}$ es una función real. Como $\sigma(A)$ es igual al rango de $\widehat{A}$, se sigue que $\sigma(A) \subseteq \mathbb{R}$. $\square$

3.4. Representación GNS.

En esta sección probaremos que toda C^*-álgebra admite una representación como una subálgebra cerrada de $\mathcal{B}(H)$ para algún espacio de Hilbert $\mathcal{H}$ adecuado. Esto se prueba mediante una construcción debida a Gelfand, Neimark y Segal (GNS).

Definición 3.4.1. *Una funcional ρ definida sobre $\mathcal{U}$ es* positiva *si $\rho(A^*A) \geq 0 \; \forall A \in \mathcal{U}$. Una funcional de norma unitaria es un* estado*. Al conjunto de los estados de $\mathcal{U}$ se lo nota $\mathcal{S}(\mathcal{U})$.*

Lema 3.4.2. *Sea $\mathcal{U}$ una C^*-álgebra con unidad. Una funcional $\rho : \mathcal{U} \to \mathbb{C}$ es positiva si y sólo si es acotada y $\rho(1) = \|\rho\|$.*

Demostración. Supongamos que ρ es positiva, sea $A \in \mathcal{U}$, y $\Theta \in [0, 2\pi)$ tal que

$$e^{i\Theta}\rho(A) = |\rho(A)| \geq 0,$$

sea

$$h = Re(e^{i\Theta}A) = \frac{1}{2}(e^{i\Theta}A + e^{-i\Theta}A^*).$$

Luego $h = h^*$ y entonces $h \leq \|h\| \leq \|A\|$, con lo cual

$$\|A\|\rho(1) - \rho(h) = \rho(\|A\| - h) \geq 0.$$

Así

$$|\rho(A)| = \rho(e^{i\Theta}A) = \overline{\rho(e^{i\Theta}A)} = \rho(e^{-i\Theta}A^*) = \rho(h) \leq \rho(1)\|A\|,$$

por lo que ρ es acotada, con $\|\rho\| = \rho(1)$.

Supongamos ahora que ρ es acotada y que $\rho(1) = \|\rho\|$. Sin pérdida de generalidad, podemos suponer $\|\rho\| = 1$. Sea $A \geq 0$, y $\rho(A) = \alpha + i\beta$. Veamos que $\alpha \geq 0$ y $\beta = 0$. Sea $s \in \mathbb{R}$ positivo y suficientemente pequeño de modo que

$$\sigma(1 - sA) = \{1 - st : t \in \sigma(A)\} \subset [0, 1].$$

Como $1 - sA$ es autoadjunto, $\|1 - sA\| = spr(1 - sA) \leq 1$, luego

$$0 \leq |1 - s\alpha| \leq |1 - s\alpha - is\beta| = |\rho(1 - sA)| \leq \|1 - sA\| \leq 1,$$

de donde se deduce inmediatamente que $\alpha \geq 0$. Definiendo $B_n = A - \alpha + in\beta$, con $n \in \mathbb{N}$, se tiene

$$\|B_n\|^2 = \|B_n^* B_n\| = \|(A - \alpha)^2 + n^2\beta^2\| \leq \|A - \alpha\|^2 + n^2\beta^2.$$

Entonces,

$$(n^2 + 2n + 1)\beta^2 = |i(n + 1)\beta|^2 = |\rho(B_n)|^2 \leq \|A - \alpha\|^2 + n^2\beta^2$$

para todo n, lo cual sólo es posible si $\beta = 0$. □

Lema 3.4.3. *Dado $A \in \mathcal{U}$ normal, existe $\phi \in \mathcal{S}(\mathcal{U})$ que es* normizante *para A, es decir*

$$|\phi(A)| = \|A\|.$$

Demostración. Como A es normal, $\|A\| = spr(A)$, luego existe $\lambda \in \sigma(A)$ tal que $|\lambda| = \|A\|$. Sea $h_0 \in \mathcal{M}_{C^*(A)}$, tal que $h_0(A) = \lambda$. Por Hahn-Banach, lo extendemos a una funcional lineal ϕ sobre $\mathcal{U}$, de modo que $\|\phi\| = \|h_0\| = 1$. Además, como

$$\phi(1) = h_0(1) = 1 = \|\phi\|,$$

por el lema anterior resulta que ϕ es un estado. □

Definición 3.4.4. *Una* *-representación *de una C^*-álgebra $\mathcal{U}$ es un *-morfismo $\pi : \mathcal{U} \to \mathcal{B}(H)$ con $\mathcal{H}$ un espacio de Hilbert.*

Proposición 3.4.5. *Sea $A \in \mathcal{U}$, entonces existe una *-representación de $\mathcal{U}$ que preserva la norma de A.*

Demostración. Sea ϕ un estado normizante para A^*A (existe por el lema anterior). Sea $\langle , \rangle_\phi : \mathcal{U} \times \mathcal{U} \to \mathbb{C}$ la forma sesquilineal dada por

$$\langle X, Y \rangle_\phi = \phi(Y^*X).$$

Sea $\|X\|_\phi = \sqrt{\langle X, X \rangle_\phi}$. Como ϕ es positiva, la forma sesquilineal es definida positiva, y por lo tanto vale la desigualdad de Cauchy-Schwarz

$$|\langle X, Y \rangle| \leq \|X\|_\phi \|Y\|_\phi.$$

Sea

$$\mathcal{N}_\phi = \{X \in \mathcal{U} : \|X\|_\phi = 0\} \subset \mathcal{U}.$$

Por Cauchy-Schwarz, si $x \in \mathcal{N}_\phi$, entonces $XY, YX \in \mathcal{N}_\phi$ para todo $Y \in \mathcal{U}$, o sea que $\mathcal{N}_\phi$ es un ideal bilátero propio (pues $A \notin \mathcal{N}_\phi$). Sea $\mathcal{U}_\phi = \mathcal{U}/\mathcal{N}_\phi$, el espacio cociente, entonces es fácil ver que $\langle , \rangle_\phi : \mathcal{U}_\phi \times \mathcal{U}_\phi \to \mathbb{C}$ es un producto interno. Sea $\mathcal{H}_\phi$ la completación de $\mathcal{U}_\phi$ respecto a la norma inducida por el producto interno, obviamente $\mathcal{H}_\phi$ es un espacio de Hilbert.

Sea $\pi_A : \mathcal{U} \to \mathcal{B}(\mathcal{H}_\phi)$ dada por $X \mapsto L_X$, donde $L_X(\nu) = \overline{X}\nu$ (la clase de X en $\mathcal{H}_\phi$). Es fácil ver que $\pi_A(X+Y) = \pi_A(X) + \pi_A(Y)$ y $\pi_A(XY) = \pi_A(X)\pi_A(Y)$ para todo $X, Y \in \mathcal{U}$, además $\pi_A(X^*) = (\pi_A(X))^*$, pues

$$\langle \pi_A(X^*)\nu, \xi \rangle_\phi = \phi(\xi^* X^* \nu) = \phi((X\xi)^*\nu) = \langle \nu, X\xi \rangle_\phi = \langle \nu, \pi_A(X)\xi \rangle_\phi.$$

Como π_A es un *-morfismo, se tiene $\|\pi_A(X)\| \leq \|X\|$. Además, como

$$\|A\|_\phi = \sqrt{\phi(A^*A)} = \sqrt{\|A\|^2} = \|A\|,$$

la representación preserva la norma de A como elemento de $\mathcal{H}_\phi$. Por último

$$|\pi_A(A)(1)| = \|L_A(1)\|^2 = \phi(A^*A) = \|A\|^2,$$

con lo que probamos que π_A preserva la norma de A. □

Corolario 3.4.6. *(Representación GNS) Toda C^*-álgebra $\mathcal{U}$ se representa como una subálgebra cerrada de $\mathcal{B}(H)$ para algún espacio de Hilbert adecuado.*

Demostración. Como para todo $A \in \mathcal{U}$ se tiene $\pi_A : \mathcal{U} \to \mathcal{B}(\mathcal{H}_A)$ que preserva la norma de A, podemos considerar la suma directa de estas aplicaciones

$$\oplus_{A \in \mathcal{U}} \pi_A : \mathcal{U} \to \mathcal{B}(\oplus_{A \in \mathcal{U}} \mathcal{H}_\phi),$$

donde en $\oplus_{A \in \mathcal{U}} \mathcal{H}_\phi$ consideramos la norma del supremo de las normas de las coordenadas. Por la Proposición 3.4.5 resulta que es una *-representación, y es inyectiva pues es una isometría. La imagen es cerrada por ser una isometría entre espacios de Banach. □

Capítulo 4

Sobre la técnica de E. Effros.

En este capítulo presentamos demostraciones publicadas por E. Effros [10] de algunos resultados conocidos, y las extendemos al contexto infinito dimensional. Si bien el trabajo de Effros se limita a matrices, las mismas ideas permiten probar los resultados correspondientes para operadores. Estas demostraciones se basan en una nueva versión de la desigualdad de Jensen, la perspectiva clásica y la de Marechal. Trabajaremos con operadores en espacios de Hilbert de dimensión infinita; muchas definiciones y demostraciones son completamente análogas al caso finito dimensional y serán omitidas.

4.1. Funciones monótonas y convexas de operadores.

Definición 4.1.1. *Decimos que una función real continua f es* monótona de operadores *en un intervalo I, si para todo par de operadores A, B con espectro en I tales que $A \leq B$, resulta $f(A) \leq f(B)$. Decimos que f es convexa de operadores en I si para todo $c \in [0, 1]$ resulta*

$$f(cA + (1 - c)B) \leq cf(A) + (1 - c)f(B).$$

Observación 4.1.2. El cálculo funcional definido en el Corolario 3.3.6 sólo nos permite definir funciones continuas sobre operadores normales en espacios de Hilbert de dimensión infinita, por lo que nos limitaremos a trabajar con ese tipo de funciones. En virtud de las Observaciones 2.4.8 y 2.4.9, esta restricción no es muy severa en lo referente a monotonía y convexidad.

Proposición 4.1.3. *Una función es convexa (monótona) de operadores en un intervalo I, si y sólo si es convexa (monótona) de matrices en I.*

Demostración. Es claro, usando proyecciones a subespacios de dimensión finita, que toda función monótona (convexa) de operadores, también lo es de matrices. Veamos que toda función monótona (convexa) de matrices, también lo es de operadores. Sea f monótona de matrices, entonces por el Teorema 2.3.27, existe una sucesión de polinomios q_n que converge uniformemente a f sobre cualquier compacto en I. Dado x en un espacio de Hilbert $\mathcal{H}$ sea p_n la proyección al subespacio generado por $x, Ax, A^2x, ..., A^jx, Bx, B^2x, ..., B^jx$, con j el grado de q_n. Así, tenemos

$$\begin{aligned} \|[f(A) - f(p_nAp_n)]x\| &\leq \|(f - q_n)(A)x\| + \|[q_n(A) - q_n(p_nAp_n)]x\| \\ &+ \|(q_n - f)(p_nAp_n)x\| \end{aligned}$$

el segundo sumando es igual a 0, y los otros tienden a 0, luego $f(p_nAp_n)x \to f(A)x$. Sean A, B operadores con espectro en I tales que $A \leq B$, luego $pAp \leq pBp$ para cualquier proyección p, en particular para proyecciones finito dimensionales, luego, por la monotonía de f para matrices, resulta $f(p_nAp_n) \leq f(p_nBp_n)$. Tomando $n \to \infty$ resulta

$$\langle x, f(A)x\rangle \leftarrow \langle x, f(p_nAp_n)x\rangle \leq \langle x, f(p_nBp_n)x\rangle \to \langle x, f(B)x\rangle.$$

Como esta construcción se puede hacer para cada x, resulta que f es monótona de operadores. Con la misma idea se prueba la propiedad correspondiente para convexidad.

□

4.2. Operadores traza y operadores Hilbert-Schmidt.

En esta sección seguiremos el libro de G.K. Pedersen [23].

El siguiente teorema, generalización del Teorema 1.1.18, será utilizado en repetidas oportunidades a lo largo de la sección.

Teorema 4.2.1 (Descomposición polar de operadores en $\mathcal{B}(\mathcal{H})$)**.** *Dado $T \in \mathcal{B}(\mathcal{H})$, existe un único operador positivo $|T|$ tal que*

$$\|Tx\| = \||T|x\| \; \forall x \in \mathcal{H}, \tag{4.1}$$

y resulta $|T| = (T^*T)^{\frac{1}{2}}$. Más aún, existe una única isometría parcial U tal que $ker(U) = ker(T)$ y $U|T| = T$. En particular, $U^*U|T| = |T|$, $U^*T = |T|$ y $UU^*T = T$.

Demostración. Si $S \geq 0$ es tal que $\|Sx\| = \|Tx\|$ para todo x, entonces

$$\langle S^2 x, x\rangle = \|Sx\|^2 = \|Tx\|^2 = \langle T^*Tx, x\rangle \ \forall x \in \mathcal{H}.$$

Luego $S^2 = T^*T$. Luego $S = (T^*T)^{\frac{1}{2}}$, por la unicidad de la raíz cuadrada. Notando $|T| = (T^*T)^{\frac{1}{2}}$ tenemos

$$ker(T) = ker(|T|) = (Im(|T|))^{\perp}, \text{ pues } |T| = |T|^*.$$

Para cada $y = |T|x$, definimos

$$U_0 y = U_0|T|x = Tx.$$

Se sigue de (4.1), que U_0 está bien definida y es una isometría entre las imágenes de $|T|$ y T. Por continuidad U_0 se extiende a una isometría $U : \overline{Im(|T|)} \to \overline{Im(T)}$. Definiendo $U = 0$ sobre $ker(T)$ $(= Im(|T|)^{\perp})$, resulta que U es una isometría parcial tal que $U|T| = T$. Si V es otra isometría parcial con $ker(V) = ker(U)$ y $V|T| = T$, entonces para todo $y = |T|x$ se tiene $Vy = Uy$, por lo que $U = V$ sobre $\overline{Im(|T|)}$. Como además, tanto U como V son nulos sobre $(Im(|T|))^{\perp}$, resulta $U = V$. Como U^*U es la proyección sobre $\overline{Im(|T|)}$, se tiene $U^*U|T| = |T|$, y como $U|T| = T$ resulta $U^*T = |T|$ y, multiplicando a izquierda por U, queda $UU^*T = T$. □

Observación 4.2.2. A diferencia del caso finito dimensional, no se puede pedir, en general, que U sea unitario.

Definición 4.2.3. *Dado un operador positivo $T \in \mathcal{B}(\mathcal{H})$, y fijada $(e_j)_j$, una b.o.n. de $\mathcal{H}$, definimos la* traza *de T como*

$$tr(T) = \sum_j \langle Te_j, e_j\rangle.$$

La traza toma valores en $[0, \infty]$.

Proposición 4.2.4. *Para todo $T \in \mathcal{B}(\mathcal{H})$ se tiene*

$$tr(TT^*) = tr(T^*T).$$

Demostración. Para cada i, j tenemos

$$\langle Te_i, e_j\rangle\langle e_j, Te_i\rangle = \langle T^*e_j, e_i\rangle\langle e_i, T^*e_j\rangle \geq 0.$$

Sumando la primera expresión sobre j, tenemos

$$\sum_j \langle\langle Te_i, e_j\rangle e_j, Te_i\rangle = \langle Te_i, Te_i\rangle = \langle T^*Te_i, e_i\rangle.$$

Del mismo modo sumando sobre i la segunda expresión tenemos

$$\sum_i \langle \langle T^* e_j, e_i \rangle e_i, T^* e_j \rangle = \langle TT^* e_j, e_j \rangle.$$

Como los sumandos de las series son positivos, las sumas no dependen del orden de los sumandos, luego

$$tr(T^*T) = \sum_i \langle T^*Te_i, e_i \rangle = \sum_j \langle TT^* e_j, e_j \rangle = tr(TT^*).$$

□

Corolario 4.2.5. *Si U es unitario y $T \geq 0$, entonces*

$$tr(UTU^*) = tr(T).$$

En particular, la traza es independiente de la base.

Demostración. Como $T = (T^{\frac{1}{2}})^2$, cambiando T por $UT^{\frac{1}{2}}$ en la proposición anterior, se tiene el corolario. □

Definición 4.2.6. *Definimos los conjuntos de* operadores traza *(o* nucleares*) y de* operadores Hilbert-Schmidt *como*

$$\mathcal{B}_1(\mathcal{H}) = span\{T : T \text{ compacto, positivo y } tr(T) < \infty\},$$

$$\mathcal{B}_2(\mathcal{H}) = \{T : T \text{ compacto, } tr(T^*T) < \infty\}.$$

Como claramente $tr(A + B) = tr(A) + tr(B)$ y $tr(\alpha A) = \alpha tr(A)$ para todo par de operadores positivos A, B, y $\alpha \geq 0$ y como todo operador de traza T se puede escribir como $T = \sum_{k=0}^{3} i^k T_k$, con $T_k \geq 0$, se sigue que definiendo

$$tr(T) = \sum_{k=0}^{3} i^k tr(T_k),$$

se obtiene una extensión de la función traza, que resulta ser una funcional lineal en $\mathcal{B}_1(\mathcal{H})$. Con esto podemos aplicar la función traza a cualquier operador del tipo $P+T$ con $P \geq 0$ y $T \in \mathcal{B}_1(\mathcal{H})$.

Lema 4.2.7. *Dados $S, T \in \mathcal{B}(\mathcal{H})$, se verifica la* ley del paralelogramo*:*

$$(S + T)^*(S + T) + (S - T)^*(S - T) = 2(S^*S) + (T^*T). \tag{4.2}$$

De esto se deduce trivialmente la acotación

$$(S+T)^*(S+T) \leq 2(S^*S + T^*T). \tag{4.3}$$

También vale la siguiente igualdad, conocida como identidad de polarización

$$4T^*S = \sum_{k=0}^{3} i^k (S + i^k T)^*(S + i^k T). \tag{4.4}$$

Demostración. Simplemente desarrollando las expresiones se obtienen (4.2) y (4.4). De (4.2) se deduce (4.3) pues $(S-T)^*(S-T) \geq 0$. □

Observación 4.2.8. De la ley del paralelogramo se deduce que los operadores de Hilbert-Schmidt forman un subespacio.

Lema 4.2.9. *Si $T \in \mathcal{B}(\mathcal{H})$ satisface $tr(|T|^p) < \infty$ para algún $p > 0$, entonces T es compacto.*

Demostración. Dada una b.o.n. $(e_j)_j$ y $\epsilon > 0$, existe un subconjunto finito $\mathbb{F}$ de índices tal que $\sum_{j \notin \mathbb{F}} \langle |T|^p e_j, e_j \rangle < \epsilon$. Notando $P_{\mathbb{F}}$ a la proyección sobre $span\{e_j : j \in \mathbb{F}\}$ tenemos

$$\begin{aligned} \| |T|^{\frac{p}{2}} (I - P_{\mathbb{F}}) \|^2 &= \| (I - P_{\mathbb{F}}) |T|^p (I - P_{\mathbb{F}}) \| \\ &\leq tr((I - P_{\mathbb{F}}) |T|^p (I - P_{\mathbb{F}})) < \epsilon. \end{aligned}$$

Como ϵ es arbitrario, resulta que $|T|^{\frac{p}{2}}$ es límite de operadores de rango finito y por lo tanto compacto. Luego existe una b.o.n., a la que también llamaremos $(e_j)_j$ formada por autovectores de $|T|^{\frac{p}{2}}$, con e_j asociado al autovalor λ_j. Así resulta $|T| e_j = \lambda^{\frac{2}{p}} e_j$ y por lo tanto $|T|$ es compacto. Por la descomposición polar $T = U|T|$ y el hecho de que los operadores compactos forman un ideal en $\mathcal{B}(\mathcal{H})$ resulta que T es compacto. □

Corolario 4.2.10. *Dado un operador $T > 0$ se tiene*

$$\|T\| \leq tr(T).$$

Demostración. Si $tr(T) = \infty$ no hay nada que probar. Si $tr(T) < \infty$, por el lema anterior resulta que T es compacto. Entonces, como además T es autoadjunto, hay una b.o.n. $(e_i)_i$ de $\mathcal{H}$ formada por autovectores de T. Como la traza no depende de la base (Corolario 4.2.5), usando en particular la base $(e_i)_i$, es claro que $tr(T) \geq \|T\|$. □

Proposición 4.2.11. *Las clases $\mathcal{B}_1(\mathcal{H})$ y $\mathcal{B}_2(\mathcal{H})$ son ideales autoadjuntos de $\mathcal{B}(\mathcal{H})$ y valen las siguientes contenciones*

$$\mathcal{B}_f(\mathcal{H}) \subseteq \mathcal{B}_1(\mathcal{H}) \subseteq \mathcal{B}_2(\mathcal{H}) \subseteq \mathcal{B}_0(\mathcal{H}),$$

donde $\mathcal{B}_f(\mathcal{H})$ es la clase de operadores de rango finito y $\mathcal{B}_0(\mathcal{H})$ es la de operadores compactos.

Demostración. La inclusión $\mathcal{B}_2(\mathcal{H}) \subseteq \mathcal{B}_0(\mathcal{H})$ se deduce del lema anterior. Sean $T \geq 0$ tal que $tr(T) < \infty$, y $S \in \mathcal{B}(\mathcal{H})$. Luego, por (4.4) se tiene

$$4TS = 4T^{\frac{1}{2}}T^{\frac{1}{2}}S = \sum_{k=0}^{3} i^k(S + i^k I)^* T(S + i^k I),$$

y por la Proposición 4.2.4 se tiene

$$\begin{aligned} tr(V^*TV) &= tr(V^*T^{\frac{1}{2}}T^{\frac{1}{2}}V) = tr(T^{\frac{1}{2}}VV^*T^{\frac{1}{2}}) \\ &= \sum_i \langle T^{\frac{1}{2}}VV^*T^{\frac{1}{2}}e_i, e_i\rangle = \sum_i \langle VV^*T^{\frac{1}{2}}e_i, T^{\frac{1}{2}}e_i\rangle \\ &\leq \sum_i \|VV^*\|\|T^{\frac{1}{2}}e_i\|^2 = \|VV^*\| \sum_i \langle T^{\frac{1}{2}}e_i, T^{\frac{1}{2}}e_i\rangle \\ &= \|VV^*\| tr(T), \end{aligned} \tag{4.5}$$

donde (e_i) es una b.o.n. de $\mathcal{H}$. Tomando en particular $V = S + i^k I$, se deduce que $TS \in \mathcal{B}_1(\mathcal{H})$ y por lo tanto $\mathcal{B}_1(\mathcal{H})$ es un ideal a derecha autoadjunto. Esto implica que es un ideal bilátero, pues $ST = (T^*S^*)^*$. Veamos ahora que

$$\mathcal{B}_1(\mathcal{H}) = \{T \in \mathcal{B}(\mathcal{H}) : tr(|T|) < \infty\}.$$

Si $|T| \in \mathcal{B}_1(\mathcal{H})$, por la descomposición polar $T = U|T|$, se ve por la primera parte de la demostración que $T \in \mathcal{B}_1(\mathcal{H})$. Por otro lado, $|T| = U^*T$, luego si $T \in \mathcal{B}_1(\mathcal{H})$, entonces $|T| \in \mathcal{B}_1(\mathcal{H})$ también. De (4.3) se sigue que $\mathcal{B}_2(\mathcal{H})$ es un subespacio de $\mathcal{B}_0(\mathcal{H})$, y la Proposición 4.2.4 muestra que el subespacio es autoadjunto. Como $\mathcal{B}_1(\mathcal{H})$ es ideal $\mathcal{B}(\mathcal{H})$, se sigue de la definición de $\mathcal{B}_2(\mathcal{H})$ que los operadores de Hilbert-Schmidt también son un ideal. Si $|T| \in \mathcal{B}_f(\mathcal{H})$, entonces $|T|$ es diagonalizable y de rango finito, luego $|T|$ (y T) pertenece a $\mathcal{B}_1(\mathcal{H})$. Si $T \in \mathcal{B}_1(\mathcal{H})$, entonces se tiene

$$T^*T = |T|^2 = |T|^{\frac{1}{2}}|T||T|^{\frac{1}{2}} \leq \|T\||T|,$$

lo que prueba que $tr(T^*T) < \infty$, i.e. $T \in \mathcal{B}_2(\mathcal{H})$. $\square$

Lema 4.2.12. *Si $S, T \in \mathcal{B}_2(\mathcal{H})$, entonces*

$$tr(ST) = tr(TS).$$

Lo mismo vale si $S \in \mathcal{B}(\mathcal{H})$ y $T \in \mathcal{B}_1(\mathcal{H})$.

Demostración. Por la identidad de polarización (4.4) y la Proposición 4.2.4 se tiene

$$\begin{aligned} 4tr(T^*S) &= \sum_{k=0}^{3} i^k tr((S + i^k T)^*(S + i^k T)) \\ &= \sum_{k=0}^{3} i^k tr((S^* + i^{-k} T^*)^*(S^* + i^{-k} T^*)) \\ &= \sum_{k=0}^{3} i^k tr((T^* + i^k S^*)^*(T^* + i^k S^*)) = 4tr(ST^*), \end{aligned}$$

lo que prueba la primera afirmación. Para la segunda, podemos suponer, sin pérdida de generalidad, que $T \geq 0$, pues $tr(TS)$ y $tr(ST)$ son lineales en T. Luego, por la primera parte resulta

$$\begin{aligned} tr(ST) &= tr((ST^{\frac{1}{2}})T^{\frac{1}{2}}) = tr(T^{\frac{1}{2}}(ST^{\frac{1}{2}})) \\ &= tr((T^{\frac{1}{2}}S)T^{\frac{1}{2}}) = tr(T^{\frac{1}{2}}(T^{\frac{1}{2}}S)) = tr(TS). \end{aligned}$$

□

Proposición 4.2.13. *$\mathcal{B}_2(\mathcal{H})$ es un espacio de Hilbert con el producto interno*

$$\langle S, T \rangle = tr(T^*S).$$

Demostración. De (4.4) se sigue que $T^*S \in \mathcal{B}_1(\mathcal{H})$. Luego $\langle\ ,\ \rangle$ está bien definida, y además es claro que es sesquilineal, definida positiva y $\langle A, B \rangle = \overline{\langle B, A \rangle}$. Más aún, es un producto interno en $\mathcal{B}_2(\mathcal{H})$, pues la 2-norma asociada verifica

$$\|T\|_2^2 = tr(T^*T) \geq \|T^*T\| = \|T\|^2,$$

por el Corolario 4.2.5. Esta desigualdad implica también que toda sucesión de Cauchy $(T_n)_n \in \mathcal{B}_2(\mathcal{H})$ para la 2-norma converge en norma a un elemento $T \in \mathcal{B}_0(\mathcal{H})$ (la clase de operadores compactos). Para toda proyección P sobre un subespacio de dimensión infinita de $\mathcal{H}$ tenemos

$$\begin{aligned} \|P(T - T_n)\|_2^2 &= tr((T - t_n)^*P(T - T_n)) = tr(P(T - T_n)(T - T_n)^*P) \\ &= \lim_m tr(P(T_m - T_n)(T_m - T_n)^*P) \\ &= \lim_m tr((T_m - T_n)^*P(T_m - T_n)) \\ &\leq \limsup_m tr((T_m - T_n)^*(T_m - T_n)) = \limsup_m \|T_m - T_n\|_2^2, \end{aligned}$$

y, como P es arbitrario, concluimos que

$$\|T - T_n\|_2 \leq \limsup_m \|T_m - T_n\|_2,$$

lo que implica que $T \in \mathcal{B}_2(\mathcal{H})$ y $T_n \to T$ en 2-norma. $\square$

Lema 4.2.14. *Si $T \in \mathcal{B}_1(\mathcal{H})$ y $S \in \mathcal{B}(\mathcal{H})$, entonces*

$$|tr(ST)| \leq \|S\| tr(|T|).$$

Demostración. Sea $T = U|T|$ la descomposición polar de T. Luego $(SU|T|^{\frac{1}{2}})^* \in \mathcal{B}_2(\mathcal{H})$ (pues $|T|^{\frac{1}{2}} \in \mathcal{B}_2(\mathcal{H})$), así, por la desigualdad de Cauchy-Schwarz y la Proposición 4.2.4 tenemos

$$\begin{aligned} |tr(ST)|^2 &= |tr(SU|T|^{\frac{1}{2}}|T|^{\frac{1}{2}})|^2 = |\langle |T|^{\frac{1}{2}}, (SU|T|^{\frac{1}{2}})^* \rangle|^2 \\ &\leq \||T|^{\frac{1}{2}}\|_2^2 \|(SU|T|^{\frac{1}{2}})^*\|_2^2 = tr(|T|) tr(|T|^{\frac{1}{2}} U^* S^* SU |T|^{\frac{1}{2}}) \\ &\leq tr(|T|) tr(\|U^* S^* SU\| |T|) \leq \|S\|^2 (tr(|T|))^2. \end{aligned} \tag{4.6}$$

$\square$

4.3. La desigualdad de Jensen-Hansen-Pedersen.

Lema 4.3.1. *Sea $E = diag(\xi, \xi^2, ..., 1) \in M_n(\mathbb{C}) \subset \mathcal{B}(\mathcal{H}^n)$, con $\xi = exp(\frac{2\pi i}{n})$ y $\mathcal{H}^n = \mathcal{H} \oplus ... \oplus \mathcal{H}$. Entonces para todo $A = (a_{ij}) \in \mathcal{B}(\mathcal{H}^n)$ se tiene*

$$\frac{1}{n} \sum_{k=1}^{n} E^{-k} A E^k = diag(a_{11}, a_{22}, ..., a_{nn}).$$

Demostración. Observemos que

$$(\sum_{k=1}^{n} E^{-k} A E^k)_{ij} = \frac{1}{n} \sum_{k=1}^{n} (\xi^{j-i})^k a_{ij},$$

y esta suma da 0 si $i \neq j$ y da a_{ii} en el caso $i = j$. $\square$

Definición 4.3.2. *Una n-upla $a = (a_1, ..., a_n)$ de operadores en $\mathcal{B}(\mathcal{H})$ es una columna unital (contractiva) si $\sum_{i=1}^n a_i^* a_i = 1$ (≤ 1). Análogamente una fila es unital (contractiva) si $\sum_{i=1}^n a_i a_i^* = 1$ (≤ 1).*

Decimos que a es una columna unitaria *si existe una matriz de operadores unitaria $U = (u_{ij})$, tal que una de sus columnas es a. Análogamente se definen las filas unitarias.*

Lema 4.3.3. *Toda n-columna contractiva se puede extender a una (n+1)-columna unitaria.*

Demostración. Dada $(a_1, ..., a_n)$ n-columna contractiva, afirmamos que si tomamos $a = -(1 - \sum_{i=1}^n a_i^* a_i)^{\frac{1}{2}}$, entonces $(a_1, ..., a_n, a)$ es una columna unitaria.

Tomemos A la matriz de $n \times n$ cuya primera columna es $(a_1, ..., a_n)$ y el resto de sus coordenadas es cero. Entonces A es una contracción.

La matriz

$$U = \begin{pmatrix} A & (1 - AA^*)^{\frac{1}{2}} \\ (1 - A^*A)^{\frac{1}{2}} & -A^* \end{pmatrix}$$

es unitaria (ver Lema 2.3.18). Pero $(1 - A^*A)^{\frac{1}{2}} = \begin{pmatrix} (1 - \sum_{i=1}^n a_i)^{\frac{1}{2}} & 0 \\ 0 & 1 \end{pmatrix}$ (escrito por bloques).

Intercambiando las primeras n columnas de U con las últimas (la i con la $n+i$), obtenemos la matriz por bloques

$$\begin{pmatrix} & & & a_1 & 0 \\ & (1 - AA^*)^{\frac{1}{2}} & & \vdots & 0 \\ & & & a_n & \\ -a_1^* & ... & -a_n^* & I - \sum_{i=1}^n a_i^* a_i & 0 \\ 0 & ... & ... & 0 & I \end{pmatrix}.$$

□

El siguiente teorema, publicado por F. Hansen y G. K. Pedersen en [13], es un refinamiento del Teorema 2.3.20. Elimina la necesidad de que 0 pertenezca al dominio de f y, en consecuencia, la necesidad de que $f(0) \leq 0$.

Teorema 4.3.4. *(Desigualdad de operadores de Jensen.)*

Sea f continua en un intervalo I. Son equivalentes:

1. *f es convexa de operadores.*

2. *Para cada $n \in \mathbb{N}$ se tiene*

$$f(\sum_{i=1}^n a_i^* x_i a_i) \leq \sum_{i=1}^n a_i^* f(x_i) a_i$$

con $(x_1, ..., x_n)$ operadores autoadjuntos y acotados en un espacio de Hilbert $\mathcal{H}$, con espectro en I, y $(a_1, ..., a_n)$ operadores acotados en $\mathcal{H}$ tales que $\sum_{i=1}^n a_i^* a_i = 1$.

3. *$f(v^*xv) \leq v^*f(x)v$ para toda isometría v en un espacio de Hilbert de dimensión infinita, y todo operador autoadjunto x con espectro en I.*

4. *$pf(pxp + s(1-p))p \leq pf(x)p$ para toda proyección p en un espacio de Hilbert de dimensión infinita, todo operador autoadjunto x con espectro en I, y todo $s \in I$.*

Demostración. $(1 \Rightarrow 2)$.

Supongamos que $(a_1, ..., a_n)$ es una columna unitaria, sea $U_n = (u_{ij}) \in B(\mathcal{H}^n)$ tal que $u_{kn} = a_k$. Sea $E = diag(\xi, \xi^2, ..., 1)$, y sea $X = diag(x_1, ..., x_n)$, entonces tenemos

$$\begin{aligned}
f(\sum_{k=1}^{n} a_k^* x_k a_k) &= f((U_n^* X U_n)_{nn}) \\
&= f((\sum_{k=1}^{n} \frac{1}{n} E^{-k} U_n^* X U_n E^k)_{nn}) \qquad (4.7) \\
&= (f(\sum_{k=1}^{n} \frac{1}{n} E^{-k} U_n^* X U_n E^k))_{nn} \leq (\frac{1}{n} \sum_{k=1}^{n} f(E^{-k} U_n^* X U_n E^k))_{nn} \\
&= (\frac{1}{n} \sum_{k=1}^{n} E^{-k} U_n^* f(X) U_n E^k)_{nn} = (U_n f(X) U_n)_{nn} = \sum_{k=1}^{n} a_k^* f(x_k) a_k
\end{aligned}$$

Donde en la igualdad (4.7) usamos el Lema 4.3.1.

Veamos el caso general, en el que la columna es unital. En este caso consideramos la columna unitaria $(a_1, ..., a_n, 0)$, y x_n arbitrario con espectro en I. Luego por el caso anterior tenemos

$$\begin{aligned}
f(\sum_{k=1}^{n} a_k^* x_k a_k) &= f(\sum_{k=1}^{n+1} a_k^* x_k a_k) \\
&\leq \sum_{k=1}^{n+1} a_k^* f(x_k) a_k = \sum_{k=1}^{n} a_k^* f(x_k) a_k.
\end{aligned}$$

$(2 \Rightarrow 3)$.

Es inmediato pues si v es una isometría, $v^*v = 1$.

$(3 \Rightarrow 4)$.

Sean x un operador autoadjunto con espectro en I, y p una proyección a un subespacio de dimensión infinita. Entonces podemos tomar una isometría v tal que $vv^* = p$. Por hipótesis $f(v^*xv) \leq v^*f(x)v$, luego

$$vf(v^*xv)v^* \leq vv^*f(x)vv^* = pf(x)p. \qquad (4.8)$$

Sean $g(t) = t^m$, y $s \in I$,

$$\begin{aligned} pvg(v^*xv)v^*p &= pv(v^*xv)^m v^*p \\ &= p(vv^*xvv^*)^m p \\ &= pg(pxp)p \\ &= pg(pxp + s(1-p))p. \end{aligned}$$

Luego, por el teorema de Stone-Weierstrass, resulta que

$$pvf(v^*xv)v^*p = pf(pxp + s(1-p))p$$

y usando (4.8) resulta

$$pf(pxp + (1-s)p)p \leq pf(x)p.$$

Si p es una proyección de rango finito, definimos $q = p \otimes 1_\infty$ en $\mathcal{H}^\infty$, es decir $q(x_1, x_2, x_3, ...) = (p(x_1), x_2, x_3, ...)$ y definimos $y = x \otimes 1_\infty$, para cualquier operador autoadjunto x con espectro en I. Así q resulta una proyección de rango infinito. Usando lo probado para este tipo de proyecciones, y que $f(a) \otimes 1_\infty = f(a \otimes 1_\infty)$ resulta

$$\begin{aligned} pf(pxp + s(1-p))p \otimes 1_\infty &= qf(qyq + s(1-q))q \\ &\leq qf(y)q = pf(x)p \otimes 1_\infty, \end{aligned}$$

lo que implica que el resultado es válido para proyecciones de rango finito también.

$(4 \Rightarrow 1)$.

Dados x, y operadores autoadjuntos con espectro en I, y $\lambda \in [0,1]$, definimos los siguientes operadores en $\mathcal{H}^2 = \mathcal{H} \times \mathcal{H}$:

$$X = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}, U = \begin{pmatrix} \lambda^{\frac{1}{2}} & (1-\lambda)^{\frac{1}{2}} \\ -(1-\lambda)^{\frac{1}{2}} & \lambda^{\frac{1}{2}} \end{pmatrix}, p = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Entonces por 4 tenemos que si $s \in I$

$$pf(pU^*XUp + s(1-p))p \leq pf(U^*XU)p = pU^*f(X)Up.$$

Pero como

$$U^*XU = \begin{pmatrix} \lambda x + (1-\lambda)y & (\lambda - \lambda^2)^{\frac{1}{2}}(y-x) \\ (\lambda-\lambda^2)^{\frac{1}{2}}(y-x) & \lambda y + (1-\lambda)x \end{pmatrix},$$

resulta

$$\begin{aligned}\begin{pmatrix} f(\lambda x+(1-\lambda)y) & 0 \\ 0 & 0 \end{pmatrix} &= pf(pU^*XUp+s(1-p))p \\ &\leq pU^*\begin{pmatrix} f(x) & 0 \\ 0 & f(y) \end{pmatrix}Up \\ &= \begin{pmatrix} \lambda f(x)+(1-\lambda)f(y) & 0 \\ 0 & 0 \end{pmatrix},\end{aligned}$$

lo que prueba que f es convexa de operadores. □

Como corolario, se obtiene una nueva demostración de alguna de las equivalencias del Teorema 2.3.20.

Corolario 4.3.5. *Sea f es una función continua definida en un intervalo I, con $0 \in I$. Son equivalentes*

1. *f es convexa de operadores y $f(0) \leq 0$.*

2. *Para todo $n \in \mathbb{N}$, se tiene*

$$f(\sum_{i=1}^{n} a_i^* x_i a_i) \leq \sum_{i=1}^{n} a_i^* f(x_i) a_i$$

para toda n-upla $(x_1, ..., x_n)$ de operadores acotados autoadjuntos en $\mathcal{H}$ con espectro en I, y toda n-upla de operadores en $\mathcal{H}$ con $\sum_{i=1}^{n} a_i^* a_i \leq 1$.

Demostración. (1⇒ 2)

Sea a_{n+1} tal que $\sum_{i=1}^{n+1} a_i^* a_i = 1$, y $x_{n+1} = 0$. Luego, por el Teorema 4.3.4,

$$f(\sum_{i=1}^{n} a_i^* x_i a_i) = f(\sum_{i=1}^{n+1} a_i^* x_i a_i) \leq \sum_{i=1}^{n+1} a_i^* f(x_i) a_i \leq \sum_{i=1}^{n} a_i^* f(x_i) a_i,$$

donde la última desigualdad vale pues $f(0) \leq 0$ □

(2⇒ 1)

En particular, tomando $n = 1$ y $a_1 = 0$, resulta $f(0) \leq 0$, (0 es el operador 0). Como el espectro del operador 0 es 0 ($\sigma(0) = 0$), y $\sigma(f(0)) = f(\sigma(0))$ resulta $f(0) \leq 0$ (el número 0). Además f resulta convexa de operadores pues la condición 2) de este corolario es más restrictiva que la condición 2 del Teorema 4.3.4.

4.4. La función perspectiva y sus aplicaciones.

Definición 4.4.1. *Dada una función convexa f definida en un conjunto convexo $K \subseteq \mathbb{R}^n$, definimos la* perspectiva *g en el conjunto $L = \{(x,t) : t > 0, x/t \in K\}$ como $g(x,t) = f(x/t)t$.*

Proposición 4.4.2. *La función $g(x,t)$ es conjuntamente convexa, en el sentido siguiente: si $0 < c < 1$, entonces*

$$g(cx_1 + (1-c)x_2, ct_1 + (1-c)t_2) \leq cg(x_1,t_1) + (1-c)g(x_2,t_2),$$

para (x_1,t_1), $(x_2,t_2) \in L$.

Demostración. Sean $p, q \geq 0$ tal que $p + q = 1$, entonces

$$\begin{aligned}
g(px_1 + qx_2, pt_1 + qt_2) &= f\left(\frac{px_1 + qx_2}{pt_1 + qt_2}\right)(pt_1 + qt_2) \\
&= f\left(\frac{x_1}{t_1}\frac{pt_1}{pt_1 + qt_2} + \frac{x_2}{t_2}\frac{qt_2}{pt_1 + qt_2}\right)(pt_1 + qt_2) \\
&\leq \left[\frac{pt_1}{pt_1 + qt_2} f\left(\frac{x_1}{t_1}\right) + \frac{qt_2}{pt_1 + qt_2} f\left(\frac{x_2}{t_2}\right)\right](pt_1 + qt_2) \\
&= pt_1 f\left(\frac{x_1}{t_1}\right) + qt_2 f\left(\frac{x_2}{t_2}\right) = pg(x_1,t_1) + qg(x_2,t_2).
\end{aligned}$$

En la desigualdad de la prueba usamos la convexidad de f. □

Corolario 4.4.3. *Sean $p = (p_i)$ y $q = (q_i)$ medidas de probabilidad finitas con $p_i, q_i > 0$, entonces la* entropía clásica

$$H(p) = -\sum p_i log(p_i)$$

es cóncava, y la entropía relativa

$$H(q||p) = \sum p_i log(p_i) - p_i log(q_i)$$

es conjuntamente convexa.

Demostración. Como la función $f(x) = x\log(x)$, $f(0) = 0$ es convexa en $[0, \infty)$, se tiene que la entropía es cóncava. La perspectiva de f es $g(x,t) = t\frac{x}{t}\log(\frac{x}{t}) = x\log(x) - x\log(t)$, que es conjuntamente convexa por la Proposición 4.4.2, por lo tanto la entropía relativa es conjuntamente convexa. □

Observación 4.4.4. Dados $L, R \in \mathcal{B}(\mathcal{H})$, definimos $[L, R] = LR - RL$ como el conmutador usual. Sean $L > 0$ y $R > 0$. Si $[L, R] = 0$, o sea, si los operadores conmutan, entonces $[L, p(R)] = 0$ para cualquier polinomio p. Tomando límite resulta que $[L, f(R)] = 0$ para cualquier función continua sobre el espectro de R. En particular, podemos definir sin ambigüedades $\frac{L}{R}$.

Proposición 4.4.5 (Convexidad conjunta de la perspectiva en operadores). *Sea f convexa de operadores. Entonces, restringida a operadores estrictamente positivos que conmutan entre sí, la función perspectiva, $g(L, R) = f\left(\frac{L}{R}\right) R$ es conjuntamente convexa. Es decir que*

$$g(cL_1 + (1-c)L_2, cR_1 + (1-c)R_2) \leq cg(L_1, R_1) + (1-c)g(L_2, R_2),$$

donde $0 \leq c \leq 1$ *y* $[L_i, R_i] = 0$.

Demostración. Sean $R = cR_1 + (1-c)R_2$, $L = cL_1 + (1-c)L_2$. Sean $A = (cR_1)^{\frac{1}{2}} R^{-\frac{1}{2}}$, $B = ((1-c)R_2)^{\frac{1}{2}} R^{-\frac{1}{2}}$. Se verifica que $A^*A + B^*B = 1$. Usando el Teorema 4.3.4 y el hecho de que $f(X)$ conmuta con cualquier operador que conmute con X, para f continua se tiene,

$$\begin{aligned}
g(L, R) &= Rf\left(\frac{L}{R}\right) = R^{\frac{1}{2}} f(R^{-\frac{1}{2}} L R^{-\frac{1}{2}}) R^{\frac{1}{2}} \\
&= R^{\frac{1}{2}} f(A^* \left(\frac{L_1}{R_1}\right) A + B^* \left(\frac{L_2}{R_2}\right) B) R^{\frac{1}{2}} \\
&\leq R^{\frac{1}{2}} (A^* f\left(\frac{L_1}{R_1}\right) A + B^* f\left(\frac{L_2}{R_2}\right) B) R^{\frac{1}{2}} \\
&= (cR_1)^{\frac{1}{2}} f\left(\frac{L_1}{R_1}\right) (cR_1)^{\frac{1}{2}} + ((1-c)R_2)^{\frac{1}{2}} f\left(\frac{L_2}{R_2}\right) ((1-c)R_2)^{\frac{1}{2}} \\
&= cg(L_1, R_1) + (1-c)g(L_2, R_2).
\end{aligned}$$

□

Como corolario de la convexidad conjunta de la perspectiva se obtiene una nueva prueba de la convexidad de la entropía relativa. Otras demostraciones de esto se pueden encontrar en [19], [24] y [25] trabajos G. Lindblad y de M. B. Ruskai, de los cuales el último es un inédito en el que reescribe una demostración más general de Lieb, que se encuentra en [17], tratando de simplificarla.

Corolario 4.4.6 (Convexidad conjunta de la entropía relativa). *La* entropía relativa *$S(\rho||\sigma) = Tr(\rho log(\rho) - \rho log(\sigma))$ es conjuntamente convexa en operadores estrictamente positivos de la forma $I + A$, con A operador de traza autoadjunto.*

Demostración. Sean, $\rho_1, \rho_2, \sigma_1, \sigma_2$ de la forma $0 < I+A$, con A nuclear y autoadjunto. Queremos ver que

$$S(c\rho_1 + (1-c)\rho_2 || c\sigma_1 + (1-c)\sigma_2) \leq cS(\rho_1||\sigma_1) + (1-c)S(\rho_2||\sigma_2).$$

Notemos $\rho_0 = c\rho_1 + (1-c)\rho_2$ y $\sigma_0 = c\sigma_1 + (1-c)\sigma_2$. Veamos que $\log(I+A)$ es de traza. Como A es compacto y autoadjunto existe una b.o.n. $(e_i)_i$ de $\mathcal{H}$ formada por autovectores de A, asociados a finitos o numerables autovalores $(\lambda_n)_n$ que se acumulan en 0. Esos autovectores también lo son de $\log(I+A)$, pero asociados a los autovalores $(\log(1+\lambda_n))_n$. Sea $(e_n)_n \subset (e_i)_i$ la colección formada por los (a lo sumo numerables) autovectores (e_i) asociados a autovalores no nulos. Luego, si existe,

$$tr(\log(I+A)) = \sum_n \langle \log(I+A)e_n, e_n \rangle = \sum_n \log(1+\lambda_n),$$

Usando que $\sum_n \lambda_n$ converge absolutamente si y sólo si $\sum_n \log(1+\lambda_n)$ también lo hace (con $\lambda_n > -1$, ver, por ejemplo, [5], Capítulo VII, Proposición 5.4) y que A es de traza, se tiene que $\log(I+A)$ es de traza. Del razonamiento anterior se deduce que existen b.o.n. (e_i^k), (f_i^k), con $k = 0, 1, 2$ de $\mathcal{H}$ formadas por autovectores de $\log \sigma_0, \log \sigma_1, \log \sigma_2, \log \rho_0, \log \rho_1$ y $\log \rho_2$ respectivamente. Llamemos (e_n^k) y (f_n^k) a los subconjunto finitos o numerables de (e_i^k) y (f_n^k) respectivamente formados por los autovectores asociados a los autovalores no nulos, que notaremos (λ_n^k) y (γ_n^k).

Sean, para $i = 1, 2$, $L_i(X) = \rho_i X$ y $R_i(X) = X\sigma_i$ operadores sobre el espacio de Hilbert $\mathcal{B}_2(\mathcal{H})$. Claramente R_i conmuta con L_i y son operadores positivos. Como $x\log(x)$ es convexa de operadores positivos (Proposición 2.3.25), su función perspectiva es conjuntamente convexa sobre operadores positivos. En particular, notando $L = cL_1 + (1-c)L_2$, $R = cR_1 + (1-c)R_2$, se tiene, para todo operador $p \in \mathcal{B}_2(\mathcal{H})$,

$$\langle g(L,R)p, p \rangle \leq c\langle g(L_1, R_1)p, p \rangle + c\langle g(L_2, R_2)p, p \rangle$$

es decir

$$\begin{aligned} tr(\rho_0 \log(\rho_0)p - \rho_0 p \log(\sigma_0)p) &\leq ctr(\rho_1 \log(\rho_1)p - \rho_1 p \log(\sigma_1)p) \\ &+ (1-c)tr(\rho_2 \log(\rho_2)p - \rho_2 p \log(\sigma_2)p). \end{aligned}$$

Veamos que podemos tomar p_n, de manera tal que, cuando n tiende a infinito,

$$tr(\rho_k \log(\rho_k)p_n - \rho_k p_n \log(\sigma_k)p_n) \to tr(\rho_k \log \rho_k - \rho_k \log \sigma_k)$$

para todo $k = 0, 1, 2$, lo que terminaría la prueba. Para eso, por la linealidad de la traza, basta ver que

$$tr(\rho_k \log \rho_k p_n) \to tr(\rho_k \log \rho_k)$$

y
$$tr(\rho_k p_n \log \sigma_k p_n) \to tr(\rho_k \log \sigma_k).$$
Dado $\epsilon > 0$, tomamos n_0 tal que
$$\sum_{i \geq n_0} |\langle \log \sigma_k e_i^k, e_i^k \rangle| < \epsilon,$$
$$\sum_{i \geq n_0} |\langle \log \rho_k f_i^k, f_i^k \rangle| < \epsilon,$$
para $k = 0, 1, 2$. Tal n_0 existe pues $\log \sigma_k$ y $\log \rho_k$ son de traza. Sea p la proyección al espacio
$$span\{e_n^k, f_n^k, \rho_k \log \sigma_k e_n^k : k = 0, 1, 2;\ n \leq n_0\}.$$
Tenemos,
$$\begin{aligned} |tr(\rho_k \log \sigma_k - p\rho p \log \sigma_k)| &\leq |\sum_{n=0}^{n_0} \langle (\rho_k \log \sigma_k - p\rho_k p \log \sigma_k) e_n^k, e_n^k \rangle| \\ &+ |\sum_{n > n_0} \langle (\rho_k \log \sigma_k - p\rho_k p \log \sigma_k) e_n^k, e_n^k \rangle|. \end{aligned}$$
Por la definición de p, y recordando que e_n^k es autovector de $\log \sigma_k$, el primer sumando es cero. Acotemos el segundo:
$$\begin{aligned} |\sum_{n > n_0} \langle (\rho_k \log \sigma_k - p\rho_k p \log \sigma_k) e_n^k, e_n^k \rangle| &\leq \sum_{n > n_o} \|\rho_k\| \|\log \sigma_k e_n^k\| \|e_n^k\| \\ &+ \|p\| \|\rho_k\| \|p\| \|\log \sigma_k e_n^k\| \|e_n^k\| \\ &\leq 2\|\rho_k\| \sum_{n > n_0} |\lambda_n^k| \leq 2\|\rho_k\| \epsilon. \end{aligned}$$
De manera similar, pero usando las bases (f_i^k), se prueba que $tr(\rho_k \log \rho_k p_n)$ tiende a $tr(\rho_k \log \rho_k)$. □

Corolario 4.4.7. *Si $0 < s < 1$, $K \in \mathcal{B}_2(\mathcal{H})$, la función $F(A, B) = Tr(A^s K^* B^{1-s} K)$ es conjuntamente cóncava en operadores positivos.*

Demostración. Probemos primero el corolario para operadores estrictamente positivos. Como $f(t) = -t^s$ es convexa de operadores (Lema 2.3.24), se tiene que $g(L, R) = -L^s R^{1-s}$ es conjuntamente convexa para operadores estrictamente positivos que conmutan. Tomando $L(X) = AX$ y $R(X) = XB$, se tiene
$$\langle g(L, R)(K^*), K^* \rangle = -Tr(A^s K^* B^{1-s} K) = -F(A, B)$$

es conjuntamente convexa por la Proposición 4.4.5. Notemos que como $\mathcal{B}_2(\mathcal{H})$ es un ideal, A^sK^* y $B^{1-s}K$ están en $\mathcal{B}_2(\mathcal{H})$; luego $A^sK^*B^{1-s}K$ es un operador de traza. Veamos el caso general, es decir, tomando operadores positivos pero no necesariamente inversibles. Dados A un operador positivo y $\epsilon > 0$, $A+\epsilon I$ resulta un operador estrictamente positivo, luego, por el caso ya probado tenemos que la aplicación

$$(A+\epsilon I, B+\epsilon I) \mapsto F(A+\epsilon I, B+\epsilon I)$$

es conjuntamente cóncava sobre operadores positivos. Veamos que

$$\lim_{\epsilon\to 0} F(A+\epsilon I, B+\epsilon I) = F(A,B),$$

lo cual termina la demostración:

$$\begin{aligned}
|F(A+\epsilon I, B+\epsilon I) - F(A,B)| &\leq |F(A+\epsilon I, B+\epsilon I) - F(A, B+\epsilon I)| \\
&+ |F(A, B+\epsilon I) - F(A,B)| \\
&= |tr(((A+\epsilon I)^s - A^s)K^*(B+\epsilon I)^{1-s}K)| \\
&+ |tr(((B+\epsilon I)^{1-s} - B^{1-s})KA^sK^*)| \\
&\leq \|(A+\epsilon I)^s - A^s\| tr(|K^*(B+\epsilon I)^{1-s}K|) \\
&+ \|(B+\epsilon I)^{1-s} - B^{1-s}\| tr(|KA^sK^*|) \to 0,
\end{aligned}$$

donde la última desigualdad vale por el Lema 4.2.14.

□

Definición 4.4.8. *Dadas f,h funciones continuas, con h estrictamente positiva, y L,R operadores positivos que conmutan entre sí, definimos la* perspectiva de Marechal *como*

$$(f\Delta h)(L,R) = f\left(\frac{L}{h(R)}\right) h(R).$$

Teorema 4.4.9. *Sea f convexa de operadores, $f(0)\leq 0$, y h cóncava de operadores y estrictamente positiva, entonces*

$$(L,R) \mapsto (f\Delta h)(L,R)$$

es conjuntamente convexa en operadores positivos L,R que conmutan.

Demostración. Sea $L = cL_1 + (1-c)L_2$ y $R = cR_1 + (1-c)R_2$ con $[L_i, R_i] = 0$ para $i = 1,2$. Entonces $ch(R_1)+(1-c)h(R_2) \leq h(R)$, luego tomando $A = c^{\frac{1}{2}}h(R_1)^{\frac{1}{2}}h(R)^{-\frac{1}{2}}$ y $B = (1-c)^{\frac{1}{2}}h(R_2)^{\frac{1}{2}}h(R)^{-\frac{1}{2}}$ se cumple

$$\begin{aligned}
A^*A + B^*B &= h(R)^{-\frac{1}{2}}ch(R_1)h(R)^{\frac{1}{2}} + h(R)^{-\frac{1}{2}}(1-c)h(R_2)h(R)^{-\frac{1}{2}} \\
&\leq h(R)^{-\frac{1}{2}}h(R)h(R)^{-\frac{1}{2}}I = I.
\end{aligned}$$

Por el Corolario 4.3.5 nos queda

$$\begin{aligned}
(f\Delta h)(L,R) &= h(R)^{\frac{1}{2}} f(h(R)^{-\frac{1}{2}} L h(R)^{-\frac{1}{2}}) h(R)^{\frac{1}{2}} \\
&= h(R)^{\frac{1}{2}} f(A^* \left(\frac{L_1}{h(R_1)}\right) A + B^* \left(\frac{L_2}{h(R_2)}\right) B) h(R)^{\frac{1}{2}} \\
&\leq h(R)^{\frac{1}{2}} A^* f\left(\frac{L_1}{h(R_1)}\right) A h(R)^{\frac{1}{2}} + h(R)^{\frac{1}{2}} B^* f\left(\frac{L_2}{h(R_2)}\right) B h(R)^{\frac{1}{2}} \\
&= c h(R_1)^{\frac{1}{2}} f\left(\frac{L_1}{h(R_1)}\right) h(R_1)^{\frac{1}{2}} + (1-c) h(R_2)^{\frac{1}{2}} f\left(\frac{L_2}{h(R_2)}\right) h(R_2)^{\frac{1}{2}} \\
&= c(f\Delta h)(L_1,R_1) + (1-c)(f\Delta h)(L_2,R_2).
\end{aligned}$$

□

Corolario 4.4.10. *Sean $0 < p, q$ tales que $p + q \leq 1$. Sea $K \in \mathcal{B}_2(\mathcal{H})$. Entonces la función $(A,B) \mapsto Tr(A^q K^* B^p K)$ es conjuntamente cóncava en operadores positivos.*

Demostración. Probemos el corolario primero para operadores son estrictamente positivos. Notemos que $p+q$ es una combinación convexa de q y 1 (tomando $c = \frac{p}{1-q}$). Sea $0 \leq t \leq 1$ tal que $p + q = (1-t)q + t1$. Tomando $s = q$, resulta $p = -tq + t = (1-q)t = (1-s)t$, entonces basta probar que $(A,B) \mapsto -Tr(A^s K^* B^{(1-s)t} K)$ es conjuntamente convexa $\forall\ 0 \leq s, t \leq 1$. Recordemos que $f(x) = -x^s$ y $h(y) = y^t$ son respectivamente convexa y cóncava de operadores (Proposición 2.3.24). Si tomamos $L(X) = AX$ y $R(X) = XB$, por el Teorema 4.4.9 resulta que $(f\Delta h)(L,R)$ es conjuntamente convexa. Entonces

$$(A,B) \mapsto \langle (f\Delta h)(L,R)(K^*), K^* \rangle$$

también es conjuntamente convexa, pero

$$\begin{aligned}
\langle [(f\Delta h)(L,R)](K^*), K^* \rangle &= Tr\left(\left[h(R) f\left(\frac{L}{h(R)}\right)\right](K^*)K\right) \\
&= Tr\left(\left[R^t\left(-\left(\frac{L}{R^t}\right)^s\right)\right](K^*)K\right) \\
&= -Tr\left(\left[L^s R^{(1-s)t}\right](K^*)K\right) \\
&= -Tr(A^s K^* B^{(1-s)t} K)
\end{aligned}$$

por lo que el corolario queda demostrado para operadores estrictamente positivos. El caso general, es decir considerando operadores positivos no necesariamente inversibles, se deduce del caso ya demostrado, considerando operadores de la forma $A + \epsilon I$, $B + \epsilon I$, y tomando ϵ tendiendo a 0, de manera análoga al corolario 4.4.7. □

Otra demostración de este corolario, muy técnica en comparación, puede encontrarse en el trabajo de E. H. Lieb [17].

Bibliografía

[1] S. G. Akl, M. Nagy, *Quantum computation and quantum information.* Int. J. Parallel Emergent Distrib. Syst. 21, 2006.

[2] J. Antezana, D. Stojanoff, *Análisis matricial.* Cursos y seminarios de matemática-serie B, departamento de matemática, FCEyN-UBA, Bs. As., 2009.

[3] A. Berthiaume, *Quantum computation.* Complexity theory retrospective, II, 23–51, Springer, New York, 1997.

[4] R. Bhatia, *Matrix analysis.* Graduate Texts in Mathematics, 169. Springer-Verlag, New York, 1997.

[5] J. Conway, *Functions of one complex variable.* Second edition. Graduate Texts in Mathematics, 11. Springer-Verlag, New York-Berlin, 1978.

[6] J. Conway, *A course in functional analysis.* Second edition. Graduate Texts in Mathematics, 96. Springer-Verlag, New York, 1990.

[7] E. Chiumiento, *Ideales Schatten.* Departamento de matemática, UNLP, 2006.

[8] K. R. Davidson, *C^*-algebras by example.* Fields Institute Monographs, 6. American Mathematical Society, Providence, RI, 1996.

[9] J. Dieudonne, *Foundations of modern analysis.* Academic Press, 1960.

[10] E. G. Effros, *A matrix convexity approach to some celebrated quantum inequalities.* Proc. Natl. Acad. Sci. USA 106 (2009), no. 4, 1006–1008.

[11] J. Hadamard, *Sur les caractères de convergence des séries a termes positifs et sur les fonctions indéfiniment croissantes.* (French) Acta Math. 18 (1894), no. 1, 319–336.

[12] F. Hansen, G. K. Pedersen, *Jensen's inequality for operators and Löwner's theorem.* Math. Ann. 258 (1981/82), no. 3, 229–241.

[13] F. Hansen, G. K. Pedersen, *Jensen's operator inequality. Bull. London Math. Soc. 35 (2003), no. 4, 553–564.*

[14] K. Hoffman, R. Kunze, *Linear algebra.* Second edition. Prentice-Hall, Inc., Englewood Cliffs, N.J. 1971.

[15] J. L. W. V. Jensen, *Sur les fonctions convexes et les inégalités entre les valeurs moyennes. (French) Acta Math. 30 (1906), no. 1, 175–193.*

[16] G. Larotonda, *Estructuras geométricas para las variedades de Banach*, 2009.

[17] E. H. Lieb, *Convex trace functions and the Wigner-Yanase-Dyson conjecture.* Advances in Math. 11 (1973), 267–288.

[18] E. H. Lieb, M.B. Ruskai, *Proof of the strong subadditivity of quantum-mechanical entropy.* Con apéndice de B. Simon. J Mathemathical Phys. 14 (1973), 1938–1941.

[19] G. Lindblad, *Entropy, information and quantum measurements.* Comm. Math. Phys. 33 (1973), 305–322.

[20] K. Loewner, *Über monotone Matrixfunktionen.* (German) Math. Z. 38 (1934), no. 1, 177–216.

[21] P. Marechal, *On a functional operation generating convex functions. I. Duality.* J. Optim. Theory Appl. 126 (2005), no.1, 175–189.

[22] P. Marechal, *On a functional operation generating convex functions. II. Algebraic properties.* J. Optim. Theory Appl. 126 (2005), no.2, 357–366.

[23] G. K. Pedersen, *Analysis now.* Graduate Texts in Mathematics, 118. Springer-Verlag, New York, 1989.

[24] M. B. Ruskai, *Another short and elementary proof of strong subadditivity of quantum entropy.* Rep. Math. Phys. 60 (2007), no. 1, 1–12.

[25] M. B. Ruskai, *Lieb's simple proof of concavity of* $(A, B) \mapsto Tr(A^p K^* B(1-p)K)$ *and remarks on related inequalities.* Preprint (2006), arxiv:quant-ph/0404126.

[26] B. Schumacher, *Quantum coding.* Phys. Rev. A (3) 51 (1995), no. 4, 2738–2747.

[27] I.Schur, *Über Gruppen linearer Substitutionen mit Koeffizienten aus einem algebraischen Zahlkörper.* (German) Math. Ann. 71 (1911), no. 3, 355–367.

[28] R. Wheeden, A. Zygmund, *Measure and integral. An introduction to real analysis.* Pure and Applied Mathematics, Vol. 43. Marcel Dekker, Inc., New York-Basel, 1977.