ANÁLISIS FUNCIONAL

E. Llorens Fuster

14 de diciembre de 2016

Índice

1. TEMA 1. ESPACIOS NORMADOS

1.1. Nociones básicas

Si X es un espacio vectorial real o complejo, una norma sobre X es una aplicación que a cada vector v de X asocia un número real $\|v\|$ de modo que, para cualesquiera vectores $v, w \in X$ y para cada escalar λ se verifique

N1) $\|v\| \geq 0$ (positividad).

N2) $\|v\| = 0 \Rightarrow v = 0$ (separación).

N3) $\|\lambda v\| = |\lambda|\|v\|$, (homogeneidad).

N4) $\|v + w\| \leq \|v\| + \|w\|$ (desigualdad triangular).

El valor absoluto en los números reales, o el módulo en los números complejos, son ejemplos de normas en los espacios vectoriales reales $\mathbb{R}$ y $\mathbb{C}$ respectivamente, y también sobre el espacio vectorial complejo $\mathbb{C}$.

De esta definición axiomática de norma se deducen directamente las siguientes consecuencias:

a) $\|0_X\| = 0$, (basta tomar $\lambda = 0$ en (N3)).

b) $\|-v\| = \|v\|$ (tomando $\lambda = -1$ en (N3)).

c) $|\|v\| - \|w\|| \leq \|v - w\|$.

Prueba de (c): Al ser $v = (v - w) + w$, de la desigualdad triangular resulta $\|v\| \leq \|v - w\| + \|w\|$ o sea $\|v\| - \|w\| \leq \|v - w\|$, e intercambiando los papeles de v y de w también queda

$$\|w\| - \|v\| \leq \|w - v\| \stackrel{(b)}{=} \|v - w\|,$$

con lo que se tiene

$$|\|v\| - \|w\|| = \text{máx}\{\|v\| - \|w\|, \|w\| - \|v\|\} \leq \|v - w\|.$$

Se llama espacio normado a un espacio vectorial real o complejo X en el que se ha definido una norma $\|\cdot\|$, lo que suele simbolizarse escribiendo el par $(X, \|\cdot\|)$. A veces no se hace referencia a la norma, si queda clara por el contexto. Naturalmente que dos espacios normados $(X, \|\cdot\|_1)$, $(Y, \|\cdot\|_2)$ son iguales si X coincide con Y en cuanto espacio vectorial, y además son iguales las aplicaciones $\|\cdot\|_1$ $\|\cdot\|_2$.

Ejemplo 1 Todo espacio con producto escalar es un espacio normado con la norma asociada al producto escalar.

Definición 1 *Si* $(X, \|\cdot\|)$ *es un espacio normado y* $\emptyset \neq D \subset X$, *está definida en* D *una distancia* d *de forma natural sin más que tomar, para cualesquiera* x, y *de* D,

$$d(x, y) := \|x - y\|.$$

La métrica d *así definida se llama* asociada a la norma $\|.\|$. *Si es necesario poner énfasis sobre la norma usaremos para esta distancia el símbolo* $d_{\|\cdot\|}$.

Efectivamente d así definida es una métrica sobre D, pues

m1) Para cada $x \in D$, $d(x,x) = \|x - x\| = \|0_X\| = 0$.

m2) Si $d(x,y) = 0$, $\|x - y\| = 0$, y por (N.2) $x = y$.

m3) $d(x,y) = \|x - y\| = |-1|\|x - y\| = \|-x + y\| = d(y,x)$.

m4) $d(x,y) = \|x - y\| = \|x - z + z - y\| \leq \|x - z\| + \|z - y\| = d(x,z) + d(z,y)$.

Esta métrica tiene en X dos propiedades más:

e) Es homogénea, es decir, para cualesquiera $x, y \in X$, $\lambda \in \mathbb{K}$,

$$d(\lambda x, \lambda y) = \|\lambda x - \lambda y\| = |\lambda|\|x - y\| = |\lambda| d(x,y).$$

f) Es invariante por traslaciones, es decir, para cualesquiera $x, y, z \in X$,

$$d(x + z, y + z) := \|(x + z) - (y + z)\| = \|x - y\| = d(x,y).$$

No toda distancia en un espacio vectorial real o complejo puede ser definida a partir de una norma sobre él. (Un ejemplo inmediato es la métrica discreta d en $\mathbb{R}$, pues si x, y son números reales distintos y d estuviera asociada a la norma $\|.\|$

$$1 = d(x/2, y/2) = \|(x/2) - (y/2)\| = \frac{1}{2}\|x - y\| = \frac{1}{2}d(x,y) = 1/2,$$

lo que es absurdo).

Definición 2 *Sea* $(X, \|\cdot\|)$ *un espacio normado; si* D *es cualquier subconjunto no vacío de* X*, entonces* $(D, d_{\|\cdot\|})$ *es un espacio métrico, y por tanto puede considerarse en* D *la topología de esta métrica, que suele llamarse* topología asociada a la norma $\|\cdot\|$.

En particular, una sucesión (v_n) en D converge al elemento $v \in X$ en esta topología de la norma si y sólo si $\|v_n - v\| \to 0$.

Definición 3 *Si el espacio métrico* $(X, d_{\|\cdot\|})$ *es completo se dice que* $(X, \|\cdot\|)$ *es un* espacio de Banach .

Definición 4 *Si* Y *es un subespacio vectorial de* X*, la restricción de la norma* $\|.\|$ *al espacio* Y *es también una norma sobre* Y*. Se dice por esta razón que* $(Y, \|.\|)$ *es un* subespacio normado *de* $(X, \|\cdot\|)$.

Si $(X, \|\cdot\|)$ es un espacio normado, usaremos la siguiente terminología y notación, que son bastante comunes.

Definición 5 *Dados, $v \in X$ y $r > 0$,*

$$B(v,r) := \{x \in X : \|x - v\| < r\}$$

es la bola abierta *de centro v y radio r.*

$$B[v,r] := \{x \in X : \|x - v\| \leq r\}$$

es la bola cerrada *de centro v y radio r. En lugar de $B[0_X, 1]$ se escribe a veces B_X (o simplemente B, si no hay lugar a confusiones).*

$$S[v,r] := \{x \in X : \|x - v\| = r\},$$

es la esfera *de centro v y radio r* [1]

Como todo subconjunto no vacío de un espacio normado es un espacio métrico, y por tanto un espacio topológico (con la topología asociada a la métrica asociada a la norma), en el marco de los espacios normados tienen sentido todas las nociones topológicas, como la clausura e interior de conjuntos, la continuidad y la continuidad uniforme de aplicaciones (definidas entre espacios normados), la convergencia de sucesiones, la compacidad, conexión, completitud, separabilidad, separación etc. etc.

1.1.1. Continuidad de la norma

En particular, si $(X, \|\cdot\|)$ es un espacio normado, la propiedad de la norma que hemos visto antes, a saber que para cualesquiera $v, w \in X$,

$$|\|v\| - \|w\|| \leq \|v - w\|,$$

nos permite ver que la función real $x \mapsto \|x\|$ es siempre continua, pues si (v_n) es cualquier sucesión con $v_n \to v$, entonces

$$0 \leq |\|v_n\| - \|v\|| \overset{(c)}{\leq} \|v_n - v\| \longrightarrow 0.$$

1.1.2. Continuidad de traslaciones y homotecias

Si $(X, \|\cdot\|)$ es un espacio normado y $x \in X$, se llama traslación según x a la aplicación $T_x : X \to X$ dada por $T_x(v) := v + x$.

Es muy sencillo ver que, si $v \neq 0_X$, T_x es continua y más aun es un homeomorfirmo:

Sea (x_n) una sucesión en X convergente a cierto $x_0 \in X$. Entonces $(x_n + x)$ converge a $x_0 + x$, o en otras palabras:

$$T_x(x_n) := x_n + x \to x_0 + x =: T_x(x_0),$$

luego T_x es continua en cualquier x_0 de X. Además, si $X \neq 0_X$, para todo $v \in X$,

$$T_x(T_{-x}(v)) := T_x(v + (-x)) := v - x + x = v$$

[1] En un espacio métrico cualquiera, también pueden definirse bolas y esferas, aunque estas últimas pueden ser vacías. En un espacio normado las esferas nunca son vacías.

es decir que $T_x \circ T_{-x} = Id$. De igual modo se ve que $T_{-x} \circ T_x = Id$. Como tanto T_x como T_{-x} son continuas y biyectivas, hemos probado que T_x es un homeomorfismo de X.

De igual modo, si $(X, \|\cdot\|)$ es un espacio normado y $\lambda \in \mathbb{K}$, se llama homotecia de razón λ a la aplicación $H_\lambda : X \to X$ dada por $H_\lambda(v) := \lambda v$.

Es muy sencillo ver que, si $\lambda \neq 0_X$, H_λ es continua y más aun es un homeomorfirmo:

Sea (x_n) una sucesión en X convergente a cierto $x_0 \in X$. Entonces (λx_n) converge a λx_0, o en otras palabras:

$$H_\lambda(x_n) := \lambda x_n \to \lambda x_0 =: H_\lambda(x_0),$$

luego H_λ es continua en cualquier x_0 de X. Además, si $\lambda \neq 0_X$, para todo $v \in X$,

$$H_\lambda(H_{1/\lambda}(v)) := H_\lambda((1/\lambda)v) := \lambda(1/\lambda)v = v$$

es decir que $H_\lambda \circ H_{1/\lambda} = Id$. De igual modo se ve que $H_{1/\lambda} \circ H_\lambda = Id$. Como tanto H_λ como $H_{1/\lambda}$ son continuas y biyectivas, hemos probado que H_λ es un homeomorfismo de X.

En todo subconjunto no vacío de un espacio normado, en cuanto que es un espacio métrico, tienen también sentido los conceptos métricos, como, por ejemplo:

Definición 6 *Si* $(X, \|\cdot\|)$ *es un espacio normado y* D *un subconjunto no vacío de* X*, se llama diámetro de* D *al elemento de* $[0, \infty]$ *definido por*

$$\mathrm{diam}(D) := \sup\{\|x - y\| : x, y \in D\}.$$

Un conjunto $B \subset X$ *es acotado si es vacío, o si* $\mathrm{diam}(B) < \infty$*.*

Los espacios normados son también espacios vectoriales, y en ese aspecto, en ellos tienen también importancia nociones algebraicas como la dependencia o independencia lineal de vectores, la convexidad, y otras.

1.2. Normas naturales sobre $\mathbb{K}^n$

Si $n \in \mathbb{N}$ consideremos el espacio vectorial real $\mathbb{R}^n$. Si $1 \leq p < \infty$, se definen, para $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$\|x\|_p := (|x_1|^p + \ldots + |x_n|^p)^{\frac{1}{p}}.$$

$$\|x\|_\infty := \max\{|x_1|, \ldots, |x_n|\}.$$

Es relativamente fácil comprobar que se trata de normas sobre $\mathbb{R}^n$. No lo haremos porque veremos más adelante que son subespacios naturales de otros espacios que sí estudiaremos con detalle. El caso particular $p = 2$, es decir $(\mathbb{R}^n, \|.\|_2)$, se llama espacio euclídeo n-dimensional.

De igual modo se pueden definir sobre $\mathbb{C}^n$, y de este modo se obtienen los espacios normados complejos correspondientes.

1.3. Normas del supremo

Sean S un conjunto no vacío, $\mathbb{K} = \mathbb{R}$ ó $\mathbb{C}$, y $\mathbb{K}^S$ el conjunto de las aplicaciones con valores en $\mathbb{K}$ definidas en S. En $\mathbb{K}^S$ puede definirse una estructura natural de espacio vectorial: Si f, g son elementos de este conjunto y $\lambda \in \mathbb{K}$, las aplicaciones

$$f + g : S \to \mathbb{K},\ \lambda f : S \to \mathbb{K}$$

dadas por

$$(f + g)(s) := f(s) + g(s) \ \ (\lambda f)(s) := \lambda f(s)$$

para cada $s \in S$, son elementos de $\mathbb{K}^S$. Es rutinario comprobar que de este modo se tiene en $\mathbb{K}^S$ una estructura de espacio vectorial sobre $\mathbb{K}$.

Un subespacio especialmente importante de $\mathbb{K}^S$ es $\mathcal{B}(S)$, el formado por las aplicaciones acotadas de $\mathbb{K}^S$, es decir

$$\mathcal{B}(S) := \{f \in \mathbb{K}^S : \sup\{|f(s)| : s \in S\} < \infty\}.$$

(Puede observarse que si S es finito, entonces $\mathcal{B}(S) = \mathbb{K}^S$). Si f es un elemento de $\mathcal{B}(S)$, podemos definir el número real no negativo

$$\|f\|_\infty := \sup\{|f(s)| : s \in S\}.$$

Si f, g son elementos de $\mathcal{B}(S)$ y λ un escalar, para cada $s \in S$ se tiene

$$|f(s) + g(s)| \leq |f(s)| + |g(s)| \leq \|f\|_\infty + \|g\|_\infty < \infty$$

$$|(\lambda f)(s)| = |\lambda||f(s)| \leq |\lambda|\|f\|_\infty < \infty$$

lo que permite ver que tanto $f + g$ como λf son elementos de $\mathcal{B}(S)$ y (tomando supremos en las partes izquierdas de las desigualdades), que

$$\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$$

$$\|\lambda f\|_\infty = |\lambda|\|f\|_\infty.$$

Por último, si

$$\|f\|_\infty = \sup\{|f(s)| : s \in S\} = 0$$

entonces f es la aplicación contante nula sobre S, es decir el vector nulo de $\mathbb{K}^S$.

En definitiva hemos probado que $(\mathcal{B}(S), \|.\|_\infty)$ es un espacio normado.

Se trata de un espacio normado que podríamos llamar abstracto, con interesantes casos particulares entre los que podemos destacar:

a) $S = \{1, \ldots, n\}$. En este caso toda aplicación definida sobre S y con valores en $\mathbb{K}$ es acotada, luego $\mathcal{B}(S) = \mathbb{K}^S$. Luego puede identificarse $\mathcal{B}(S)$ con $\mathbb{K}^n$, pues si $x : S \to \mathbb{K}$ es un elemento de $\mathcal{B}(S) = \mathbb{K}^S$ que unívocamente descrito por $(x(1), x(2), \ldots, x(n)) \in \mathbb{K}^n$. Además

$$\|x\|_\infty = \max\{|x(i)| : i = 1, \ldots, n\}.$$

En definitiva, en este caso el espacio normado $(\mathcal{B}(S), \|\cdot\|_\infty)$ es (se identifica con) $(\mathbb{K}^n, \|\cdot\|_\infty)$.

b) $S = \{1, 2, \ldots\} = \mathbb{N}$. En este caso, como toda aplicación definida sobre el conjunto de los números naturales es una sucesión $\mathcal{B}(\mathbb{N})$ no es más que el espacio vectorial de las sucesiones acotadas de elementos de $\mathbb{K}$. Si $x = (x(n))_{n=1}^{\infty}$ es una de estas sucesiones, manteniendo la costumbre de escribir x_n en lugar de $x(n)$, su norma será

$$\|x\|_\infty = \|(x_n)\|_\infty = \sup\{|x_n| : n = 1, 2, \ldots\}.$$

Para el espacio normado $(\mathcal{B}(\mathbb{N}), \|.\|_\infty)$ suele utilizarse el símbolo $(\ell_\infty, \|\cdot\|)_\infty)$ y también m ($m(\mathbb{R})$ ó $m(\mathbb{C})$ si es necesario concretar).

El espacio normado $(\ell_\infty, \|\cdot\|_\infty)$ tiene algunos subespacios de interés:

φ : El subespacio formado por todas las sucesiones *casi nulas* es decir con sólo un número finito de términos distintos de cero.

c_0 : Formado por las sucesiones convergentes a cero en $\mathbb{K}$.

c: Formado por las sucesiones convergentes a algún elemento de $\mathbb{K}$.

Señalemos que toda sucesión de cuadrado sumable es acotada, luego ℓ_2 es subespacio vectorial de ℓ_∞ . La restricción de la norma $\|.\|_\infty$ a este subespacio no coincide con la norma $\|.\|_2$, como se comprueba fácilmente.

c) $S = [a, b]$, $\mathbb{K} = \mathbb{R}$. En tal caso $(\mathcal{B}(S), \|.\|_\infty)$ es el espacio de las funciones reales acotadas definidas sobre el intervalo $[a, b]$, que tiene un subespacio especialmente interesante, el $\mathcal{C}([a, b])$ formado por las funciones continuas en $[a, b]$.

Teorema 1 *Los espacios* $(\mathcal{B}(S), \|.\|_\infty)$ *son completos.*

Prueba : Sea (f_n) una sucesión de Cauchy en $(\mathcal{B}(S), \|.\|_\infty)$. Fijado s en S, de la desigualdad evidente

$$|f_p(s) - f_q(s)| \leq \|f_p - f_q\|_\infty,$$

resulta inmediatamente que la sucesión $(f(s))$ es de Cauchy en $\mathbb{K}$. Como $\mathbb{K}$ es completo existe un número real, que podemos llamar $f(s)$ tal que

$$f(s) = \lim_{n\to\infty} f_n(s).$$

(Es razonable llamar a f función *límite puntual* de la sucesión (f_n), y lo que hemos visto hasta ahora es que por ser completo $\mathbb{K}$, cualquier sucesión de Cauchy en $(\mathcal{B}(S), \|.\|_\infty)$ tiene límite puntual. Falta probar además que esta función límite puntual es un elemento de $\mathcal{B}(S)$ y también que la sucesión dada (f_n) converge hacia f en $(\mathcal{B}(S), \|.\|_\infty)$, es decir que (f_n) converge a f uniformemente en S.

Por ser (f_n) una sucesión de Cauchy, dado $\varepsilon > 0$ existe un natural n_0 tal que si $p, q \geq n_0$, y $s \in S$,

$$|f_p(s) - f_q(s)| \leq \|f_p - f_q\|_\infty < \varepsilon,$$

de donde

$$|f_p(s) - f(s)| = \lim_{q\to\infty} |f_p(s) - f_q(s)| \leq \varepsilon.$$

luego, para todo $s \in S$,

$$-\varepsilon \leq f_p(s) - f(s) \leq \varepsilon$$

o sea

$$f_p(s) - \varepsilon \leq f(s) \leq f_p(s) + \varepsilon$$

lo que nos dice que f es acotada en S (por serlo f_p) y también que

$$\|f_p - f\|_\infty = \sup\{|f_p(s) - f(s)| : s \in S\} \leq \varepsilon.$$

En definitiva hemos probado que, dado $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que si $p \geq n_0$, $\|f - f_p\|_\infty \leq \varepsilon$ o sea que

$$\lim_{p\to\infty} \|f_p - f\|_\infty = 0.$$

Por tanto, hemos visto que toda sucesión de Cauchy (f_n) en el espacio normado $(\mathcal{B}(S), \|.\|_\infty)$ converge a un elemento f de dicho espacio, o en otras palabras, que los espacios $(\mathcal{B}(S), \|.\|_\infty)$ son de Banach.

1.3.1. Los espacios de Banach $(c, \|\cdot\|_\infty)$, $(c_0, \|\cdot\|_\infty)$

Sabemos que $(\ell_\infty, \|\cdot\|_\infty) = (\mathcal{B}(\mathbb{N}), \|\cdot\|_\infty)$ es un espacio de Banach. Como c es un subespacio vectorial de ℓ_∞, para ver que $(c, \|\cdot\|)_\infty$ es un espacio de Banach bastará ver que el conjunto c es cerrado en $(\ell_\infty, \|\cdot\|_\infty)$.

Teorema 2 *c es cerrado en $(\ell_\infty, \|\cdot\|_\infty)$.*

Prueba : Sea pues $(x^{(n)})_{n=1}^\infty$ una sucesión en c con

$$x^{(m)} := (x_1^{(m)}, x_2^{(m)}, \ldots, x_n^{(m)}, \ldots)$$

tal que converge en $(\ell_\infty, \|.\|_\infty)$ a cierto vector $y = (y_1, \ldots, y_n, \ldots) \in \ell_\infty$. Hemos de ver que $y \in c$, es decir que la sucesión (y_n) es convergente en $\mathbb{K}$. Como para cada entero positivo p, $|y_p - x_p^{(m)}| \leq \|y - x^{(m)}\|_\infty$, al ser $(x^{(m)})$ convergente hacia y, se sigue que hay convergencia en la componente p-ésima es decir que $y_p = \lim_{m\to\infty} x_p^{(m)}$. Dado $\varepsilon > 0$ existe m_0 tal que para todo $m \geq m_0$ $\|x^{(m)} - y\|_\infty < \frac{\varepsilon}{3}$. Como $x^{(m_0)} \in c$, la sucesión $(x_p^{(m_0)})_{p=1}^\infty$ es de Cauchy en $\mathbb{K}$, por lo que existirá un natural n_0 de modo que si $p, q \geq n_0$,

$$|x_p^{(m_0)} - x_q^{(m_0)}| \leq \frac{\varepsilon}{3}.$$

Entonces si $p, q \geq n_0$,

$$\begin{aligned} |y_p - y_q| &\leq |y_p - x_p^{(m_0)}| + |x_p^{(m_0)} - x_q^{(m_0)}| + |x_q^{(m_0)} - y_q| \\ &\leq \|y - x^{(m_0)}\|_\infty + |x_p^{(m_0)} - x_q^{(m_0)}| + \|x^{(m_0)} - y\|_\infty \\ &< \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3} \end{aligned}$$

lo que prueba que la sucesión (y_n) es de Cauchy en $\mathbb{K}$. Como $\mathbb{K}$ es completo, (y_n) es convergente, es decir $y = (y_n) \in c$.

□

Corolario 1 *$(c, \|\cdot\|_\infty)$ es un espacio de Banach.*

TEOREMA 3 *El espacio vectorial c_0 es cerrado en $(\ell_\infty, \|\cdot\|_\infty)$ (y por tanto en $(c, \|.\|_\infty)$).*

Prueba : Sea pues $(x^{(n)})_{n=1}^\infty$ una sucesión de elementos de c_0 con

$$x^{(m)} := (x_1^{(m)}, x_2^{(m)}, \ldots, x_n^{(m)}, \ldots)$$

y tal que converge en $(\ell_\infty, \|.\|_\infty)$ a cierto vector $y = (y_1, \ldots, y_n, \ldots) \in \ell_\infty$. Hemos de ver que $y \in c_0$, es decir que la sucesión (y_n) es convergente a 0 en $\mathbb{K}$.

Dado $\varepsilon > 0$ existe m_0 tal que si $m \geq m_0$

$$\|x^{(m)} - y\|_\infty < \frac{\varepsilon}{2}.$$

Como $x^{(m_0)} \in c_0$, dado el mismo $\varepsilon > 0$ existe n_0 tal que si $n \geq n_0$, $|x_n^{(m_0)}| < \frac{\varepsilon}{2}$.

Entonces, si $n \geq \max\{n_0, m_0\}$,

$$|y_n| \leq |y_n - x_n^{(m_0)}| + |x_n^{(m_0)}| \leq \|y - x^{(m_0)}\|_\infty + |x_n^{(m_0)}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

1.3.2. El espacio de Banach $(\mathcal{C}(I), \|\cdot\|_\infty)$

Sea $I = [a, b]$ un intervalo compacto (no trivial) de la recta real. El espacio de Banach de las funciones (reales o complejas) acotadas $(\mathcal{B}(I), \|.\|_\infty)$ tiene un subespacio vectorial importante que es el formado por aquellas funciones que además son continuas, conjunto al que simbolizaremos con $\mathcal{C}(I)$.

Una sucesión (f_n) de funciones acotadas $f_n : I \to \mathbb{K}$ verifica que $\|f_n - f\|_\infty \to 0$ para cierta $f \in \mathcal{B}(I)$ si y sólo si (f_n) converge *uniformemente* a f en I.

De los cursos elementales de Análisis Matemático es sabido que si cada f_n es continua en I, es uniformemente continua.

Dado $\varepsilon > 0$, existe un entero positivo n_0 tal que si $n \geq n_0$, $\|f_n - f\|_\infty < \frac{\varepsilon}{3}$.

Como f_{n_0} es uniformemente continua en I, dado el mismo $\varepsilon > 0$, existe $\delta > 0$ tal que si $t, t' \in I$ con $|t - t'| < \delta$, entonces $|f_{n_0}(t) - f_{n_0}(t')| < \frac{\varepsilon}{3}$

Por tanto, si $t, t' \in I$ con $|t - t'| < \delta$,

$$\begin{aligned} |f(t) - f(t')| &\leq |f(t) - f_{n_0}(t)| + |f_{n_0}(t) - f_{n_0}(t')| + |f_{n_0}(t') - f(t')| \\ &\leq \|f - f_{n_0}\|_\infty + |f_{n_0}(t) - f_{n_0}(t')| + \|f_{n_0} - f\|_\infty \\ &< \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3} \\ &= \varepsilon \end{aligned}$$

Luego, f es uniformemente continua en I. En otras palabras, $\mathcal{C}(I)$ es cerrado en $(\mathcal{B}(I), \|.\|_\infty)$. Por tanto $(\mathcal{C}(I), \|\cdot\|_\infty)$ es un espacio de Banach.

$\square$

En esta prueba no se ha usado para nada el hecho de que I sea un intervalo. Sí, en cambio, que I sea un conjunto compacto.

1.3.3. El espacio de Banach $(\mathcal{C}^1([a,b]), \|\cdot\|_\infty)$

Se define como

$$\mathcal{C}^1([a,b]) = \{f \in C([a,b]) : \forall t \in [a,b]\, \exists f'(t), \text{ y } f' \in C([a,b])\}.$$

En este espacio es habitual considerar la siguiente norma:

$$\|f\| := \|f\|_\infty + \|f'\|_\infty.$$

Comprobar que ciertamente se trata de una norma sobre $\mathcal{C}^1([a,b])$ es inmediato. Vemos a ver que:

Teorema 4 *$(\mathcal{C}^1([a,b]), \|\cdot\|)$ es un espacio de Banach.*

Prueba : Si (f_n) es una sucesión de Cauchy en $(\mathcal{C}^1([a,b]), \|\cdot\|)$ se ve inmediatamente que (f_n) y (f'_n) también lo son en en $(\mathcal{C}([a,b]), \|\cdot\|_\infty)$, luego existe $f \in (\mathcal{C}([a,b]), \|\cdot\|)$ tal que

$$\|f_n - f\|_\infty \to 0,$$

y existe $h \in (\mathcal{C}([a,b]), \|\cdot\|)$ tal que

$$\|f'_n - h\|_\infty \to 0.$$

Naturalmente, existe $y_0 = \lim_n f'_n(a)$ Por el teorema fundamental del Cálculo, la expresión

$$\varphi(x) := y_0 + \int_a^x h(t)dt$$

define una función derivable en $[a,b]$, con derivada continua h en este intervalo.

Afirmamos que $f \in (\mathcal{C}^1([a,b])$ y que $f' = h$. Por ser las f'_n continuas y por el teorema fundamental del Cálculo

$$f_n(x) = f_n(a) + \int_a^x f'_n(t)dt$$

se tiene que

$$|\varphi(t) - f_n(t)| = |y_0 + \int_a^t h(s)ds - f_n(a) - \int_a^t f'_n(s)ds| \leq |y_0 - f_n(a)| + |\int_a^t (f'_n(s) - h(s))ds|$$
$$\leq |y_0 - f_n(a)| + (b-a)\|f'_n - h\|_\infty \to 0.$$

De aquí se sigue que la función φ es límite puntual de la sucesión (f_n), y por unicidad de límite $\varphi(x) = f(x)$ para todo $x \in [a,b]$. Por tanto, para todo $x \in [a,b]$,

$$f(x) = y_0 + \int_a^x h(t)dt$$

lo que nos dice que $f \in (\mathcal{C}^1([a,b])$, y que $f' = h$. Entonces,

$$\|f_n - f\| := \|f_n - f\|_\infty + \|f'_n - f'\|_\infty = \|f_n - f\|_\infty + \|f'_n - h\|_\infty \to 0,$$

luego (f_n) converge a f en $(\mathcal{C}^1([a,b]), \|\cdot\|_\infty)$.

$\square$

Ejemplo 2 El el espacio $\mathcal{C}^1([-1,1])$ definimos para $n > 2$, la sucesión (f_n) por

$$f_n(x) = \frac{x}{1 + n^2 x^2}.$$

Se tiene que converge uniformemente hacia la función idénticamente nula, $\Theta(x) \equiv 0$, pero en el punto $x = 0$, las derivadas $f'_n(0) = 1$ no convergen hacia $\Theta'(0) = 0$.

1.4. Los espacios de sucesiones ℓ_p

Como caso particular de espacio con producto escalar ya hemos tenido ocasión de conocer el espacio de Hilbert ℓ_2 de las sucesiones de números reales de cuadrado sumable, que puede ser considerado la más inmediata extension infinito-dimensional del espacio euclídeo $\mathbb{R}^n$. También hemos visto el espacio de sucesiones ℓ_∞ , que extiende al $(\mathbb{R}^n, \|.\|_\infty)$. Parece natural que exista un análogo de dimensión infinita para los espacios $(\mathbb{R}^n, \|.\|_p)$. Tales son los espacios de sucesiones ℓ_p .

Así pues si $1 \leq p < \infty$, se definen los espacios vectoriales reales o complejos

$$\ell_p := \{(x_n) \in \mathbb{K}^{\mathbb{N}} : \sum_{n=1}^{\infty} |x_n|^p < \infty\}.$$

Si $x = (x_n) \in \ell_p$ y $\lambda \in \mathbb{K}$,

$$\sum_{n=1}^{\infty} |\lambda x_n|^p = |\lambda|^p \sum_{n=1}^{\infty} |x_n|^p < \infty,$$

luego $\lambda x \in \ell_p$.

Además, si $x = (x_n), y = (y_n)$ son elementos de ℓ_p, como

$$|x_n + y_n|^p \leq [|x_n| + |y_n|]^p \leq [2 \max\{|x_n|, |y_n|\}]^p = 2^p \max\{|x_n|^p, |y_n|^p\} \leq 2^p[|x_n|^p + |y_n|^p]$$

resulta

$$\sum_{n=1}^{\infty} |x_n + y_n| \leq 2^p \sum_{n=1}^{\infty} [|x_n|^p + |y_n|^p] = 2^p[\sum_{n=1}^{\infty} |x_n|^p + \sum_{n=1}^{\infty} |y_n|^p] < \infty,$$

luego $x + y \in \ell_p$, y por tanto ℓ_p es un subespacio vectorial de c_0. (Recordemos que si $\sum_{n=1}^{\infty} |x_n|^p < \infty$, entonces $|x_n|^p \to 0$, luego $x_n \to 0$, es decir $x \in c_0$).

Veremos a continuación que la expresión

$$\|(x_n)\|_p := \left[\sum_{n=1}^{\infty} |x_n|^p\right]^{\frac{1}{p}}$$

define una norma que es la standard sobre ℓ_p. Para ello se necesitan los siguientes lemas:

Lema 1 *(Desigualdad de Young[2]) Sean a y b números reales no negativos, $p > 1$ y q el conjugado de p, (es decir $\frac{1}{p} + \frac{1}{q} = 1$). Entonces se verifica la relación*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

[2]William Henry Young, (Londres, 1863 – Lausanne, 1942) fue un matemático inglés.)

Prueba : (1) La desigualdad es trivial si $a = 0$ ó $b = 0$. Si $a, b > 0$ y ponemos $\lambda = \frac{1}{p} \in (0,1)$ entonces $1 - \lambda = 1 - \frac{1}{p} = \frac{1}{q}$. Como la función logarítmica es cóncava,

$$\log(\lambda a^p + (1-\lambda)b^q) \geq \lambda \log(a^p) + (1-\lambda)\log(b^q) = p\lambda \log(a) + q(1-\lambda)\log(b) = \log(ab)$$

luego

$$e^{\log(\lambda a^p + (1-\lambda)b^q)} \geq e^{\log(ab)}$$

es decir

$$\frac{a^p}{p} + \frac{b^q}{q} \leq ab.$$

Prueba : (2) Para $t > 0$ consideremos la función real

$$g(t) := \frac{1}{p}t + \frac{1}{q} - t^{\frac{1}{p}}.$$

Como

$$g'(t) := \frac{1}{p} - \frac{1}{p}t^{\frac{1}{p}-1},$$

entonces la ecuación $g'(t) = 0$ tiene como única solución $t_0 = 1$.

Dado que $g''(1) = -\frac{1}{p}\left(\frac{1}{p} - 1\right) = \frac{1}{pq} > 0$, este valor $t_0 = 1$ es el único mínimo de g en $(0, \infty)$.

Si $b > 0$, entonces $0 = g(1) \leq g\left(\frac{a^p}{b^q}\right)$, es decir

$$0 \leq \frac{1}{p}\left(\frac{a^p}{b^q}\right) + \frac{1}{q} - \left(\frac{a^p}{b^q}\right)^{\frac{1}{p}}.$$

Multiplicando por b^q queda,

$$0 \leq \frac{a^p}{p} + \frac{b^q}{q} - \frac{a}{b^{\frac{q}{p}-q}},$$

es decir,

$$ab^{q-\frac{q}{p}} \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Como p y q son conjugados, $q - \frac{q}{p} = \frac{qp-q}{p} = 1$, la desigualdad anterior es

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q},$$

que también se cumple (trivialmente) si $b = 0$, lo que completa la prueba.

□

TEOREMA 5 *(Desigualdad de Hölder para sucesiones). Sean $p, q > 1$ números conjugados. Si $x \in \ell_p$ e $y \in \ell_q$ entonces*

$$\sum_{n=1}^{\infty} |x_n y_n| \leq \|x\|_p \|y\|_q.$$

Prueba : Si $\|x\|_p = 0$ entonces $x_n = 0$ para todo natural n, y la desigualdad se cumple de manera evidente. Lo mismo ocurre si $\|y\|_q = 0$. Si $\|x\|_p \neq 0$ y $\|y\|_q \neq 0$ entonces, para cada $n \in \mathbb{N}$, es

$$\frac{|x_n|}{\|x\|_p} \geq 0, \quad \frac{|y_n|}{\|y\|_q} \geq 0.$$

Por el lema anterior,

$$\frac{|x_n|}{\|x\|_p}\frac{|y_n|}{\|y\|_q} \leq \frac{1}{p}\left(\frac{|x_n|}{\|x\|_p}\right)^p + \frac{1}{q}\left(\frac{|y_n|}{\|y\|_q}\right)^q.$$

Como $x \in \ell_p$ y $g \in \ell_q$, la serie que tiene por término general la expresión definida por el segundo miembro de la desigualdad anterior es convergente, luego también lo es la que define el primer miembro. Por monotonía

$$\sum_{n=1}^{\infty} \frac{|x_n|}{\|x\|_p}\frac{|y_n|}{\|y\|_q} \leq \sum_{n=1}^{\infty}\left[\frac{1}{p}\left(\frac{|x_n|}{\|x\|_p}\right)^p + \frac{1}{q}\left(\frac{|y_n|}{\|y\|_q}\right)^q\right],$$

es decir

$$\frac{1}{\|x\|_p\|y\|_q}\sum_{n=1}^{\infty}|x_n y_n| \leq \frac{1}{p}\frac{1}{\|x\|_p^p}\sum_{n=1}^{\infty}|x_n|^p + \frac{1}{q}\frac{1}{\|y\|_q^q}\sum_{n=1}^{\infty}|y_n|^q = \frac{1}{p} + \frac{1}{q} = 1$$

de donde resulta inmediatamente la desigualdad del enunciado.

□

Más arriba hemos visto que si $x, y \in \ell_p$ se cumple que

$$\begin{aligned} \textstyle\sum_{n=1}^{\infty}|x_n + y_n|^p &\leq 2^p \textstyle\sum_{n=1}^{\infty}[|x_n|^p + |y_n|^p] \\ &= 2^p[\textstyle\sum_{n=1}^{\infty}|x_n|^p + \sum_{n=1}^{\infty}|y_n|^p] < \infty, \end{aligned}$$

lo que también puede escribirse

$$\|x + y\|_p \leq 2[\|x\|_p^p + \|y\|_p^p]^{\frac{1}{p}}.$$

La llamada desigualdad de Minkowski mejora la anterior:

Teorema 6 *(Desigualdad de Minkowski en ℓ_p). Sea $p \in [1, \infty)$. Si $x, y \in \ell_p$, entonces*

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

Prueba : El caso $p = 1$ es consecuencia inmediata de que, para cada $n \in \mathbb{N}$,

$$|x_n + y_n| \leq |x_n| + |y_n|.$$

Sea pues $1 < p < \infty$, y sea q el conjugado de p. Supondremos que $\|x + y\|_p > 0$, pues si $\|x + y\|_p = 0$ la desigualdad se cumple trivialmente.

Para cada $n \in \mathbb{N}$,

$$|x_n + y_n|^p \leq (|x_n| + |y_n|)|x_n + y_n|^{p-1} = |x_n||x_n + y_n|^{p-1} + |y_n||x_n + y_n|^{p-1}$$

Consideremos la sucesión (h_n) dada por

$$h_n := |x_n + y_n|^{p-1}.$$

Como

$$|h_n|^q = [|x_n + y_n|^{p-1}]^q = |x_n + y_n|^{pq-q} = |x_n + y_n|^p$$

y $x + y \in \ell_p$, se sigue que $h \in \ell_q$, con

$$\|h\|_q^q = \sum_{n=1}^{\infty} |h_n|^q = \sum_{n=1}^{\infty} |x_n + y_n|^p = \|x + y\|_p^p.$$

De aquí que $\|h\|_q = \|x + y\|_p^{\frac{p}{q}}$.

Como $x \in \ell_p$ y $h \in \ell_q$, por la desigualdad de Hölder

$$\sum_{n=1}^{\infty} |x_n h_n| \leq \|x\|_p \|h\|_q = \|x\|_p \|x + y\|_p^{\frac{p}{q}}. \tag{1}$$

Del mismo modo, como $y \in \ell_p$ y $h \in \ell_q$, por la desigualdad de Hölder

$$\sum_{n=1}^{\infty} |y_n h_n| \leq \|y\|_p \|h\|_q = \|y\|_p \|x + y\|_p^{\frac{p}{q}}. \tag{2}$$

Según hemos visto más arriba, para cada $n \in \mathbb{N}$ es

$$|x_n + y_n|^p \leq |x_n||x_n + y_n|^{p-1} + |y_n||x_n + y_n|^{p-1} = |x_n h_n| + |y_n h_n|.$$

Por monotonía, y teniendo en cuenta (1) y (2) se cumplirá que

$$\begin{aligned} \textstyle\sum_{n=1}^{\infty} |x_n + y_n|^p &\leq \textstyle\sum_{n=1}^{\infty} |x_n h_n| + \sum_{n=1}^{\infty} |y_n h_n| \\ &\leq \|x\|_p \|x + y\|_p^{\frac{p}{q}} + \|y\|_p \|x + y\|_p^{\frac{p}{q}} \\ &= [\|x\|_p + \|y\|_p] \left(\|x + y\|_p^{\frac{p}{q}} \right). \end{aligned}$$

Por lo tanto,

$$\|x + y\|_p^p \leq [\|x\|_p + \|y\|_p] \left(\|x + y\|_p^{\frac{p}{q}} \right),$$

o sea

$$\|x + y\|_p^{p - \frac{p}{q}} \leq \|x\|_p + \|y\|_p.$$

Al ser p y q conjugados $p - \frac{p}{q} = \frac{pq-p}{q} = 1$, luego hemos probado la desigualdad del enunciado.

Los espacios $(\ell_p, \|.\|_p)$, $(1 \leq p < \infty)$, como espacios de Banach Sea $(x^{(n)})_{n=1}^{\infty}$ una sucesión de Cauchy en ℓ_p con

$$x^{(m)} := (x_1^{(m)}, x_2^{(m)}, \ldots, x_n^{(m)}, \ldots).$$

Al ser, para cualesquiera enteros positivos m, q, i,

$$|x_i^{(m)} - x_i^{(q)}| = \left(|x_i^{(m)} - x_i^{(q)}|^p\right)^{\frac{1}{p}} \leq \|x^{(m)} - x^{(q)}\|_p,$$

se deduce, para cada $i \in \mathbb{N}$, que la sucesión de números reales o complejos $(x_i^{(n)})_n$ es de Cauchy, y tendrá por tanto un límite que será un número (real o complejo, respectivamente), digamos y_i.

Sea $y := (y_i)$. Vamos a ver que $y \in \ell_p$ y que $\lim_{n\to\infty} \|x^n - y\|_p = 0$. Por definición de sucesión de Cauchy, dado $\varepsilon > 0$, existirá un natural n tal que si $m, q \geq n$, entonces

$$\|x^{(m)} - x^{(q)}\|_p < \varepsilon$$

es decir

$$\sum_{i=1}^{\infty} |x_i^{(m)} - x_i^{(q)}|^p < \varepsilon^p,$$

de donde para cualquier suma parcial

$$\sum_{i=1}^{k} |x_i^{(m)} - x_i^{(q)}|^p < \varepsilon^p,$$

y de aquí se sigue que

$$\lim_{q\to\infty} \left[\sum_{i=1}^{k} |x_i^{(m)} - x_i^{(q)}|^p\right] \leq \varepsilon^p,$$

o sea

$$\sum_{i=1}^{k} |x_i^{(m)} - y_i|^p \leq \varepsilon^p. \tag{3}$$

Por tanto la serie $\sum_{i=1}^{\infty} |x_i^{(m)} - y_i|^p$ es convergente y con suma menor o igual que ε^p . A partir de aquí resulta que $x^{(m)} - y$ es un elemento de ℓ_p, por lo que también pertenecerá a ℓ_p el vector

$$y = y - x^{(m)} + x^{(m)}.$$

Además también se sigue de (3)

$$\|x^{(m)} - y\|_p \leq \varepsilon$$

siempre que $m \geq n$, lo que demuestra (siendo $\varepsilon > 0$ arbitrario), que $\|x^{(m)} - y\|_p \longrightarrow 0$, es decir

$$x^{(m)} \to y.$$

en $(\ell_p, \|\cdot\|)$. En definitiva hemos probado que toda sucesión de Cauchy en ℓ_p converge a un elemento de ℓ_p, el cual es, por otra parte, límite de la sucesión dada *componente a componente*.

$\square$

1.5. Series en un espacio normado

Como ejemplo de la interacción entre la estructura vectorial y la estructura topológica puede servir la noción de serie, que tiene sentido en cualquier espacio normado. Si (x_n) es una sucesión en X, y $(X, \|\cdot\|)$ es un espacio normado, en él pueden considerarse los procesos de sumar y tener límite (converger). Esto motiva la siguiente:

DEFINICIÓN 7 *Si* $(X, \|\cdot\|)$ *es un espacio normado y* (x_n) *es una sucesión de elementos de* X*, y la sucesión* (v_n) *dada por* $v_k := x_1 + \ldots + x_k = \sum_{n=1}^{k} x_n$ *converge en* $(X, \|\cdot\|)$ *a cierto vector* $v \in X$*, a dicho vector se le denota por* $\sum_{n=1}^{\infty} x_n$*, y se le llama* suma de la serie $\sum_{n=1}^{\infty} x_n$.

Es decir que

$$v = \sum_{n=1}^{\infty} x_n := \lim_{k\to\infty} \sum_{n=1}^{k} x_n$$

y debe entenderse que, esto es equivalente a que

$$\lim_{k\to\infty} \left\| v - \sum_{n=1}^{k} x_n \right\| = 0.$$

Ejemplo 3 *En* ℓ_p *consideremos la sucesión de vectores* $(e^{(n)})$ *definida por*

$$e^{(n)} := (0, \overset{n-1}{\ldots}, 0, 1, 0 \ldots).$$

Vamos a ver que, para cada $v = (v_n) \in \ell_p$,

$$v = \sum_{n=1}^{\infty} v_n e^{(n)}.$$

En efecto,

$$\|v - \sum_{n=1}^{k} v_n e^{(n)}\|_p^p = \|v - (v_1, \ldots, v_k, 0 \ldots)\|_p^p = \|(0, \ldots, 0, v_{k+1}, v_{k+2}, \ldots)\|_p^p = \sum_{k+1}^{\infty} |v_n|^p$$

Como $v \in \ell_p$, $\sum_{n=1}^{\infty} |v_n|^p < \infty$; por una propiedad bien conocida de las series numéricas,

$$\lim_{k\to\infty} \sum_{n=k+1}^{\infty} |v_n|^p = 0$$

luego

$$\lim_{k\to\infty} \|v - \sum_{n=1}^{k} v_n e^{(n)}\|_p = 0,$$

es decir,

$$\sum_{n=1}^{\infty} v_n e^{(n)} = v.$$

Observar que, en cambio, si $v = (1, 1, 1, \ldots) \in c$ se tiene que

$$\left\| v - \sum_{n=1}^{k} v_n e^{(n)} \right\|_\infty = \|(0, 0, \ldots, 0, 1, 1, \ldots)\|_\infty = 1$$

es decir, que $v \neq \sum_{n=1}^{\infty} v_n e^{(n)}$. Por tanto, en c **es falso** que $x = \sum_{k=1}^{\infty} x_k e^{(k)}$ se cumpla para todo $x = (x_n) \in c$.

A veces puede escribirse el símbolo $\sum_{n=1}^{\infty} x_n$ cuando (x_n) es una sucesión en $(X, \|\cdot\|)$, sin establecer si existe en este espacio el límite de las sumas parciales $\sum_{n=1}^{k} x_n$. Si esto se hace así, es preferible llamar a la expresión $\sum_{n=1}^{\infty} x_n$ serie formal asociada a la sucesión (x_n).

En ese caso se dice que $\sum_{n=1}^{\infty} x_n$ es convergente si existe como serie, es decir si existe en $(X, \|\cdot\|)$ el límite de las sumas parciales, según acabamos de ver.

Definición 8 *Una serie (formal)* $\sum_{n=1}^{\infty} x_n$ *en* $(X, \|\cdot\|)$ *se dice* absolutamente convergente *si la serie numérica*

$$\sum_{n=1}^{\infty} \|x_n\|$$

es convergente.

Teorema 7 *Sea* $(X, \|\cdot\|)$ *un espacio de Banach . Sea* $\sum_{n=1}^{\infty} x_n$ *una serie (formal) en* X*. Si la serie* $\sum_{n=1}^{\infty} x_n$ *es absolutamente convergente en* $(X, \|\cdot\|)$ *, entonces es convergente en* $(X, \|\cdot\|)$

Prueba : Sean σ_k y s_k las respectivas sumas parciales de orden k de las series del enunciado, es decir:

$$\sigma_k := \sum_{n=1}^{k} x_n, \qquad s_k := \sum_{n=1}^{k} \|x_n\|.$$

Si la serie numérica $\sum_{n=1}^{\infty} \|x_n\|$ es convergente, entonces la sucesión de sumas parciales (s_k) es de Cauchy en $\mathbb{R}$. Dado $\varepsilon > 0$ existe un entero positivo n_0 tal que si $n_0 \leq p < p$, $s_q - s_p < \varepsilon$.

Como para cualesquiera enteros positivos $p < q$,

$$\|\sigma_q - \sigma_\|\| = \left\| \sum_{n=1}^{q} x_n - \sum_{n=1}^{p} x_n \right\| \leq \sum_{n=p+1}^{q} \|x_n\| = s_q - s_p < \ varepsilon,$$

se tendrá que la sucesión (σ_k) es de Cauchy en $(X, \|\cdot\|)$. Al ser $(X, \|\cdot\|)$ un espacio de Banach, esta sucesión de sumas parciales $(\sum_{n=1}^{k} x_n)_{k\geq 1}$ tendrá un límite $x \in X$, que por definición será la suma de la serie $\sum_{n=1}^{\infty} x_n$.

$\square$

Es hasta cierto punto sorprendente que también se cumple una especie de recíproco del teorema anterior:

Teorema 8 *Sea* $(X, \|\cdot\|)$ *un espacio normado en el cual toda serie absolutamente convergente es convergente. Entonces* $(X, \|\cdot\|)$ *es un espacio de Banach .*

Prueba : Sea (x_n) una sucesión de Cauchy en $(X, \|\cdot\|)$.

Dado $\varepsilon = \frac{1}{2}$ existe $n_0 \in \mathbb{N}$ tal que, si $p, q \geq n_0$ $\|x_p - x_q\| < \frac{1}{2}$.

Dado $\varepsilon = \frac{1}{2^2}$ existe $n_1 \in \mathbb{N}$, que puede elegirse $n_1 > n_0$ de modo que, si $p, q \geq n_1$ $\|x_p - x_q\| < \frac{1}{2^2}$.

En particular, $\|x_{n_1} - x_{n_0}\| < \frac{1}{2}$.

Dado $\varepsilon = \frac{1}{2^3}$ existe $n_2 \in \mathbb{N}$, que puede elegirse $n_2 > n_2$ de modo que, si $p, q \geq n_2$ $\|x_p - x_q\| < \frac{1}{2^3}$.

En particular, $\|x_{n_2} - x_{n_1}\| < \frac{1}{2^2}$.

Continuando de este modo, obtenemos por inducción una subsucesión (x_{n_k}) de la sucesión original (x_n) de manera que

$$\|x_{n_k} - x_{n_{k-1}}\| < \frac{1}{2^k}.$$

Como

$$\sum_{k=1}^{\infty} \|x_{n_k} - x_{n_{k-1}}\| \leq \sum_{k=1}^{\infty} \frac{1}{2^k} < \infty,$$

es decir como la serie numérica $\sum_{k=1}^{\infty} \|x_{n_k} - x_{n_{k-1}}\|$ es convergente, la serie $\sum_{k=1}^{\infty}(x_{n_k} - x_{n_{k-1}})$ es convergente en $(X, \|\cdot\|)$, es decir, existe $x \in X$ tal que

$$x = \lim_{m\to\infty} \sum_{k=1}^{m}(x_{n_k} - x_{n_{k-1}}).$$

Pero,

$$\sum_{k=1}^{m}(x_{n_k} - x_{n_{k-1}}) = (x_{n_1} - x_{n_0}) + (x_{n_2} - x_{n_1}) + \ldots + (x_{n_m} - x_{n_{m-1}}) = x_{n_m} - x_{n_0},$$

luego

$$x = \lim_{m\to\infty} \sum_{k=1}^{m}(x_{n_k} - x_{n_{k-1}}) = \lim_{m\to\infty}(x_{n_m} - x_{n_0}).$$

Por tanto existe

$$\lim_{m\to\infty} x_{n_m} = x + x_{n_0}.$$

Recordamos ahora una conocida propiedad de las sucesiones de Cauchy en espacios métricos: Si una sucesión de Cauchy (x_n) tiene una subsucesión convergente, ella misma es convergente.

Luego existe en X el vector $\lim_n x_n$, o en otras palabras, existe $v \in X$tal que

$$\lim_{n\to\infty} \|x_n - v\| = 0.$$

□

1.6. Los espacios de Lebesgue $L^p(\Omega)$

1.6.1. Definición

Sea Ω un subconjunto (Lebesgue) medible de $\mathbb{R}^n$, con medida positiva.

Si $1 < p < \infty$ consideraremos los siguientes conjuntos de funciones

$$\mathcal{L}^p(\Omega, \mathbb{R}) := \left\{ f : \Omega \to \mathbb{R} : f \text{es medible y} \int_\Omega |f(t)|^p dt < \infty \right\}.$$

$$\mathcal{L}^p(\Omega, \mathbb{C}) := \left\{ f : \Omega \to \mathbb{C} : f \text{es medible y} \int_\Omega |f(t)|^p dt < \infty \right\}.$$

Si no es necesario distinguir entre $\mathcal{L}^p(\Omega, \mathbb{R})$ y $\mathcal{L}^p(\Omega, \mathbb{C})$ escribiremos simplemente $\mathcal{L}^p(\Omega)$.

Para cada $f \in \mathcal{L}^p(\Omega)$ escribiremos

$$\|f\|_p := \left(\int_\Omega |f(t)|^p dt \right)^{\frac{1}{p}}.$$

Es importante tener en cuenta que, a pesar de la notación usada, la aplicación $f \mapsto \|f\|_p$ no es una norma sobre $\mathcal{L}^p(\Omega)$, puesto que de ser $\|f\|_p = 0$, es decir, de ser

$$\int_\Omega |f(t)|^p dt = 0,$$

no se sigue que f es la función contante cero (sobre Ω), sino sólo que $f(t) = 0$ *para casi todo punto* $t \in \Omega$.)

Naturalmente, siempre podemos considerar que una función que es nula para casi todo punto de Ω, es *como si fuera* la función nula. Si hacemos esto, en realidad estamos definiendo en $\mathcal{L}^(\Omega)$ la relación binaria de equivalencia

$$f_1 \sim f_2 \Leftrightarrow m[\{x \in \Omega : f(x) \neq g(x)\}] = 0$$

donde m es la medida de Lebesgue de $\mathbb{R}^n$. Si establecemos esta equivalencia de funciones, estamos considerando como iguales, a dos que son diferentes a lo sumo en los valores que toman en un conjunto de medida nula. Por ejemplo, si $\Omega = \mathbb{R}^1$, la función idénticamente cero θ y la función característica de los racionales $\chi_\mathbb{Q}$, serán equivalentes según $\sim$.

Esta relación binaria de equivalencia da lugar a una partición de $\mathcal{L}^p(\Omega)$ en clases. Concretamente, dada cualquier $f \in \mathcal{L}^p(\Omega)$, la clase de equivalencia $[f]$ a la que pertenece f viene dada por

$$[f] = \{g \in \mathcal{L}^p(\Omega) : g \sim f\} = \{g \in \mathcal{L}^p(\Omega) : g(t) = f(t) \text{ p.c.t. } t \in \Omega\}.$$

El conjunto formado por estas clases de equivalencia, es decir el conjunto cociente $\mathcal{L}^p(\Omega)/\sim$ se llama espacio de Lebesgue $L^p(\Omega)$. (También se usan los símbolos $L^p(\Omega, \mathbb{R})$ y $L^p(\Omega, \mathbb{C})$ si es necesario especificar si se trata de (clases de) funciones reales o complejas.

Normalmente no se distingue entre un elemento de $L^p(\Omega)$ y uno de sus representantes, es decir que no se hace distinción entre funciones (elementos de $\mathcal{L}^p(\Omega)$) y clases (de equivalencia) (elementos de $L^p(\Omega)$), aunque no deben confundirse. (Del mismo modo en que pocas veces hay que distinguir entre las fracciones $\frac{1}{3}$ o $\frac{3}{9}$ -que consideramos *equivalentes*, es decir como la misma a casi todos los efectos-, y el número racional $\frac{1}{3}$).

Observemos que si f_1, f_2 son funciones de $\mathcal{L}^p(\Omega)$, entonces, como la integral de una función no cambia si la modificamos en los puntos de un conjunto de medida nula,

$$\|f_1\|_p := \left(\int_\Omega |f_1(t)|^p dt\right)^{\frac{1}{p}} = \left(\int_\Omega |f_2(t)|^p dt\right)^{\frac{1}{p}} =: \|f_2\|_p.$$

Esto hace que sea consistente **definir**, para $[f] \in L^p(\Omega)$

$$\|[f]\|_p := \left(\int_\Omega |f(t)|^p dt\right)^{\frac{1}{p}},$$

es decir que el valor del número real $\|[f]\|_p$ no depende de la función elegida para calcular la integral del segundo miembro, es decir no depende del representante elegido para la clase $[f]$.

Siguiendo una costumbre general, no usaremos en adelante la notación habitual para las clases de equivalencia, como $[f]$, para referirnos a los elementos de $L^p(\Omega)$. De este modo, expresiones como $f \in L^p(\Omega)$, deben entenderse en realidad como $[f] \in L^p(\Omega)$. Por ejemplo, la definición anterior quedaría: Si $f \in L^p(\Omega)$,

$$\|f\| := \left(\int_\Omega |f(t)|^p dt\right)^{\frac{1}{p}},$$

(y debe entenderse que la misma letra 'f' se usa como símbolo de una clase de equivalencia en la parte izquierda de la igualdad, y de una función (de dicha clase) en la parte derecha).

Así pués, (teniendo presente este convenio de notaciones) si $f \in L^p(\Omega)$ y $\|f\|_p = 0$, entonces

$$\int_\Omega |f(t)|^p dt = 0,$$

y si la integral (en Ω) de una función no negativa es 0 sabemos que esta función es nula para casi todo punto de Ω. Esto que implica que, *como función*, $f \sim \theta \Leftrightarrow [f] = [\theta]$, (donde θ es la función constante 0 sobre Ω). Pero si no distinguimos entre funciones y clases, entonces concluimos simplemente que $f = 0 \in L^p(\Omega)$.

1.6.2. Relación entre $L^{p_1}(\Omega)$ y $L^{p_2}(\Omega)$

PROPOSICIÓN 1 *Si la medida de Lebesgue $m(\Omega)$ es finita entonces si $1 \leq p_1 \leq p_2 < \infty$, se tiene que*

$$L^{p_2}(\Omega) \subset L^{p_1}(\Omega).$$

Prueba : Tomemos $f \in L^{p_2}(\Omega)$. Consideremos los conjuntos medibles y disjuntos

$$A := \{x \in \Omega : |f(x)| \leq 1\} \;\; B := \{x \in \Omega : |f(x)| > 1\}.$$

Entonces

$$\begin{aligned}\int_\Omega |f(t)|^{p_1}dt &= \int_A |f(t)|^{p_1}dt + \int_B |f(t)|^{p_1}dt \\ &\leq \int_A 1dt + \int_B |f(t)|^{p_1}dt \\ &= m(A) + \int_B |f(t)|^{p_1}dt \\ &\leq m(A) + \int_B |f(t)|^{p_2}dt \\ &\leq m(A) + \int_\Omega |f(t)|^{p_2}dt < \infty.\end{aligned}$$

Por tanto, todo elemento de $\mathcal{L}^{p_2}(\Omega)$ pertenece a $\mathcal{L}^{p_1}(\Omega)$, lo que completa la prueba.

$\square$

Por otra parte, si Ω es un conjunto de medida de Lebesgue no finita esta inclusión no se mantiene: La función $h : [1,\infty) \to \mathbb{R}$ dada por $h(t) = \frac{1}{t}$ pertenece a $\mathcal{L}^2([1,\infty))$, pues

$$\int_{[1,\infty)} \left|\frac{1}{t}\right|^2 dt = 1 < \infty,$$

mientras que $\int_{[1,\infty)} \left|\frac{1}{t}\right| dt$ no es un número real, luego $h \notin \mathcal{L}^1([1,\infty))$.

1.6.3. Estructura vectorial

Es claro que la suma de dos funciones medibles reales o complejas definidas sobre Ω, es también una función medible (real o compleja respectivamente. También es claro que el producto de un escalar real por una función real medible sobre Ω es también una función medible sobre Ω. Esto hace del conjunto de las funciones reales medibles sobre Ω un espacio vectorial real, y del mismo modo el conjunto de las funciones complejas medibles sobre Ω tiene una estructura natural de espacio vectorial sobre el cuerpo de los complejos.

Veremos a continuación que $\mathcal{L}^p(\Omega,\mathbb{K})$ (donde $\mathbb{K} = \mathbb{R}$ õ $\mathbb{C}$ indistintamente) es un subespacio de este espacio vectorial.

Efectivamente, sean $f, g \in \mathcal{L}^p(\Omega,\mathbb{K})$. Para cada $t \in \Omega$,

$$\begin{aligned}|f(t)+g(t)|^p &\leq [|f(t)| + |g(t)|]^p \\ &= [2\,\text{máx}\{|f(t)|,|g(t)|\}]^p \\ &= 2^p[\text{máx}\{|f(t)|,|g(t)|\}]^p \\ &= 2^p[\text{máx}\{|f(t)|^p,|g(t)|^p\}] \\ &\leq 2^p[|f(t)|^p + |g(t)|^p].\end{aligned}$$

Como $\int_\Omega |f(t)|^p dt < \infty$ y $\int_\Omega |g(t)|^p dt < \infty$, por monotonía de la integral, se sigue inmediatamente que $\int_\Omega [|f(t)+g(t)|^p]dt < \infty$:

$$\begin{aligned}\int_\Omega |f(t)+g(t)|^p dt &\leq 2^p \int_\Omega [|f(t)|^p + |g(t)|^p]dt \\ &= 2^p[\int_\Omega |f(t)|^p dt + \int_\Omega |g(t)|^p dt] < \infty.\end{aligned}$$

También $L^p(\Omega)$ tiene estructura natural de espacio vectorial. Si $f, g \in \mathcal{L}^p(\Omega)$ podemos definir correctamente la suma de clases de equivalencia, y el producto de un escalar (real o complejo) por una clase, por las fórmulas

$$[f] + [g] = [f + g]$$
$$\lambda[f] = [\lambda f]$$

donde $\lambda \in \mathbb{R}$ ó $\lambda \in \mathbb{C}$ según los casos.

Estas definiciones son correctas porque no dependen del representante elegido para cada una de las clases, es decir:

Si $[f_1] = [f_2], \in L^p(\Omega)$, entonces

$$m[\{t \in \Omega : f_1(t) \neq f_2(t)\}] = 0.$$

Análogamente, si $[g_1] = [g_2], \in L^p(\Omega)$, entonces

$$m[\{t \in \Omega : g_1(t) \neq g_2(t)\}] = 0.$$

Pero entonces, como

$$\{t \in \Omega : f_1(t) + g_1(t) \neq f_2(t) + g_2(t)\} \subset \{t \in \Omega : f_1(t) \neq f_2(t)\} \cup \{t \in \Omega : g_1(t) \neq g_2(t)\},$$

se sigue que

$$\begin{aligned} m[\{t \in \Omega : f_1(t) + g_1(t) \neq f_2(t) + g_2(t)\}] \leq \\ \leq m[\{t \in \Omega : f_1(t) \neq f_2(t)\} \cup \{t \in \Omega : g_1(t) \neq g_2(t)\}] \leq \\ \leq m[\{t \in \Omega : f_1(t) \neq f_2(t)\}] + [\{t \in \Omega : g_1(t) \neq g_2(t)\}] = 0. \end{aligned}$$

Luego $f_1 + g_1 \sim f_2 + g_2$, o sea $[f_1 + g_1] = [f_2 + g_2]$.

De modo parecido podemos probar la consistencia de la fórmula $\lambda[f] = [\lambda f]$.

A partir de aquí es de rutina demostrar que $L^p(\Omega)$ es un espacio vectorial sobre $\mathbb{R}$ o sobre C, según sea el caso.

1.6.4. Desigualdad de Hölder para espacios de funciones.

Teorema 9 *(Desigualdad de Hölder). Sean $p, q > 1$ números conjugados. Si $f \in \mathcal{L}^p(\Omega)$ y $g \in \mathcal{L}^q(\Omega)$ entonces $fg \in \mathcal{L}^1(\Omega)$, y además*

$$\|fg\|_1 := \int_\Omega |f(t)g(t)|dt \leq \|f\|_p \|g\|_q.$$

Prueba : Si $\|f\|_p = 0$ entonces $f(t) = 0$ para casi todo t en Ω y la desigualdad se cumple de manera evidente. Lo mismo ocurre si $\|g\|_q = 0$. Si $\|f\|_p \neq 0$ y $\|g\|_q \neq 0$ entonces, para cada $t \in \Omega$, es

$$\frac{|f(t)|}{\|f\|_p} \geq 0, \quad \frac{|g(t)|}{\|g\|_q} \geq 0.$$

Por el lema anterior,

$$\frac{|f(t)|}{\|f\|_p}\frac{|g(t)|}{\|g\|_q} \leq \frac{1}{p}\left(\frac{|f(t)|}{\|f\|_p}\right)^p + \frac{1}{q}\left(\frac{|g(t)|}{\|g\|_q}\right)^q.$$

Como $f \in \mathcal{L}^p(\Omega)$ y $g \in \mathcal{L}^q(\Omega)$, la función definida por el segundo miembro de la desigualdad anterior es de $\mathcal{L}^1(\Omega)$, luego también lo es la que define el primer miembro. Por la monotonía de la integral se tendrá

$$\int_\Omega \frac{|f(t)|}{\|f\|_p}\frac{|g(t)|}{\|g\|_q}dt \leq \int_\Omega \left[\frac{1}{p}\left(\frac{|f(t)|}{\|f\|_p}\right)^p + \frac{1}{q}\left(\frac{|g(t)|}{\|g\|_q}\right)^q\right] dt,$$

es decir

$$\frac{1}{\|f\|_p\|g\|_q}\int_\Omega |f(t)g(t)|dt \leq \frac{1}{p}\frac{1}{\|f\|_p^p}\int_\Omega |f(t)|^p dt + \frac{1}{q}\frac{1}{\|g\|_q^q}\int_\Omega |g(t)|^q dt = \frac{1}{p} + \frac{1}{q} = 1,$$

de donde resulta inmediatamente la desigualdad del enunciado.

$\square$

1.6.5. Desigualdad de Minkowski

Más arriba hemos visto que si $f, g \in \mathcal{L}^p(\Omega)$ se cumple que

$$\begin{aligned}\textstyle\int_\Omega |f(t)+g(t)|^p dt &\leq 2^p \textstyle\int_\Omega [|f(t)|^p + |g(t)|^p]dt \\ &= 2^p[\textstyle\int_\Omega |f(t)|^p dt + \int_\Omega |g(t)|^p dt] < \infty,\end{aligned}$$

lo que también puede escribirse

$$\|f+g\|_p \leq 2[\|f\|_p^p + \|g\|_p^p]^{\frac{1}{p}}.$$

La llamada desigualdad de Minkowski mejora la anterior:

Teorema 10 *(Desigualdad de Minkowski). Sea $p \in [1, \infty)$. Si $f, g \in \mathcal{L}^p(\Omega)$, entonces*

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

Prueba : El caso $p = 1$ es consecuencia inmediata de que, para cada $t \in \Omega$,

$$|f(t)+g(t) \leq |f(t)| + |g(t)|.$$

Sea pues $1 < p < \infty$, y sea q el conjugado de p. Supondremos que $\|f+g\|_p > 0$, pues si $\|f+g\|_p = 0$ la desigualdad se cumple trivialmente.

Para cada $t \in \Omega$,

$$|(f+g)(t)|^p \leq (|f(t)|+|g(t)|)|f(t)+g(t)|^{p-1} = |f(t)||f(t)+g(t)|^{p-1} + |g(t)||f(t)+g(t)|^{p-1}$$

Consideremos la función

$$h(t) := |f(t)+g(t)|^{p-1}.$$

Como

$$|h(t)|^q = [|f(t) + g(t)|^{p-1}]^q = |f(t) + g(t)|^{pq-q} = |f(t) + g(t)|^p$$

y $f + g \in \mathcal{L}^p(\Omega)$, se sigue que $h \in \mathcal{L}^q(\Omega)$, con

$$\|h\|_q^q = \int_\Omega |h(t)|^q dt = \int_\Omega |f(t) + g(t)|^p dt = \|f + g\|_p^p.$$

De aquí que $\|h\|_q = \|f + g\|_p^{\frac{p}{q}}$.

Como $f \in \mathcal{L}^p(\Omega)$ y $h \in \mathcal{L}^q(\Omega)$, por la desigualdad de Hölder $fh \in L^1(\Omega)$ con

$$\int_\Omega |f(t)h(t)|dt \leq \|f\|_p \|h\|_q = \|f\|_p \|f + g\|_p^{\frac{p}{q}}. \tag{4}$$

Del mismo modo, como $g \in \mathcal{L}^p(\Omega)$ y $h \in \mathcal{L}^q(\Omega)$, por la desigualdad de Hölder

$$\int_\Omega |g(t)h(t)|dt \leq \|f\|_p \|h\|_q = \|g\|_p \|f + g\|_p^{\frac{p}{q}}. \tag{5}$$

Según hemos visto más arriba, para cada $t \in \Omega$ es

$$|(f + g)(t)|^p \leq |f(t)||f(t) + g(t)|^{p-1} + |g(t)||f(t) + g(t)|^{p-1} = |f(t)h(t)| + |g(t)h(t)|.$$

Por monotonía y aditividad de la integral, y teniendo en cuenta (4) y (5) se cumplirá que

$$\begin{aligned} \textstyle\int_\Omega |(f + g)(t)|^p dt &\leq \textstyle\int_\Omega |f(t)h(t)|dt + \int_\Omega |g(t)h(t)|dt \\ &\leq \|f\|_p \|f + g\|_p^{\frac{p}{q}} + \|g\|_p \|f + g\|_p^{\frac{p}{q}} \\ &= [\|f\|_p + \|g\|_p] \left(\|f + g\|_p^{\frac{p}{q}} \right). \end{aligned}$$

Por lo tanto,

$$\|f + g\|_p^p \leq [\|f\|_p + \|g\|_p] \left(\|f + g\|_p^{\frac{p}{q}} \right),$$

o sea

$$\|f + g\|_p^{p - \frac{p}{q}} \leq \|f\|_p + \|g\|_p.$$

Al ser p y q conjugados $p - \frac{p}{q} = \frac{pq-p}{q} = 1$, luego hemos probado la desigualdad del enunciado.

□

1.6.6. El espacio de Banach $(L^p(\Omega), \|\cdot\|_p)$

Vamos a comprobar que el espacio normado $(L^p(\Omega), \|\cdot\|_p)$ es completo. Para verlo basta con demostrar que toda serie absolutamente convergente en este espacio, es convergente, y luego aplicar el Teorema 8.

Teorema 11 *Sea (f_n) una sucesión en $L^p(\Omega)$ tal que la serie $\sum_{n=1}^\infty \|f_n\|_p$ es convergente. Entonces existe $f \in L^p(\Omega)$ tal que*

$$\sum_{n=1}^\infty f_n = f,$$

es decir,

$$\lim_{k\to\infty}\left\|f-\sum_{n=1}^{k}f_n\right\|_p=0.$$

Prueba : Formamos la sucesión (g_k) de elementos de $L^p(\Omega)$ dados, para cada entero positivo k, por

$$g_k:=\sum_{n=1}^{k}|f_n|.$$

Para cada entero positivo k, $g_k^p=|g_k|^p\in L^1(\Omega)$. Además para cada $t\in\Omega$,

$$g_k(t)\leq g_{k+1}(t)\Rightarrow[g_k(t)]^p\leq[g_{k+1}(t)]^p,$$

es decir, para cada t de Ω la sucesión de números reales no negativos $([g_k(t)]^p)_k$ es monótona creciente.

Por otra parte, por la desigualdad triangular y por la hipótesis del teorema,

$$\begin{aligned}(\textstyle\int_\Omega[g_k(t)]^p dt)^{\frac{1}{p}} &= \|g_k\|_p\\ &\leq \textstyle\sum_{n=1}^{k}\|f_n\|_p\\ &\leq \textstyle\sum_{n=1}^{\infty}\|f_n\|_p<\infty.\end{aligned}$$

Luego la sucesión $(g_k^p)_k$ es no negativa y monótona creciente en $L^1(\Omega)$, y sus integrales en Ω tienen una cota superior común:

$$\int_\Omega[g_k(t)]^p dt\leq\left[\sum_{n=1}^{\infty}\|f_n\|_p\right]^p. \tag{6}$$

Una de las muchas versiones del teorema de la convergencia monótona de Dini nos garantiza entonces que existe $g\in L^1(\Omega)$ tal que

(a) Para casi todo $t\in\Omega$, $g_k(t)^p\longrightarrow g(t)$.

(b) $\int_\Omega g_k(t)^p dt\longrightarrow\int_\Omega g(t)dt$.

De (a) se sigue que $g(t)\geq 0$ para casi todo t de Ω, por lo que está definido (c.p.t. $t\in\Omega$) $g^{\frac{1}{p}}(t)\in[0,\infty]$.

De (b) y de (6) se sigue que

$$\int_\Omega g(t)dx\leq\left[\sum_{n=1}^{\infty}\|f_n\|_p\right]^p<\infty,$$

luego $g(t)$ es un número real no negativo para casi todo punto t de Ω.

Como $[g^{\frac{1}{p}}]^p=g\in L^1(\Omega)$, tenemos que $g^{\frac{1}{p}}\in L^p(\Omega)$.

De (a) se sigue también que para casi todo $t\in\Omega$,

$$g_k(t)\longrightarrow[g(t)]^{\frac{1}{p}}$$

o, en otras palabras,

$$\sum_{n=1}^{k} |f_n(t)| \overset{k\to\infty}{\longrightarrow} [g(t)]^{\frac{1}{p}}.$$

Esto nos dice que la serie de números reales (o complejos)

$$\sum_{n=1}^{\infty} f_n(t)$$

es absolutamente convergente (y por tanto convergente) para casi todo $t \in \Omega$. Definamos pues para $t \in \Omega$

$$f(t) := \sum_{n=1}^{\infty} f_n(t).$$

(lo que tendrá sentido y dará un número real o complejo, según acabamos de ver, para casi todo $t \in \Omega$, y completaremos, como es habitual por $f(t) := 0$ en los puntos de Ω donde la serie anterior no converja).

Vamos a ver que $f \in L^p(\Omega)$ y que $\lim_{k\to\infty} \left\| f - \sum_{n=1}^{k} f_n \right\|_p = 0$.

En efecto, para casi todo $t \in \Omega$,

$$\begin{aligned} |f(t)| &= \left| \lim_{k\to\infty} \textstyle\sum_{n=1}^{k} f_n(t) \right| \\ &= \lim_{k\to\infty} \left| \textstyle\sum_{n=1}^{k} f_n(t) \right| \\ &\leq \lim_{k\to\infty} \textstyle\sum_{n=1}^{k} |f_n(t)| \\ &= \lim_{k\to\infty} g_k(t) \\ &= [g(t)]^{\frac{1}{p}}, \end{aligned}$$

luego para casi todo $t \in \Omega$, $|f(t)|^p \leq g(t)$. Como $g \in \mathcal{L}^1(\Omega)$, tenemos que $f \in L^p(\Omega)$.

Para ver que (f_n) converge a f en $(L^p(\Omega), \|\cdot\|_p)$, aplicaremos el Teorema de la convergencia dominada a la sucesión de funciones (h_k) con

$$h_k := \left| f - \sum_{n=1}^{k} f_n \right|^p.$$

Para casi todo $t \in \Omega$,

$$\begin{aligned} \left| f(t) - \textstyle\sum_{n=1}^{k} f_n(t) \right|^p &\leq \left[|f(t)| + \left| \textstyle\sum_{n=1}^{k} f_n(t) \right| \right]^p \\ &\leq \left[2 \max\{|f(t)|, \left| \textstyle\sum_{n=1}^{k} f_n(t) \right| \} \right]^p \\ &\leq 2^p \left[\max\{|f(t)|, \textstyle\sum_{n=1}^{k} |f_n(t)| \} \right]^p \\ &= 2^p \left[\max\{|f(t)|^p, \left(\textstyle\sum_{n=1}^{k} |f_n(t)| \right)^p \} \right] \\ &= 2^p \max\{|f(t)|^p, (g_k(t))^p\} \\ &\leq 2^p \max\{g(t), (g_k(t))^p\} \\ &= 2^p g(t) \end{aligned}$$

luego para cada entero positivo k, $|h_k|$ está mayorada (dominada) en casi todo punto de Ω por $2^p g \in L^1(\Omega)$.

Podemos pues aplicar a esta sucesión el Teorema de la convergencia dominada y tendremos,

$$\int_\Omega [\lim_k h_k(t)]dt = \lim_k [\int_\Omega h_k(t)dt].$$

es decir,

$$0 = \lim_k \left[\|f - \sum_{n=1}^{k} f_n\|_p^p \right].$$

lo que implica que $0 = \lim_{k\to\infty} \|f - \sum_{n=1}^{k} f_n\|_p$ y completa la prueba.

□

1.6.7. El espacio $L^\infty(\Omega)$

Dada una función real medible $f : \Omega \to \mathbb{R}$, se dice que está **esencialmente acotada** si existe un número real M tal que

$$\mu(\{x \in \Omega : |f(x)| > M\}) = 0.$$

En ese caso M suele llamarse **cota esencial** de f. El conjunto

$$\mathcal{L}^\infty(\Omega) := \{f : \Omega \to \mathbb{R} : f \text{esencialmente acotada}\}$$

es un espacio vectorial. Si $f \in \mathcal{L}^\infty(\Omega)$ llamaremos **supremo esencial** de f al ínfimo de las cotas superiores esenciales. Es decir, el elemento de $[0, \infty)$ dado por

$$\inf\{M > 0 : \mu(\{x \in \Omega : |f(x)| > M\}) = 0\}$$

se llama **supremo esencial** de f, y que simbolizaremos por $\sup\text{ess}(f)$.

Se comprueba fácilmente que $\sup\text{ess}(f)$ es a su vez una cota superior esencial de f, y que $f \mapsto \sup\text{ess}\ (f)$ cumple todas las condiciones para ser una norma sobre $\mathcal{L}^\infty(\Omega)$ excepto la de separación (N2).

Igual que en el caso anterior, se define en este conjunto de funciones esencialmente acotadas la misma relación de equivalencia que ya hemos considerado, y en el espacio vectorial cociente, que suele llamarse espacio de Lebesgue $L^\infty(\Omega)$, la aplicación

$$[f] \mapsto \|[f]\|_\infty := \sup\text{ess}(f)$$

está bien definida y es una norma.

Vamos a comprobar que $(L^\infty(\Omega), \|.\|_\infty)$ es un espacio de Banach.

Sea pues (f_n) una sucesión de Cauchy en $(L^\infty(\Omega), \|.\|_\infty)$. Para cada entero positivo k existe otro, n_k, tal que para $m, n \geq n_k$,

$$\|f_m - f_n\|_\infty \leq \frac{1}{k}.$$

Entonces $\frac{1}{k}$ es una cota esencial de $f_m - f_n$, y por tanto existe un conjunto de medida nula M^k_{mn} tal que pata todo $t \in \Omega \backslash M^k_{mn}$ se verifica

$$|f_m(t) - f_n(t)| \leq \frac{1}{k}.$$

Sea

$$M^k := \bigcup_{m,n \geq n_k} M^k_{mn}.$$

Claramente M^k es de medida nula y para todo $t \in \Omega \backslash M^k$ y para cualesquiera $m, n \geq n_k$, $|f_m(t) - f_n(t)| \leq \frac{1}{k}$. Finalmente si $M = \bigcup_k M^k$ también M es de medida nula y para todo $t \in \Omega \backslash M$ es claro que $(f_n(t))$ es una sucesión de Cauchy en $\mathbb{K}$ que tendrá un límite, digamos $f(t)$. La función f será entonces medible y esencialmente acotada en Ω. Definimos como es habitual $f(t) = 0$ para $t \in M$ y tenemos que (f_n) converge uniformemente a f en $\Omega \backslash M$.

1.7. Espacios normados producto y cociente

Sean $(X, \|.\|)$, $(F, \|.\|)$ espacios normados. En el espacio vectorial $X \times F$ podemos definir las normas

$$\|(x,y)\|_1 := \|x\| + \|y\|,$$

$$\|(x,y)\|_p := (\|x\|^p + \|y\|^p)^{\frac{1}{p}} \quad (1 < p < \infty),$$

$$\|(x,y)\|_\infty := \text{máx}\{\|x\|, \|y\|\},$$

y los correspondientes espacios normados $(X \times F, \|.\|_p)$, $(1 \leq p \leq \infty)$ se llaman espacio ℓ_p -producto de los espacios normados X y F.

De igual modo si damos una sucesión de espacios normados $(X_n, \|.\|_n)$ podemos considerar en el espacio vectorial producto $\Pi_{i=1}^\infty X_i$, los subespacios

$$\ell_p(X_i) := \{(x_i) \in \Pi_{i=1}^\infty X_i : \sum_{i=1}^\infty \|x_i\|_i^p < \infty\}$$

$(1 \leq p < \infty)$

$$\ell_\infty(X_i) := \{(x_i) \in \Pi_{i=1}^\infty X_i : \sup\{\|x_i\|_i : i = 1, 2, \ldots\} < \infty\}.$$

y en ellos las aplicaciones $\|.\|_p$ definidas respectivamente por

$$\|(x_i)\|_p := \left(\sum_{i=1}^\infty \|x_i\|_i^p\right)^{\frac{1}{p}}.$$

$$\|(x_i)\|_\infty := \sup\{\|x_i\|_i : i = 1, 2, \ldots\}.$$

que dan lugar a los espacios normados $(\ell_p(X_i), \| \cdot \|_p)$ $(1 \leq p \leq \infty)$. Es claro que cada X_i está isométricamente contenido en $\ell_p(X_i)$. También es claro que las comprobaciones de todo cuanto se ha afirmado para construir estos espacios son perfectamente paralelas con las hechas en el caso de los espacios ℓ_p (que en realidad son un caso particular de esta construcción, si se considera $X_i = \mathbb{R}$ $(i = 1, 2, \ldots)$). Por esta razón es corriente escribir ℓ_p^n en lugar de $(\mathbb{R}^n, \|.\|_p)$ $(1 \leq p \leq \infty$).

Sea $(X, \|.\|)$ un espacio normado, y sea F un subespacio vectorial cerrado de X. En el espacio vectorial cociente X/F sabemos que la clase de equivalencia del elemento $x \in X$ es $[x] := x + F$, y podemos "definir"

$$\|x + F\| := \text{ínf}\{\|x + v\| : v \in F\}.$$

Es de rutina comprobar que de este modo se ha definido correctamente una aplicación de X/F en $[0, \infty)$, que además es una norma. Al correspondiente espacio normado $(X/F, \|.\|)$ se le llama espacio normado cociente.

1.8. Espacios normados separables

DEFINICIÓN 9 *Un espacio métrico (X, d) se llama* separable *si admite un subconjunto numerable D que además es denso en X, es decir que verifica $cl(D) = X$. Un espacio normado es separable si lo es como espacio métrico con la métrica asociada a la norma.*

El ejemplo fundamental de espacio métrico separable es $\mathbb{R}$ con la métrica usual, que admite como subconjunto numerable denso a $\mathbb{Q}$, el conjunto de los números racionales.

La estructura vectorial de los espacios normados permite dar un teorema que caracteriza cómodamente a los que son separables.

Notación: Si X es un espacio vectorial sobre $\mathbb{K}$ y $M \subset X$ es no vacío, escribiremos $< M >$, o bien $span(M)$ para referirnos al menor subespacio vectorial de X que contiene a M. Es la intersección de todos los subespacios vectoriales que contienen a M.

TEOREMA 12 *(Caractización de los espacios normados separables).*

El espacio normado $(X, \|\cdot\|)$ es separable si y sólo si existe un conjunto numerable

$$\{e^1, e^2, \ldots, e^n, \ldots\} \subset X$$

tal que

$$cl[< \{e^1, e^2, \ldots, e^n, \ldots\} >] = X,$$

donde $< \{e^1, e^2, \ldots, e^n, \ldots\} >$ es el subespacio vectorial de X generado por $\{e^1, e^2, \ldots, e^n, \ldots\}$.

Prueba : Si $(X, \|\cdot\|)$ es separable, existe $D \subset X$ con D numerable y denso en X. En ese caso,

$$X = cl(D) \subset cl(< D >) \subset X,$$

luego poniendo $D = \{e^1, e^2, \ldots, e^n, \ldots\}$ se cumple el enunciado.

Recíprocamente:

Supongamos que $X = cl[< \{e^1, e^2, \ldots, e^n, \ldots\} >]$. Sea

$$D := \{\sum_{i \in F} r_i e^{n_i} : r_i \in \mathbb{Q},\ F \text{ finito } F \subset \mathbb{N}\}.$$

Como $\mathbb{Q}$ es numerable, D es numerable. Vamos a ver que D es denso en X. Efectivamente, sean $x \in X$ y $\varepsilon > 0$. Como $X = cl[< \{e^1, e^2, \ldots, e^n, \ldots\} >]$, existe un vector $y \in < \{e^1, e^2, \ldots, e^n, \ldots\} >$ tal que $d(x,y) < \frac{\varepsilon}{2}$.

Por otra parte, al ser $y \in < \{e^1, e^2, \ldots, e^n, \ldots\} >$, existen $a_{n_1}, \ldots, a_{n_k} \in \mathbb{R}$ tales que

$$y = \sum_{i=1}^{k} a_{n_i} e^{n_i}.$$

Como $\mathbb{Q}$ es denso en $\mathbb{R}$ existen $r_1, \ldots, r_k$ tales que

$$\max_{i=1,\ldots,k} |a_{n_i} - r_i| < \frac{\varepsilon}{2k \max_{i=1,\ldots,k}\{\|e^{n_i}\|\}}.$$

El elemento $z := \sum_{i=1}^{k} r_i e^{n_i}$ pertenece al conjunto D.

$$d(x,z) \leq d(x,y) + d(y,z) < \frac{\varepsilon}{2} + \left\| \sum_{i=1}^{k} (a_{n_i} - r_i) e^{n_i} \right\|$$

$$\leq \frac{\varepsilon}{2} + \sum_{i=1}^{k} |a_{n_i} - r_i| \|e^{n_i}\| < \frac{\varepsilon}{2} + \sum_{i=1}^{k} \frac{\varepsilon}{2k \max_{i=1,\ldots,k}\{\|e^{n_i}\|\}} \|e^{n_i}\| \leq \varepsilon.$$

$\square$

Corolario 2 *Los espacios ℓ_p son separables, para $1 \leq p < \infty$*

Prueba : Observemos primero que en ℓ_p, para todo $x = (x_1, x_2, \ldots)$ se cumple

$$x = \sum_{i=1}^{\infty} x_i e^{(i)} \tag{7}$$

En efecto, fijado un entero positivo k,

$$\left\| \sum_{i=1}^{k} x_i e^{(i)} - x \right\|_p^p = \sum_{i \geq k+1} |x_i|^p$$

Pero como la serie de números reales $\sum_{i=1}^{\infty} |x_i|^p$ es convergente, sabemos que

$$\lim_{k \to \infty} \sum_{i \geq k+1} |x_i|^p = 0$$

luego $\lim_{k\to\infty} \left\| \sum_{i=1}^{k} x_i e^{(i)} - x \right\|_p = 0$, lo que prueba (7). Pero la igualdad (7) nos dice que dado cualquier $\varepsilon > 0$ existe un entero positivo k_0 de manera que

$$\left\| \sum_{i=1}^{k_0} x_i e^{(i)} - x \right\|_p < \varepsilon$$

es decir que dado cualquier $\varepsilon > 0$ existe un elemento $y := \sum_{i=1}^{k_0} x_i e^{(i)}$ de $< B >$, con $B := \{e^{(1)}, e^{(2)}, \ldots\}$, tal que dista de x menos que ε; o en otras palabras, que el conjunto de las combinaciones lineales (finitas) de los elementos del conjunto numerable B es denso en ℓ_p. Entonces, por el teorema anterior, ℓ_p es separable.

$\square$

Ejemplo 4 *Si S es cualquier conjunto infinito, $(\mathcal{B}(S), \|.\|_\infty)$ es no separable.*

Sea A un subconjunto no vacío de S. La funcion característica χ_A (que vale 1 si $t \in A$ y cero en otro caso), es acotada, y por tanto es un elemento de $\mathcal{B}(S)$. Además $\|\chi_A\|_\infty = 1$.

Si A' es cualquier subconjunto no vacío de S, con $A' \neq A$ es fácil ver que

$$\|\chi_A - \chi_{A'}\|_\infty = 1$$

(Los números reales $|\chi_A(t) - \chi_{A'}(t)|$ sólo pueden tomar el valor 0 ó el valor 1. Siendo $A \neq A'$, o bien existe $t \in A'$ con $t \notin A$ en cuyo caso $|\chi_A(t) - \chi'_A(t)| = 1$ o bien existe $t \in A$ con $t \notin A'$ y entonces también $|\chi_A(t) - \chi_{A'}(t)| = 1$. En los dos casos, $\|\chi_A - \chi_{A'}\|_\infty = 1$.)

Por tanto las bolas abiertas $B\left(\chi_A, \frac{1}{2}\right)$, $B\left(\chi_{A'}, \frac{1}{2}\right)$ serán disjuntas siempre que $A \neq A'$.

Si $\mathcal{B}(S)$ fuera separable existiría un conjunto numerable Q denso en $\mathcal{B}(S)$. Por tanto

$$Q \cap B\left(\chi_A, \frac{1}{2}\right) \neq \emptyset.$$

es decir que Q tiene al menos un elemento en $B\left(\chi_A, \frac{1}{2}\right)$. Además, si $A \neq A'$ como las bolas $B\left(\chi_A, \frac{1}{2}\right)$ y $B\left(\chi_{A'}, \frac{1}{2}\right)$ son disjuntas, Q tiene, al menos, dos elementos distintos, uno en cada una de ellas.

Este razonamiento prueba que Q ha de tener al menos tantos elementos como subconjuntos distintos admita el conjunto S.

Pero el cardinal de los subconjuntos de un conjunto infinito es *estrictamente mayor* que el de los números naturales. Luego tembién el cardinal de Q es mayor que el del conjunto de los números naturales, es decir que Q es no numerable.

En particular, $(\ell^\infty, \|.\|_\infty)$ y $(\mathcal{B}([a,b]), \|.\|_\infty)$ son no separables.

2. APLICACIONES LINEALES Y CONTINUAS

2.1. Espacios normados de aplicaciones lineales continuas

Sean $(X, \|\cdot\|)$, $(Y, \|\cdot\|)$ espacios normados sobre el mismo cuerpo. El conjunto de las aplicaciones lineales entre X e Y, $\mathcal{L}(X,Y)$ tiene una estructura natural de espacio vectorial. Como en en X y en Y existe una topología, tiene sentido considerar el subconjunto de $\mathcal{L}(X,Y)$ formado por las aplicaciones (lineales) que además sean continuas, que suele simbolizarse como $\mathcal{B}(X,Y)$ ó como $\mathcal{L}_c(X,Y)$. El vector nulo de $\mathcal{L}(X,Y)$, es decir la aplicación lineal cuyo núcleo es X, es siempre contínua, y por tanto $\mathcal{B}(X,Y)$ nunca es vacío. Las bien conocidas propiedades de las funciones continuas permiten además afirmar que $\mathcal{B}(X,Y)$ es un subespacio vectorial de $\mathcal{L}(X,Y)$.

Ejemplo 5 No toda aplicación lineal es continua, como sucede en espacios de dimensión finita. Por ejemplo, en el espacio normado $(\mathcal{C}([-1,1]), \|.\|_\infty)$ sea D el subespacio que forman las funciones de $\mathcal{C}([-1,1])$ que son derivables en el intervalo abierto $(-1,1)$. La sucesión de funciones (f_n) dada por

$$f_n(t) := \frac{1}{n}\text{sen}(nt),$$

tiene todos sus términos en D, y $\|f_n\|_\infty \leq 1/n$, por lo que converge en $(\mathcal{C}([-1,1]), \|.\|_\infty)$ a la función nula $\Theta = 0_D$. La aplicación $L : D \to \mathbb{R}$ definida por $L(f) := f'(0)$ es lineal, pero la sucesión

$$(L(f_n)) = (f_n'(0)) = (\cos(n0)) = (1,1,1,\ldots)$$

es constante y no converge a $L(\Theta) = 0$. Así pues, L no es continua en el origen de D.

PROPOSICIÓN 2 *Si $L \in \mathcal{L}(X,Y)$, los siguientes enunciados son equivalentes:*

a) L es continua en X.

b) L es continua en 0_X.

c) L es acotada en $B_X := \{v \in X : \|v\| \leq 1\}$, es decir, $\{\|L(v)\| : v \in B_X\}$ es acotado en $\mathbb{R}$.

d) Existe un número real M tal que, para todo $v \in X$, $\|L(v)\| \leq M\|v\|$.

Prueba : Es evidente que (a)$\Rightarrow$ (b).

Si L es continua en 0_X, dado $\varepsilon = 1$ existe $\delta > 0$ tal que si $x \in X$ con $\|x - 0_X\| < \delta$, entonces $\|L(x)\| = \|L(x) - L(0_X)\| < \varepsilon = 1$. Entonces, si $v \in B_X$, es no nulo, el vector $\frac{\delta}{2\|v\|}v$ tiene norma $\frac{\delta}{2} < \delta$, luego

$$\left\| L\left(\frac{\delta}{2\|v\|}v\right)\right\| < 1$$

es decir,

$$\|L(v)\| < \frac{2\|v\|}{\delta} \leq \frac{2}{\delta},$$

desigualdad que también se cumple (trivialmente) si $v = 0_X$. Así pues para todo $v \in B_X$, $\|L(v)\| \leq \frac{2}{\delta}$, luego L es acotada en B_X. Hemos probado pues (b)$\Rightarrow$ (c)

Si L es acotada en B_X, es decir si existe $M \in \mathbb{R}$ tal que $\|L(w)\| \leq M$ para todo $w \in B_X$, dado $v \in X$, $v \neq 0_X$, se tiene que $\frac{1}{\|v\|}v$ tiene norma 1, luego pertenece a B_X, por lo que

$$\left\| L\left(\frac{1}{\|v\|}v\right)\right\| \leq M,$$

lo cual implica inmediatamente que $\|L(v)\| \leq M\|v\|$ para todo $v \in B_X$. Luego hemos probado que (c)$\Rightarrow$(d).

Por último (d)$\Rightarrow$(a) es inmediato porque tomando en particular en (d) $v = x - y$, con $x, y \in X$, y teniendo en cuenta que L es lineal,

$$\|L(x) - L(y)\| = \|L(x - y)\| \leq M\|x - y\|,$$

lo que prueba que L cumple una condición de Lipschitz, o sea que es uniformemente continua (y por tanto continua) en X.

2.2. Normas de aplicaciones lineales

Para cada $L \in \mathcal{B}(X, Y)$ están bien definidos los siguientes números reales:

$$\begin{aligned}
a(L) &:= \sup\left\{\tfrac{\|L(v)\|}{\|v\|} : v \in X, v \neq 0\right\}.\\
b(L) &:= \sup\{\|L(u)\| : u \in X, \|u\| = 1\}.\\
c(L) &:= \sup\{\|L(v)\| : v \in X, \|v\| \leq 1\}.\\
d(L) &:= \text{ínf}\{M \in \mathbb{R} : \forall v \in X, \|L(v)\| \leq M\|v\|\}.
\end{aligned}$$

Proposición 3 *Para cada $L \in \mathcal{B}(X, Y)$, $a(L) = b(L) = c(L) = d(L)$.*

Prueba : Como evidentemente

$$\{\|L(u)\| : u \in X, \|u\| = 1\} \subset \{\|L(u)\| : u \in X, \|u\| \leq 1\},$$

inmediatamente se tiene que $b(L) \leq c(L)$.

Si $M_0 \in \mathcal{M} := \{M \in \mathbb{R} : \forall v \in X, \|L(v)\| \leq M\|v\|\}$ entonces para todo vector $v \in X$, con norma menor o igual que 1

$$\|L(v)\| \leq M_0\|v\| \leq M_0,$$

luego M_0 es una cota superior del conjunto $\{\|L(u)\| : u \in X, \|u\| \leq 1\}$. Por definición de supremo se sigue que $c(L) \leq M_0$. Esto es cierto para todos los elementos del conjunto $\mathcal{M}$, luego

$$c(L) \leq \text{ínf}\, \mathcal{M} = d(L).$$

También por definición de supremo, si $v \neq 0$

$$a(L) \geq \frac{\|L(v)\|}{\|v\|}$$

luego para todo $v \in X$, $\|L(v)\| \leq a(L)\|v\|$. De aquí se sigue que $a(L) \in \mathcal{M}$, luego

$$d(M) := \inf \mathcal{M} \leq a(L).$$

En definitiva tenemos $b(L) \leq c(L) \leq d(L) \leq a(L)$.

Pero dado cualquier vector no nulo $v \in X$, $u = \frac{1}{\|v\|}v$ tiene norma 1 y

$$\frac{\|L(v)\|}{\|v\|} = \|L(\frac{1}{\|v\|}v)\| = \|L(u)\|.$$

luego

$$\left\{\frac{\|L(v)\|}{\|v\|} : v \in X, v \neq 0\right\} \subset \{\|L(u)\| : u \in X, \|u\| = 1\}.$$

Por tanto, $a(L) \leq b(L)$, y se completa la prueba.

$\square$

Definición 10 *Para cada $L \in \mathcal{B}(X,Y)$, se llama norma de L al número real*

$$\|L\| := \sup\{\|L(v)\| : v \in X, \|v\| \leq 1\}.$$

Como los números reales del enunciado anterior en realidad son el mismo, la expresión de cualquiera de ellos puede tomarse como definición de $\|L\|$. Comprobar que, efectivamente, se trata de una norma, es casi un ejercicio. Antes veamos una desigualdad útil

Corolario 3 *En las condiciones de la proposición anterior, para todo $x \in X$,*

$$\|L(x)\| \leq \|L\|\|x\|.$$

Prueba : Por definición del supremo $a(L) = \|L\|$ se tiene que para todo vector no nulo $x \in X$,

$$\|L\| = a(L) \geq \frac{\|L(x)\|}{\|x\|}$$

luego, $\|L(x)\| \leq \|L\|\|x\|$ para los vectores x no nulos, y, trivialmente, también para $x = 0_X$.

$\square$

Proposición 4 *En el espacio vectorial $\mathcal{B}(X,Y)$, la aplicación*

$$L \mapsto \|L\| := \inf\{M \in \mathbb{R} : \forall v \in X, \|L(v)\| \leq M\|v\|\}.$$

es una norma.[3]

[3]En la línea anterior aparece el mismo símbolo, $\|\cdot\|$, para *tres* normas, la de X, la de Y y la que se está definiendo en $\mathcal{B}(X,Y)$. A pesar de esto, si se lee con cuidado no debe haber confusión.

Prueba : Puede observarse en primer lugar que, de la definición de $\|L\|$ se sigue inmediatamente que, para todo $v \in X$,

$$\|L(v)\| \leq \|L\|\|v\|.$$

Luego si $\|L\| = 0$, para todo $v \in X$, $\|L(v)\| = 0$, y entonces L es la aplicación lineal nula, $0_{\mathcal{B}(X,Y)}$.

Por otra parte, para todo $v \in X$, $\|0_{\mathcal{B}(X,Y)}(v)\| = 0 \leq 0\|v\|$, luego $\|0_{\mathcal{B}(X,Y)}(v)\| = 0$.

Además, si $S, L \in \mathcal{B}(X,Y)$, para todo $v \in X$,

$$\|(S+L)(v)\| \leq \|S(v)\| + \|L(v)\| \leq \|S\|\|v\| + \|L\|\|v\| = (\|S\| + \|L\|)\|v\|.$$

Por tanto

$$(\|S\| + \|L\|) \in \{M \in \mathbb{R} : \forall v \in X, \|(S+L)(v)\| \leq M\|v\|\}$$

de donde

$$\|S+L\| \leq \|S\| + \|L\|.$$

Por último está también claro que si $\lambda \neq 0$ es un escalar del cuerpo sobre el que estén definidos X e Y, para todo $v \in X$,

$$\|(\lambda L)(v)\| = |\lambda|\|L(v)\| \leq |\lambda|\|L\|\|v\|,$$

y razonando como antes,

$$\|\lambda L\| \leq |\lambda|\|L\|.$$

Aplicando dos veces esta desigualdad, tenemos

$$\|\lambda L\| \leq |\lambda|\|L\| = |\lambda|\|\lambda^{-1}\lambda L\| \leq |\lambda||\lambda^{-1}|\|\lambda L\| = \|\lambda L\|$$

de donde se sigue que $\|\lambda L\| = |\lambda|\|L\|$.

Vamos a calcular la norma de algunas aplicaciones lineales que se usan frecuentemente, o que pueden ayudar a comprender mejor el concepto de norma de una aplicación lineal.

La primera idea es que la norma de las aplicaciones lineales en cierto sentido es una medida de la contracción o dilatación que produce la aplicación a los vectores de su dominio.

Ejemplo 6 *Homotecias. Sea* $(X, \|\cdot\|)$ *un espacio normado, y sea* λ *un escalar no nulo. Si* $H_\lambda : X \to X$ *es la homotecia*

$$H_\lambda(x) = \lambda x$$

entonces $\|H_\lambda\| = |\lambda|$.

Basta aplicar directamente la definición:

$$\|H_\lambda\| = \sup\{\frac{\|H_\lambda(x)\|}{\|x\|} : x \in X, x \neq 0\} = \sup\{\frac{\|\lambda x\|}{\|x\|} : x \in X, x \neq 0\} = |\lambda|.$$

Si S es un conjunto no vacío sabemos que $(\mathcal{B}(S), \|.\|_\infty)$ es un espacio de Banach . Sobre este espacio hay una aplicaciones lineales especialmente sencillas: Si $s \in S$, y $f \in \mathcal{B}(S)$ se llama *evaluación en* s la aplicación lineal $E_s : \mathcal{B}(S) \to \mathbb{K}$ dada por $E_s(f) := f(s)$.

Ejemplo 7 *Evaluaciones. Consideremos el espacio normado* $(\mathcal{B}(S), \|.\|_\infty)$ *y en él el funcional evaluación en* $s \in S$*, es decir*

$$E_s(f) := f(s).$$

La norma de cualquier E_s *es* 1*.*

Basta aplicar la definición, y recordar la definición de $\|f\|_\infty$:

$$\|E_s\| := \sup\{\frac{|E_s(f)|}{\|f\|_\infty} : f \in \mathcal{B}(S), f \neq 0_{\mathcal{B}(S)}\} = \sup\{\frac{|f(s)|}{\|f\|_\infty} : f \in \mathcal{B}(S), f \neq 0_{\mathcal{B}(S)}\} \leq 1.$$

Por otra parte, para la función constante 1 sobre S (a la que llamaremos u), evidentemente se tiene $\|u\|_\infty = 1$, luego

$$\|E_s\| := \sup\{|E_s(f)\| : f \in \mathcal{B}(S), \|f\|_\infty = 1\} \geq |E_s(u)| = 1.$$

Ejemplo 8 *Sean* $L : \mathbb{R}^2 \to R^2$ *la aplicación lineal y continua dada por*

$$L(x, y) := (x + 2y, 3x + 4y).$$

Consideremos L como un elemento de $\mathcal{B}((\mathbb{R}^2, \|\cdot\|_\infty), (\mathbb{R}^2, \|\cdot\|_\infty))$.

Como $\|(1, 1)\|_\infty = 1$,

$$\|L\| \geq \|L(1, 1)\|_\infty = \|(3, 7)\|_\infty = 7.$$

Por otra parte, para cualquier $(x, y) \in \mathbb{R}^2$,

$$\|(x + 2y, 3x + 4y)\|_\infty \leq \text{máx}\{|x| + 2|y|, 3|x| + 4|y|\} \leq 7 \,\text{máx}\{|x|, |y|\} = 7\|(x, y)\|_\infty,$$

luego $\|L\| \leq 7$. En definitiva $\|L\| = 7$.

Consideremos ahora L como un elemento de $\mathcal{B}((\mathbb{R}^2, \|\cdot\|_1), (\mathbb{R}^2, \|\cdot\|_1))$. Para cualquier $(x, y) \in \mathbb{R}^2$,

$$\|(x + 2y, 3x + 4y)\|_1 \leq |x| + 2|y| + 3|x| + 4|y| \leq 4|x| + 6|y| \leq 6(|x| + |y|) = 6\|(x, y)\|_1$$

Luego $\|L\| \leq 6$.

Por otra parte, $\| (0, 1) \| = 1$, luego

$$\|L\| \geq \|L((0, 1))\|_1 = \|(2, 4)\|_1 = 6.$$

En definitiva $\|L\| = 6$.

Este ejemplo pone de manifiesto que la norma definida en $\mathcal{B}(X, Y)$ depende fuertemente de (las normas de) los espacios $(X, \|\cdot\|)$, $(Y, \|\cdot\|)$.

Ejemplo 9 *Un operador integral. Sea* $\sigma \in \mathcal{C}([a, b])$ *una función dada y sea* $T : \mathcal{C}([a, b]) \to \mathbb{R}$ *la aplicación lineal dada por*

$$T(f) := \int_a^b \sigma(t)f(t)dt.$$

Entonces $\|T\| = \int_a^b |\sigma(t)|dt$*, si en* $\mathcal{C}([a, b])$ *se considera la norma del supremo.*

La desigualdad

$$|T(f)| \leq \int_a^b |\sigma(t)f(t)|dt \leq [\int_a^b |\sigma(t)|dt]\|f\|_\infty$$

nos da inmediatamente que $\|T\| \leq \int_a^b |\sigma(t)|dt$.

Si σ fuera no negativa, bastaría tomar la función constante 1 en $[a,b]$ para tener $\|T\| \geq \int_a^b \sigma(t)1dt = \int_a^b |\sigma(t)|dt$. Pero en al caso general hay que proceder con un poco más de cuidado. Sea $g : (-\infty, \infty) \to [-1,1]$ definida por

$$g(x) := \begin{cases} -1 & x \leq -1 \\ x & -1 \leq x \leq 1 \\ 1 & 1 \leq x \end{cases}$$

y definamos también las funciones continuas $g_n : [a,b] \to [-1,1]$ por

$$g_n(t) := g\big(n\sigma(t)\big).$$

Observemos que si $\sigma(t) > 0$ entonces para n suficientemente grande $n\sigma(t) > 1$ y entonces $g_n(t) = 1$. Igualmente, si $\sigma(t) < 0$ entonces para n suficientemente grande $n\sigma(t) < -1$ y entonces $g_n(t) = -1$. Finalmente, si $\sigma(t) = 0$ $g_n(t) = 0$ para todo n.

En suma podemos afirmar que para cada $t \in [a,b]$,

$$g_n(t) \to sg(\sigma(t))$$

donde $sg(\sigma(t))$ es el signo del número real $\sigma(t)$.

Por otro lado es claro que $\|g_n\|_\infty \leq 1$, luego para todo entero positivo n

$$\|T\| \geq T(g_n) := \int_a^b \sigma(t)g_n(t)dt.$$

Teniendo en cuenta que para todo t en $[a,b]$, $|\sigma(t)g_n(t)| \leq |\sigma(t)|$ y que σ es sumable en $[a,b]$, por el teorema de la convergencia dominada tenemos que

$$\|T\| \geq \lim_n \int_a^b \sigma(t)g_n(t)dt = \int_a^b \lim_n(\sigma(t)g_n(t))dt = \int_a^b \sigma(t)sg(t)dt = \int_a^b |\sigma(t)|dt.$$

2.3. El espacio de Banach $(\mathcal{B}(X,Y), \|.\|)$

La proposición anterior nos ha presentado un nuevo ejemplo abstracto de espacio normado: El espacio vectorial $\mathcal{B}(X,Y)$ de las aplicaciones lineales y continuas entre dos espacios normados $(X, \|\cdot\|)$ $(Y, \|\cdot\|)$ sobre el mismo cuerpo con la norma de las aplicaciones lineales.

Para $T \in \mathcal{B}(X,Y)$, $\|T\|$ depende como hemos visto, tanto de la propia aplicación T,como de las normas que se estén considerando en X y en Y. (Esto último es origen de confusiones, porque no se refleja en la notación).

La razón de adoptar una definición así se verá enseguida, aunque ha de notarse que no es la única que podría haberse usado para "normar"$\mathcal{B}(X,Y)$.

No obstante, si no se especifica lo contrario, en $\mathcal{B}(X,Y)$ se supone definida precisamente esta norma, que suele llamarse *norma de las aplicaciones lineales* ó *norma de los operadores*.

TEOREMA 13 *Si $(Y, \|\cdot\|)$ es un espacio de Banach, entonces $\mathcal{B}(X,Y)$ es un espacio de Banach (con la norma de las aplicaciones lineales).*

Prueba : Sea (L_n) una sucesión de Cauchy en $\mathcal{B}(X,Y)$ (respecto de la norma de las aplicaciones lineales y continuas). Dado $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que si $p, q \geq n_0$

$$\|L_p - L_q\| < \varepsilon,$$

y entonces fijado $x \in X$,

$$\|L_p(x) - L_q(x)\| \leq \|L_p - L_q\|\|x\| < \varepsilon\|x\|.$$

Esto nos demuestra que $(L_n(x))$ es una sucesión de Cauchy en $(Y, \|\cdot\|)$. Como este espacio es completo, existe en Y

$$\lim L_n(x).$$

Naturalmente el elemento anterior depende del x previamente fijado en X. Es natural, por tanto, denotarlo como función de x, por ejemplo $L(x)$.

Así pues $x \mapsto L(x) := \lim_n L_n(x)$ es una aplicación entre X e Y. Además está claro por las propiedades de los límites de sucesiones que L es lineal.

De la continuidad de la norma y de la desigualdad

$$\|L_p(x) - L_q(x)\| \leq \varepsilon\|x\|$$

se sigue

$$\lim_p [\|L_p(x) - L_q(x)\|] \leq \varepsilon\|x\|$$

o sea,

$$\|L(x) - L_q(x)\| \leq \varepsilon\|x\|.$$

Por tanto, como

$$\|L(x)\| \leq \|L(x) - L_q(x)\| + \|L_q(x)\| \leq \varepsilon\|x\| + \|L_q\|\|x\|$$

tenemos que L es continua, con $\|L\| \leq \varepsilon + \|L_q\|$ para todo $q \geq n_0$.

Por último, de la mismo desigualdad

$$\|L(x) - L_q(x)\| \leq \varepsilon\|x\|.$$

se sigue también por definición de la norma de las aplicaciones lineales que

$$\|L - L_q\| \leq \varepsilon$$

para todo $q \geq \varepsilon$. En otras palabras,

$$\lim \|L - L_n\| = 0$$

lo que equivale a que (L_n) converge a L en $(\mathcal{B}(X,Y), \|\cdot\|)$.

$\square$

Como vemos la prueba se parece mucho a la desarrollada para el caso del espacio normado de las aplicaciones reales acotadas con la norma del supremo. En el caso particular en que los espacios $(X, \|\cdot\|)$, $(Y, \|.\|)$ sean el mismo, entonces se usa la notación $\mathcal{B}(X)$ en lugar de $\mathcal{B}(X,X)$. Si $S, T \in \mathcal{B}(X)$, entonces la aplicación compuesta $S \circ T$ también pertenece a $\mathcal{B}(X)$ lo que da al conjunto $\mathcal{B}(X)$ una estructura muy rica, ya que además de la estructura de espacio normado (con la norma de los operadores), existe una de anillo (en general no conmutativo) cuando se considera la composición de aplicaciones como operación multiplicativa. Además existe una çompatibilidad"de estas estructuras que es consecuencia de la desigualdad $\|S \circ T\| \leq \|S\|\|T\|$. Otras normas sobre $\mathcal{B}(X)$ no cumplen esta desigualdad, y esta es una de las razones de definir en $\mathcal{B}(X,Y)$ la norma que se maneja usualmente.

PROPOSICIÓN 5 *Si X, Y, Z son espacios normados sobre el mismo cuerpo, $S \in \mathcal{B}(X,Y)$, y $T \in \mathcal{B}(Y,Z)$*

$$\|T \circ S\| \leq \|T\|\|S\|.$$

Prueba : Para cada $x \in X$,

$$\|(T \circ S)(x)\| \leq \|T\|\|S(x)\| \leq \|T\|\|S\|\|x\|$$

y de la definición de la norma de las aplicaciones lineales se sigue inmediatamente que $\|T \circ S\| \leq \|T\|\|S\|$.

$\square$

2.3.1. Espacio dual

Uno caso particular importante se da cuando $(Y, \|\cdot\|)$ es el cuerpo $\mathbb{K}$ de los escalares con la norma del módulo. Entonces $(\mathcal{B}(X,\mathbb{K}), \|.\|)$ suele llamarse espacio (normado) dual topológico de $(X, \|\cdot\|)$, utilizándose para su notación bien X' o bien X^* según gustos. Estos espacios duales siempre son de Banach, porque $(\mathbb{K}, |\cdot|)$ es un espacio de Banach.

DEFINICIÓN 11 *Sean $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ dos espacios normados y sea $T : X \to Y$ una aplicación.*

- *Se dice que T es un isomorfismo algebraico si T es lineal y biyectiva.*
- *Se dice que T es un isomorfismo topológico si T es un isomorfismo algebraico y tanto T como T^{-1} son continuas.*
- *Se dice que T es un isomorfismo isométrico si T es un isomorfismo algebraico que, además, es una isometría, esto es, si para cualesquiera $x, y \in X$, se cumple la igualdad*
$$\|T(x) - T(y)\|_Y = \|x - y\|_X.$$
- *Si existe un isomorfismo algebraico entre $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$, se dice que estos espacios son algebraicamente isomorfos.*

- *Si existe un isomorfismo topológico entre* $(X, \|\cdot\|_X)$ *e* $(Y, \|\cdot\|_Y)$*, se dice que estos espacios son topológicamente isomorfos.*

- *Si existe un isomorfismo isométrico entre* $(X, \|\cdot\|_X)$ *e* $(Y, \|\cdot\|_Y)$*, se dice que estos espacios son isométricamente isomorfos.*

Observar que, si $T : X \to Y$ es lineal, entonces T es una isometría si, y sólo si, $\|T(x)\|_Y = \|x\|_X$ para todo $x \in X$.

Los espacios normados $\mathcal{B}(X,Y)$ son abstractos por lo que es muy conveniente, ante cada caso concreto, poder representarlo es decir, disponer de otro espacio normado isométricamente isomorfo a él y de manejo más cómodo. Por ejemplo, si $(X, \|\cdot\|)$, $(Y, \|\cdot\|)$ son de dimensión finita, $\mathcal{B}(X,Y) = \mathcal{L}(X,Y)$, como espacios vectoriales pueden ser representados por espacios de matrices.

Otro ejemplo importante es el de $\mathcal{B}(\mathbb{R}, Y)$, (si el espacio normado Y es real) que admite al Y como representación, es decir que:

Si $y \in Y$, $L_y(t) := ty$ define un elemento de $\mathcal{B}(\mathbb{R}, Y)$, y recíprocamente, si $L \in \mathcal{B}(\mathbb{R}, Y)$, $L(t) = L(t,1) = tL(1) = L_{L(1)}(t)$.

Esto sugiere que la aplicación biyectiva $y \mapsto L_y$ es el isomorfismo isométrico entre Y y $\mathcal{B}(\mathbb{R}, Y)$, pues, efectivamente $\|L_y\| = \|y\|$.

Para los espacios normados duales de los ejemplos que hasta ahora hemos considerado tenemos:

El espacio dual de...	Se identifica con
$(\mathbb{R}^n, \|.\|_2)$	$(\mathbb{R}^n, \|.\|_2)$
$(\mathbb{R}^n, \|.\|_1)$	$(\mathbb{R}^n, \|.\|_\infty)$
$(\mathbb{R}^n, \|.\|_\infty)$	$(\mathbb{R}^n, \|.\|_1)$
$(\ell_p, \|.\|_p)$, $(1 < p < \infty)$	$(\ell_q, \|.\|_q)$ $(p + q = pq)$
$(\ell_1, \|.\|_1)$	$(\ell_\infty, \|.\|_\infty)$
$(c_0, \|.\|_\infty)$	$(\ell_1, \|.\|_1)$
$(L^p(\Omega), \|.\|_p)$, $(1 < p < \infty)$	$(L^q(\Omega), \|.\|_q)$ $(p + q = pq)$
$(L^1(\Omega), \|.\|_1)$	$(L^\infty(\Omega), \|.\|_\infty)$

Esto debe entenderse con cuidado. Significa que existe un isomorfismo isométrico entre cada uno de la primera columna y el espacio de su misma fila en la segunda columna.

Si se toma un elemento f de un espacio de la segunda columna, de algún modo puede considerarse como una aplicación lineal y continua sobre el espacio de su misma fila, pero este modo no está especificado en la tabla, aunque en los tres primeros casos es el siguiente:

En los espacios de sucesiones, si $y = (y_1, \ldots, y_n, \ldots)$ pertenece a algún espacio de la segunda columna, se identifica con la aplicación lineal continua (funcional) sobre el respectivo espacio de sucesiones de su misma fila que, sobre el elemento $(x_1, \ldots, x_n, \ldots)$

toma el valor

$$y(x) := \sum_{n=1}^{\infty} x_n y_n.$$

Por último, en los espacios $(L^p(\Omega), \|.\|_p)$, si f pertenece a algún espacio de la segunda columna, se identifica con la aplicación lineal continua (funcional) sobre el respectivo espacio de su misma fila que sobre el elemento g toma el valor

$$f(g) := \int_{\Omega} f(t)g(t)dt.$$

Veremos algún caso:

Teorema 14 *El dual de ℓ_1 es ℓ_∞.*

Supongamos que $f \in \ell_1^*$. Para cada $x = (x_n) \in \ell_1$, es fácil ver que $x = \sum_{i=1}^{n} x_n e^i$, donde $e^i = (\delta_{j,i})_{j=1}^{\infty}$.

$$f(x) = f\left(\lim_{n\to\infty} \sum_{i=1}^{n} x_i e^i\right) = \lim_{n\to\infty} \sum_{i=1}^{n} x_i f(e^i) = \sum_{n=1}^{\infty} x_n f(e_n).$$

Esto sugiere definir $T : \ell_1^* \to \ell_\infty$ por $T(f) = (f(e_n))$. Veamos que T es un isomorfismo isométrico.

- En primer lugar, T está bien definido, es decir si $f \in \ell_1^*$, entonces $(f(e_n)) \in \ell_\infty$ pues

  $$|f(e_n)| \leq \|f\| \cdot \|e_n\|_1 = \|f\|$$

  para todo $n \in \mathbb{N}$.

- Que T es lineal es inmediato.

- T es inyectiva: si $T(f) = 0_{\ell_\infty} = (0, 0, \ldots)$, entonces es $f(e_n) = 0$ para todo $n \in \mathbb{N}$, de donde, para cada $x \in \ell_1$, $f(x) = \lim_{n\to\infty} \sum_{i=1}^{n} x(i) f(e^i) = 0$, o sea f es el funcional nulo sobre ℓ_1.

- T es exhaustiva: sea $a = (a_n) \in \ell_\infty$ y definamos $f : \ell_1 \to \mathbb{K}$ por $f(x) = \sum_{n=1}^{\infty} x(n) a_n$. Entonces, f es lineal y, además

  $$|f(x)| \leq \sum_{n=1}^{\infty} |x(n)| \cdot |a_n| \leq \|x\|_1 \cdot \|a\|_\infty,$$

  lo que nos dice que la serie $\sum_{n=1}^{\infty} x(n) a_n$ es absolutamente convergente, y por tanto convergente en $\mathbb{K}$. Finalmente, $T(f) = (f(e_n)) = (a_n) = a$.

- T es una isometría: sea $f \in \ell_1^*$ y veamos que $\|T(f)\|_\infty = \|f\|$.

  Para obtener la desigualdad $\|f\| \leq \|T(f)\|_\infty$, basta tener en cuenta que para todo $x \in \ell_1$ se tiene que

  $$|f(x)| = \left|\sum_{n=1}^{\infty} x(i) f(e^i)\right| \leq \|T(f)\|_\infty \cdot \|x\|_1.$$

Para obtener la otra desigualdad observemos que, al ser $\|e_n\|_1 = 1$ para todo $n \in \mathbb{N}$, es

$$\|f\| \geq |f(e_n)| \quad , \qquad n = 1, 2, \ldots,$$

de donde $\|f\| \geq \|(f(e_n))\|_\infty = \|T(f)\|_\infty$.

□

2.3.2. Aplicación conjugada

DEFINICIÓN 12 *Si X, Y son espacios de Banach sobre el mismo cuerpo $\mathbb{K}$, y $T : X \to Y$ una aplicación lineal y continua, se llama aplicación conjugada de T a $T^t : Y^* \to X^*$ dada por*

$$T^t(g) = g \circ T.$$

Ciertamente, $g \circ T$ es lineal (por ser composición de aplicaciones lineales) y continua (por ser composición de aplicaciones continuas), y aplica X en el cuerpo $\mathbb{K}$.

T^t transforma funcionales lineales en funcionales lineales, y a su vez, es lineal:

Dados $f, g \in Y^*$, $a, b \in \mathbb{K}$,

$$\begin{aligned} T^t(af + bg) &= (af + bg) \circ T \\ &= (af) \circ T + (bg) \circ T \\ &= a(f \circ T) + b(g \circ T) \\ &= aT^t(f) + bT^t(g) \end{aligned}$$

Por otra parte, si $g \in Y^*$,

$$\begin{aligned} \|T^t(g)\| &:= \sup\{|T^t(g)(x)| : x \in X, \|x\| \leq 1\} \\ &= \sup\{|(g \circ T)(x)| : x \in X, \|x\| \leq 1\} \\ &= \sup\{|g(T(x))| : x \in X, \|x\| \leq 1\} \\ &\leq \sup\{\|g\|\|T(x)| : x \in X, \|x\| \leq 1\} \\ &\leq \sup\{\|g\|\|T\|\|x| : x \in X, \|x\| \leq 1\} \\ &\leq \|T\|\|g\| \end{aligned}$$

lo que nos muestra que T^t es continua con $\|T^t\| \leq \|T\|$.

Para dar ejemplos no triviales de aplicaciones conjugadas necesitamos conocer más hechos referentes a los espacios duales.

2.4. Isomorfismos

Sean $(X, \|\cdot\|)$ $(Y, \|\cdot\|)$ espacios normados sobre el mismo cuerpo. Si existe una aplicación lineal biyectiva entre X e Y se dice que estos espacios vectoriales son algebraicamente isomorfos. Si $L : X \to Y$ es lineal uno-a-uno (inyectiva) entonces $L^{-1} : L(X) \to X$ es automáticamente lineal.

Efectivamente, si $y_1, y_2 \in L(X)$, existen $x_1, x_2 \in X$ con $y_1 = L(x_1), y_2 = L(x_2)$. Entonces

$$L^{-1}(y_1 + y_2) = L^{-1}(L(x_1) + L(x_2)) = L^{-1}(L(x_1 + x_2)) = x_1 + x_2 = L^{-1}(y_1) + L^{-1}(y_2)$$

Análogamente, para todo escalar λ,

$$L^{-1}(\lambda y_1) = L^{-1}(\lambda L(x_1)) = L^{-1}(L(\lambda x_1)) = \lambda x_1 = \lambda L^{-1}(y_1).$$

En particular, si los espacios vectoriales X, Y son algebraicamente isomorfos, tanto L como su inversa son lineales.

No toda aplicación lineal e inyectiva de un espacio normado X en sí mismo, es sobre.

Ejemplo 10 *La aplicación $S : \ell_2 \to \ell_2$ dada por*

$$S(x) = S((x_1, x_2, \ldots)) = (0, x_1, x_2, \ldots)$$

es una isometría lineal con $\|S\| = 1$, aunque no es exhaustiva (sobre).

Ejemplo 11 Consideremos los espacios normados $X := (\ell_2, \|.\|_2)$ y $Y := (\ell_2, \|.\|_\infty)$.

Trivialmente, la aplicación lineal y biyectiva $Id : \ell_2 \to \ell_2$ tiene inversa también lineal. Como para todo $x \in \ell_2$, $\|x\|_\infty \leq \|x\|_2$, la aplicación identidad $Id : X \to Y$ es continua. Pero en cambio, en este contexto Id^{-1} es no continua porque la sucesión $\left(v^{(n)}\right)$ en la bola unidad de Y definida por

$$v^{(n)} = (\overbrace{\frac{1}{n}, \ldots, \frac{1}{n}}^{n^2 \text{ veces}}, 0, \ldots)$$

verifica que $\|v^{(n)}\|_2 = 1, \|v^{(n)}\|_\infty = \frac{1}{n}$, luego $(v^{(n)})$ converge a 0_{ℓ_2} en $(\ell_2, \|\cdot\|_\infty)$, pero no en $(\ell_2, \|\cdot\|_2)$. En otras palabras, $Id^{-1} : Y \to X$ no es continua. Es claro que X e Y son espacios normados algebraicamente isomorfos. *pero la identidad Id entre estos espacios no es un homeomorfismo.*

Ejemplo 12 *Los espacios normados $(\mathbb{R}^n, \|, \|_2)$, $(\mathbb{R}^n, \|\cdot\|_\infty)$ son topológicamente isomorfos.*

Basta tomar la aplicación identidad entre ellos que es evidentemente lineal y biyectiva.

Además, para todo $x \in \mathbb{R}^n$,

$$\|I(x)\|_\infty = \|x\|_\infty \leq \|x\|_2,$$

luego I es continua.

Recíprocamente, para todo $x \in \mathbb{R}^n$,

$$\|I^{-1}(x)\|_2 = \|x\|_2 \leq \sqrt{n}\|x\|_\infty$$

luego I^{-1} es continua.

Una aplicación $L : X \to Y$ es una isometría lineal si es lineal y cumple $\|L(x)\| = \|x\|$ para todo $x \in X$. Toda isometría lineal es continua e inyectiva, y tiene norma 1. (Aunque no necesariamente es biyectiva, como puede comprobarse en el ejemplo 10.

Ejemplo 13 Sea X un espacio de Hilbert (real o complejo). Para todo $f \in X^*$ sabemos (teorema de representación de Riesz) que existe un único elemento $v \in X$ con $f(x) =< v, x >$ para todo $x \in X$. Por la desigualdad de Cauchy-Schwartz,

$$|f(x)| = | < v, x > | \leq \|v\|\|x\|,$$

luego $\|f\| \leq \|v\|$. Si $v = 0_X$ entonces $f = 0_{X^*}$ y $\|f\| = \|v\|$. Si $v \neq 0_X$,

$$\|f\| \geq \left| f\left(\frac{v}{\|v\|}\right) \right| = \|v\|.$$

Es decir que $\|f\| = \|v\|$ en todos los casos.

Podemos pues considerar la aplicación $J : X \to X^*$ dada por $J(v) = f_v$ donde $f_v(x) =< x, v >$ para cada $x \in X$.

Hemos comprobado que $\|J(v)\| = \|v\|$, o mejor que $\|J(v)\|_{X^*} = \|v\|_X$. Luego J es una isometría de X *sobre* X^*, es decir J es biyectiva. Sin embargo, en caso de ser X un espacio de Hilbert real es diferente del caso de ser X espacio de Hilbert sobre el cuerpo de los complejos.

Si $v, w \in X$ está claro que, para todo $x \in X$

$$f_{v+w}(x) = \langle x, v + w \rangle = \langle x, v \rangle + \langle x, w \rangle = f_v(x) + f_w(x)$$

y esto se cumple en el caso real y en caso complejo, es decir

$$J(v + w) = J(v) + J(w).$$

Pero si $\lambda \in \mathbb{C}$, para todo $x \in X$ se tiene que

$$f_{\lambda v}(x) = \langle x, \lambda v \rangle = \overline{\lambda}\langle x, v \rangle = \overline{\lambda} f_v(x)$$

es decir

$$J(\lambda v) = \overline{\lambda} J(v).$$

Por tanto, en el caso de un espacio de Hilbert complejo, J es una isometría no lineal, mientras que en el caso de un espacio de Hilbert real, la aplicación J es una isomorfismo isométrico, y los espacios X y X^* son isométricamente isomorfos.

No es fácil dar ejemplos de pares de espacios normados isométricamente isomorfos que no sean triviales o bien de comprobación difícil para nuestro nivel.

Proposición 6 *(Caracterización de isomorfismos topológicos.)*

Sean $(X, \|\cdot\|)$ *y* $(Y, \|\cdot\|)$ *espacios normados sobre el mismo cuerpo. La aplicación lineal y sobre* $T : X \to Y$ *es un isomorfismo topológico si y sólo si existen dos constantes positivas* m, M *tales que, para todo* $x \in X$,

$$m\|x\| \leq \|T(x)\| \leq M\|x\|.$$

Prueba : Si T es isomorfismo topológico, en particular es continua, luego

$$\|T(x)\| \leq \|T\|\|x\|$$

para todo $x \in X$. También T^{-1} es continua; así pués para todo x de X,

$$\|x\| = \|T^{-1}(T(x))\| \leq \|T^{-1}\|\|T(x)\|$$

de donde

$$\frac{1}{\|T^{-1}\|}\|x\| \leq \|T(x)\|$$

Observar que no puede ser $\|T^{-1}\| = 0$ porque en ese caso T^{-1} sería la aplicación lineal nula, que no tiene inversa.

Recíprocamente, si existen las constantes positivas m, M cumpliendo la condición del enunciado, la desigualdad

$$\|T(x)\| \leq M\|x\|$$

(válida para todo $x \in X$) demuestra que T es continua, y la desigualdad

$$m\|x - y\| \leq \|T(x - y)\| = \|T(x) - T(y)\|,$$

(válida para cualesquiera $x, y \in X$), demuestra que T es inyectiva.

Sabemos que entonces T^{-1} está definida en Y (por ser T sobre) y es lineal. Además la desigualdad

$$m\|T^{-1}(y)\| \leq \|T(T^{-1}(y))\| = \|y\|$$

(válida para todo $y \in Y$) demuestra que

$$\|T^{-1}(y)\| \leq \frac{1}{m}\|y\|$$

para todo $y \in Y$, luego T^{-1} es continua.

En suma, si existen las contantes del enunciado, T y su inversa son (lineales) biyectivas y continuas. Luego T es un isomorfismo topológico.

$\square$

El siguiente enunciado asegura la continuidad de todas las aplicaciones lineales definidas sobre determinados espacios.

Proposición 7 *(Continuidad automática). Si* $(Y, \|\cdot\|)$ *es un espacio normado sobre* $\mathbb{K}$ *toda aplicación lineal* $L : \mathbb{K}^n \to Y$ *es continua, si en* $\mathbb{K}^n$ *se considera definida la topología de la norma euclídea.*

Prueba : Sea $(e^{(1)}, \ldots, e^{(n)})$ la base canónica de $\mathbb{K}^n$. Para cada $x = x_1 e^{(1)} + \ldots + x_n e^{(n)}$ se tiene

$$\|L(x)\|| \leq |x_1|\|L(e^{(1)})\| + \ldots + |x_n|\|L(e^{(n)})\|.$$

Si consideramos los vectores de $\mathbb{R}^n$, $v := (|x_1|, \ldots, |x_n|)$ $w := (\|L(e^{(1)}\|, \ldots, \|L(e^{(n)}\|)$, vemos que el segundo miembro de esta última desigualdad es el producto escalar ordinario en $\mathbb{R}^n$, $\langle v, w\rangle$. Podemos aplicar entonces la desigualdad de Cauchy-Schwartz de $\mathbb{R}^n$ y tendremos, para cada $x \in \mathbb{K}^n$,

$$\|L(x)\| \leq \langle v, w\rangle \leq \|v\|_2\|w\|_2 = \|w\|_2\|x\|_2$$

luego L es continua, con $\|L\| \leq \|w\|_2$.

Teorema 15 *Todo espacio normado $(Y, \|\cdot\|)$ de dimensión n sobre el cuerpo $\mathbb{K}$ es topológicamente isomorfo a $(\mathbb{K}^n, \|\cdot\|_2)$.*

Si fijamos $(y^{(1)}, \ldots, y^{(n)})$, una base de Y, podemos definir la aplicación lineal $L : \mathbb{K}^n \to Y$ por

$$L(x_1, \ldots, x_n) := x_1 y^{(1)} + \ldots + x_n y^{(n)}.$$

Desde luego, L es inyectiva pues de ser $L(x_1, \ldots, x_n) = 0_Y$, se tendría que

$$x_1 y^{(1)} + \ldots + x_n y^{(n)} = 0_y,$$

y como $(y^{(1)}, \ldots, y^{(n)})$ es una base de Y, los escalares $x_1, \ldots, x_n$ han de ser todos cero, es decir $(x_1, \ldots, x_n) = 0_{\mathbb{K}^n}$.

Por la proposición anterior L es continua si consideramos en $\mathbb{K}^n$ la topología de la norma $\|.\|_2$. Entonces la función

$$(x_1, \ldots, x_n) \mapsto \|L(x_1, \ldots, x_n)\|$$

será continua en $(\mathbb{K}^n, \|\cdot\|_2)$. Como la esfera unidad S de $(\mathbb{K}^n, \|\cdot\|_2)$ es un conjunto compacto, existirán en S el máximo y el mínimo de esta función, o sea que existen $v, w \in S$ tales que para cualquier $(x_1, \ldots, x_n) \in S$,

$$\|L(v)\| \leq \|L(x_1, \ldots, x_n)\| \leq \|L(w)\|.$$

Observar que $\|L(v)\| = 0$ implicaría que $L(v) = 0_Y$, y como L es inyectiva entonces $v = 0_{\mathbb{K}^n}$ lo cual es absurdo porque $0_X \notin S$. O sea que si ponemos $m := \|L(v)\|, M := \|L(w)\|$, las constantes m, M son positivas y para cualquier $(x_1, \ldots, x_n) \in S$,

$$m \leq \|L(x_1, \ldots, x_n)\| \leq M.$$

En particular, dado cualquier vector no nulo, $z \in \mathbb{K}^n$, $\frac{1}{\|z\|_2} z \in S$, de donde

$$m \leq \|L(\frac{1}{\|z\|_2} z)\| \leq M.$$

o sea,

$$m\|z\|_2 \leq \|L(z)\| \leq M\|z\|_2.$$

Esto nos dice, según la proposición 6, que L es un isomorfismo topológico.

□

Corolario 4 *Dos espacios normados sobre el mismo cuerpo, que tengan la misma dimensión finita n, son topológicamente isomorfos.*

Prueba : Cada uno de ellos será topológicamente isomorfo a $(\mathbb{K}^n, \|\cdot\|_2)$.

□

Corolario 5 *Si $(X, \|\cdot\|)$, $(Y\|\cdot\|)$ son espacios normados sobre el mismo cuerpo, y X es de dimensión finita, entonces toda aplicación lineal $L : X \to Y$ es continua.*

Prueba : Sea n la dimensión de X. Si $J : \mathbb{K}^n \to X$ es el isomorfismo topológico que existe, según hemos visto, entre $\mathbb{K}^n$ y X.

$$\begin{tikzcd} X \arrow[r, "L"] & Y \\ \mathbb{K}^n \arrow[u, "J"] \arrow[ur, "L \circ J"'] & \end{tikzcd}$$

La aplicación
$$L \circ J : \mathbb{K}^n \to Y$$
es continua por ser lineal y estar definida en $(\mathbb{K}^n, \|\cdot\|_2)$. Entonces, $L = (L \circ J) \circ J^{-1}$ será también continua.

Teorema 16 *Sea* $(X, \|\cdot\|)$ *un espacio de Banach . Sea* $(Y, \|\cdot\|)$ *un espacio normado, y sea* $T : X \to Y$ *un isomorfismo topológico. Entonces* $(Y, \|\cdot\|)$ *es un espacio de Banach.*

Basta tener en cuenta que cualquier sucesión de Cauchy (y_n) en $(Y, \|\cdot\|)$ verificará que la sucesión $(T^{-1}(y_n))$ es también de Cauchy en $(X, \|\cdot\|)$, pues para cualesquiera enteros positivos p, q
$$\|T^{-1}(y_p) - T^{-1}(y_q)\| \leq \|T^{-1}\| \|y_p - y_q\|.$$
Luego, por ser $(X, \|\cdot\|)$ un espacio de Banach existirá $x \in X$ con
$$T^{-1}(y_n) \to x,$$
y como T es continua,
$$y_n = T(T^{-1}(y_n)) \to T(x),$$
luego existe en $(Y, \|\cdot\|)$ el límite de (y_n).

$\square$

Corolario 6 *Sea* V *un subespacio vectorial de dimensión finita del espacio normado* $(X, \|\cdot\|)$ *. Entonces* V *es un conjunto cerrado.*

Prueba : Basta tener en cuenta que si $V = 0_X$ entonces es cerrado trivialmente, y si $\dim(V) = n \in \mathbb{N}$, existe un isomorfismo topológico $J : \mathbb{K}^n \to V$. Teniendo en cuenta que $(\mathbb{K}^n, \|\cdot\|_2)$ es un espacio de Banach y aplicando el teorema anterior, se obtiene que V es un espacio de Banach. Si existiera una sucesión (v_n) en V que fuera convergente a un elemento de $X \setminus V$, por ser convergente en $(X, \|\cdot\|)$ sería de Cauchy (en V y no convergente a un elemento de V). Pero esto sería una contradicción con el hecho de que V es completo.

$\square$

2.5. Equivalencia de normas

Si los espacios normados $(X, \|\cdot\|)$, $(X, |.|)$ tienen la misma topología inducida por las normas $\|\cdot\|$, $|.|$, se dice que estas normas son (topológicamente) equivalentes. En tal caso las respectivas métricas asociadas son también equivalentes. Se demuestra con facilidad que

Teorema 17 *Las normas* $\|\cdot\|$, $|.|$ *sobre* X *son equivalentes si y sólo si existen constantes positivas* a, b *en* $\mathbb{R}$ *tales que, para cada* v *de* X

$$a\|v\| \leq |v| \leq b\|v\|.$$

Prueba : Si existen las constantes a, b del enunciado, $|y-x| < r \Rightarrow \|y-x\| < \frac{r}{a}$, luego $B_{|.|}(x, r) \subset B_{\|\cdot\|}(x, \frac{r}{a})$. Igualmente, $\|y-x\| < r \Rightarrow |y-x| < br$, luego $B_{\|\cdot\|}(x, r) \subset B_{|.|}(x, br)$. Por lo tanto, si G es $\|\cdot\|$-abierto no vacío, para todo $x \in G$ existe $\rho > 0$ con $B_{\|\cdot\|}(x, \rho) \subset G$, y entonces

$$B_{|.|}(x, a\rho) \subset B_{\|\cdot\|}(x, \frac{a\rho}{a}) \subset G$$

luego G es también $|.|$-abierto. Igual razonaríamos si G es $|.|$-abierto no vacío. Luego los conjuntos $\|\cdot\|$-abiertos y los $|.|$-abiertos son los mismos.

Recíprocamente, si las dos normas son equivalentes, $B := B_{\|\cdot\|}(0_X, 1)$ es $|.|$-abierto, luego 0 es punto $|.|$-interior de B, es decir, existe $\rho > 0$ tal que

$$B_{|.|}(0_X, \rho) \subset B.$$

Tomemos cualquier vector no nulo $x \in X$. Como $\frac{1}{2\rho|x|}x \in B_{|.|}(0_X, \rho)$, entonces

$$\left\| \frac{1}{2\rho|x|} x \right\| < 1$$

o sea, para todo $x \in X$,

$$\frac{1}{2\rho}\|x\| \leq |x|.$$

Intercambiando los papeles de las dos normas obtenemos la otra desigualdad, completándose la prueba de este modo.

$\square$

Esto puede decirse de otro modo:

Proposición 8 *Las normas* $\|\cdot\|$ *y* $|.|$ *sobre el espacio vectorial* X *son equivalentes si y sólo si la aplicación identidad es un isomorfismo topológico entre los espacios normados* $(X, \|\cdot\|)$ *y* $(X, |.|)$.

Prueba : Si las normas son equivalentes, la aplicación identidad y su inversa evidentemente son continuas, puesto que las dos normas inducen la misma topología sobre X.

Recíprocamente, si la identidad es un isomorfismo topológico, por la proposición 6 existen m, M constantes positivas tales que para todo $x \in X$,

$$m\|x\| \leq |Id(x)| \leq M\|x\|,$$

o sea

$$m\|x\| \leq |x| \leq M\|x\|,$$

y por el teorema anterior, las normas $\|.\|$ y $|.|$ son equivalentes en X.

$\square$

TEOREMA 18 *(Teorema de Tijonov). En un espacio vectorial de dimensión finita n, todas las normas son equivalentes.*

Prueba : Consideremos la bola cerrada unidad $B_{\|\cdot\|}$ de $(X, \|\cdot\|)$. Existe un isomorfismo topológico J entre $(X, \|\cdot\|)$ y $(\mathbb{K}^n, \|\cdot\|_2)$.

- $J(B_{\|\cdot\|})$ es cerrado en $(\mathbb{K}^n, \|\cdot\|_2)$:

  Si $(J(x_n))$ es una sucesión en $J(B_{\|\cdot\|})$ convergente a cierto $x \in \mathbb{K}^n$ como J^{-1} es continua

  $$x_n = J^{-1}(J(x_n)) \to J^{-1}(x)$$

  Pero como $x_n \in B_{\|\cdot\|}$, y $B_{\|\cdot\|}$ es cerrado, ha de ser $J^{-1}(x) \in B_{\|\cdot\|}$, luego $x \in J(B_{\|\cdot\|})$.

- $J(B_{\|\cdot\|})$ es acotado:

  Por la proposición 6, existen constantes $m, M > 0$ tales que para todo $x \in X$,

  $$m\|x\| \leq \|J(x)\|_2 \leq M\|x\|$$

  luego si $x \in B_{\|\cdot\|}$, $\|J(x)\| \leq M$.

Pero en $(\mathbb{K}^n, \|\cdot\|_2)$ los conjuntos cerrados y acotados son compactos. Luego $J(B_{\|\cdot\|})$ es compacto. Pero entonces $B_{\|\cdot\|} = J^{-1}(J(B_{\|\cdot\|}))$ es también compacto.

Sea $|\cdot|$ una norma sobre X. Toda norma sobre un espacio vectorial es una aplicación real continua sobre el mismo, o sea, $v \mapsto |v|$ es continua en X, y en particular en el compacto $B_{\|\cdot\|}$. Luego existe $w \in B$ tal que para todo $v \in B_{\|\cdot\|}$

$$|v| \leq |w|.$$

(Notemos que no puede ser $w = 0_X$). En particular, si $x \neq 0_X$, el vector $\frac{1}{\|x\|}x \in B_{\|\cdot\|}$, luego

$$|\frac{1}{\|x\|}x| \leq |w|.$$

o sea,

$$|x| \leq |w|\|x\|.$$

Repitiendo el mismo argumento, ahora con el espacio $(X, |.|)$ y su bola unidad $B_{|\cdot|}$, y usando la continuidad de $\|\cdot\|$, también obtendríamos

$$\|x\| \leq \|v\|\,|x|.$$

para todo $x \in X$ y cierto $v \neq 0_X$ de $B_{|\cdot|}$, la bola unidad de $(X, |.|)$.

Combinando estas desigualdades tenemos

$$\frac{1}{\|v\|}\|x\| \leq |x| \leq |w|\|x\|$$

para todo x de X, o en otras palabras, existen constantes positivas a, b tales que

$$a\|x\| \leq |x| \leq b\|x\|.$$

Luego la identidad entre $(X, \|\cdot\|)$ y $(X, |.|)$ es un isomorfismo, y las normas $\|\cdot\|$ son equivalentes en X.

□

2.6. Los conjuntos compactos en espacios de dimensión no finita

En el espacio euclídeo n-dimensional, los conjuntos cerrados y acotados (y en particular las bolas cerradas centradas en el origen) son compactos. En dimensión no finita, esto no es así.

Ejemplo 14 En el espacio $(\ell_2, \|.\|_2)$ los vectores $(e^{(n)})_{n=1}^{\infty}$, con

$$e^{(n)} := (0, \ldots, 0, \overbrace{1}^{n}, 0, \ldots)$$

forman una sucesión en el conjunto $B := \{x \in \ell_2 : \|x\|_2 \leq 1\}$ que es cerrado y acotado. Si B fuera compacto $(e^{(n)})_{n=1}^{\infty}$ admitiría una subsucesión convergente a un punto de B. Esta subsucesión convergente sería de Cauchy, y por tanto sus términos estarían entre sí a distancia arbitrariamente pequeña para índices lo bastante grandes. Pero esto es absurdo porque si $i \neq j$

$$\|e^{(i)} - e^{(j)}\|_2 = \sqrt{2}.$$

Se cumple además el siguiente teorema, de enunciado sorprendente (una condición topológica da una conclusión algebraica):

Teorema 19 *(Teorema de F. Riesz para espacios normados). Si un espacio normado $(X, \|\cdot\|)$ tiene un entorno compacto del origen entonces es de dimensión finita.*

Prueba : Si existe un entorno del origen W que sea compacto, entonces existe $r > 0$ tal que la bola cerrada $B[0_X, r]$ está contenida en W. Toda bola cerrada es un conjunto cerrado, y todo cerrado contenido en un compacto es a su vez un conjunto compacto. Luego $B[0_X, r]$ es un conjunto compacto. La aplicación $x \mapsto \frac{1}{r}x$ es continua en X, y transforma $B[0_X, r]$ en $B[0_X, 1]$, luego si W es compacto, $B_X := B[0_X, 1]$ sería compacto.

La colección de bolas abiertas $\{B(v, 1/2) : v \in B_X\}$ es un recubrimiento abierto de B_X, luego existen $v_1, \ldots, v_n$ en B_X tales que

$$B_X \subset B(v_1, 1/2) \cup \ldots \cup B(v_n, 1/2).$$

Sea M el subespacio de X generado por los vectores $v_1, \ldots, v_n$, que será un conjunto cerrado en X, (por ser M un subespacio de dimensión finita). Afirmamos que $X = M$, y el teorema estará así demostrado.

Para llegar a un absurdo, supongamos que exista x en $X\backslash M$. Por ser M cerrado con $x \notin M$, se tendrá que $d(x,M) > 0$. Por definición de $d(x,M)$ existirá y en M tal que

$$d(x,M) \leq \|x-y\| < \frac{3}{2}d(x,M).$$

Sea $z = \frac{1}{\|x-y\|}(x-y)$. Al ser $\|z\| = 1$, $z \in B_X$, luego z pertenece a alguna de las bolas abiertas que recubren a B_X, por ejemplo $z \in B(v_k, 1/2)$. Entonces

$$x = y + \|x-y\|z = y + \|x-y\|v_k + \|x-y\|(z-v_k).$$

Así pues puede escribirse

$$x = w + \|x-y\|(z-v_k)$$

con $w \in M$, de donde

$$0 < d(x,M) \leq \|x-w\| = \|x-y\|\|z-v_k\| < \frac{3}{2}d(x,M)\frac{1}{2},$$

que es una contradicción. Luego $X\backslash M$ es vacío, y entonces $X = M$ es de dimensión finita.

$\square$

Una parte de la prueba anterior puede refinarse un poco para demostrar el siguiente resultado:

Lema 2 *Lema de F. Riesz Sean M Z subespacios vectoriales de un espacio normado X. Si M es cerrado y subespacio estricto de Z entonces para cada $\theta \in (0,1)$ existe $z \in Z$ con $\|z\| = 1$ y con $\|z-y\| \geq \theta$ para todo $y \in M$.*

Prueba: Como M es distinto de Z existe $x \in Z\backslash M$. Como M es cerrado, $d(x,Y) > 0$. Por definición de distancia existe $y_0 \in M$ tal que

$$d(x,M) \leq \|x-y_0\| < \frac{d(x,M)}{\theta}.$$

Sea $z = \frac{1}{\|x-y_0\|}(x-y_0)$. Desde luego $\|x-y_0\| \neq 0$, pues $x \notin M$. Además $\|z\| = 1$ y para cualquier vector $y \in M$ se tiene que $y_0 - \|x-y_0\|y \in M$, de donde

$$\begin{aligned}
\|z-y\| &= \left\|\tfrac{1}{\|x-y_0\|}(x-y_0) - y\right\| \\
&= \tfrac{1}{\|x-y_0\|}\,\|x-y_0-\|x-y_0\|y\| \\
&= \tfrac{1}{\|x-y_0\|}\,\|x-(y_0-\|x-y_0\|y)\| \\
&\geq \tfrac{1}{\|x-y_0\|}d(x,Y) \\
&> \tfrac{d(x,M)}{\frac{d(x,M)}{\theta}} \\
&= \theta,
\end{aligned}$$

lo que completa la prueba.

$\square$

3. ÁLGEBRAS DE OPERADORES

Las aplicaciones lineales y continuas de un espacio normado en sí mismo suelen llamarse **operadores**. El espacio vectorial $\mathcal{B}(X,X)$ con la norma de las aplicaciones lineales continuas (en este caso también llamada *norma de los operadores*), tiene una estructura especialmente rica porque la composición $S \circ T$ de dos elementos $S,T \in \mathcal{B}(X,X)$ es un elemento del mismo espacio. Además,

$$\|S \circ T\| \leq \|S\|\|T\|, \quad \|Id_X\| = 1.$$

La composición de aplicaciones tiene buenas propiedades de compatibilidad con la estructura vectorial. Por ejemplo, si T_1, T_2, T_3 son operadores de $\mathcal{B}(X,X)$, y λ, μ son escalares,

a) $(T_1 + T_2) \circ T_2 = (T_1 \circ T_3) + (T_2 \circ T_3)$.

b) $T_1 \circ (T_2 + T_3) = (T_1 \circ T_2) + (T_1 \circ T_3)$.

c) $(\lambda T_1) \circ (\mu T_2) = \lambda\mu(T_1 \circ T_2)$.

PROPOSICIÓN 9 *Fijado un operador $T \in \mathcal{B}(X,X)$, las aplicaciones*

$$\mathcal{R}_T : \mathcal{B}(X,X) \to \mathcal{B}(X,X) \ \ \mathcal{L}_T : \mathcal{B}(X,X) \to \mathcal{B}(X,X),$$

dadas respectivamente por

$$\mathcal{R}_T(L) = L \circ T, \ \ \mathcal{L}_T(L) = T \circ L,$$

son (lineales y) continuas.

Sea (L_n) una sucesión de operadores de $\mathcal{B}(X,X)$, convergente en este espacio a $L \in \mathcal{B}(X,X)$.

$$\|\mathcal{R}_T(L_n) - \mathcal{R}_T(L)\| = \|(L_n \circ T) - (L \circ T)\| = \|(L_n - L) \circ T\| \leq \|L_n - L\|\|T\|.$$

Como $\|T\|$ es constante (no depende de n), y $\|L_n - L\| \to 0$, obtenemos inmediatamente que $\|\mathcal{R}_T(L_n) - \mathcal{R}_T(L)\| \to 0$, o sea que $(\mathcal{R}_T(L_n))$ converge a $\mathcal{R}_T(L)$ en $(\mathcal{B}(X,X), \|.\|)$.

El caso de $\mathcal{L}_T$ se prueba del mismo modo.

□

DEFINICIÓN 13 *Si un elemento $T \in \mathcal{L}(X,X)$ verifica que existe $T^{-1} \in \mathcal{L}(X,X)$, se dice que T es algebraicamente invertible.*

Cuando el espacio vectorial X es de dimensión finita, propiedades bien conocidas de Algebra Lineal, especialmente la *Fórmula de Grassman*

$$\dim\,(T(X)) + \dim\,(\ker(T)) = \dim\,(X)$$

nos dicen que los siguientes enunciados son equivalentes:

a) T es uno a uno (inyectiva).

b) $\ker(T) = \{0_X\}$.

c) dim $(T(X)) = $ dim (X).

d) T es sobre (exhaustiva), o sea $T(X) = X$.

e) T es algebraicamente invertible.

Sin embargo, cuando X no es de dimensión finita, en general estas equivalencias no se cumplen.

El siguiente ejemplo puede contribuir a distinguir casos en que estos enunciados no equivalen.

Ejemplo 15 *En ℓ_2 definimos los siguientes operadores:*

$$L : \ell_2 \to \ell_2, \quad L(x_1, x_2, \ldots) := (x_2, x_3, \ldots).$$

$$R : \ell_2 \to \ell_2, \quad R(x_1, x_2, \ldots) := (0, x_1, x_2, \ldots).$$

Se tiene que L es sobre, pero no inyectiva. También se tiene que R es inyectiva, pero no sobre.

Por otra parte,

$$L \circ R = Id_{\ell_2} \neq R \circ L.$$

En un lenguaje algebraico podríamos decir que R admite un inverso por la izquierda, pero no es algebraicamente invertible. Tampoco lo es L porque no es inyectiva.

Podemos notar, que respecto a la norma ordinaria de ℓ_2 las dos aplicaciones son continuas.

DEFINICIÓN 14 *Si un elemento $T \in \mathcal{B}(X, X)$ verifica que existe $T^{-1} \in \mathcal{B}(X, X)$, se dice que T es topológicamente invertible.*

En el caso de ser X de dimensión finita, bastará que exista T^{-1} para que sea lineal y continua, es decir que si T es algebraicamente invertible es topológicamente invertible, pero en dimensión infinita esto no es tan sencillo.

Ejemplo 16 Sea φ es espacio de las sucesiones de números reales con todos sus términos nulos excepto a lo sumo un número finito de ellos, con la norma del supremo.

Definimos la aplicación lineal $T : \varphi \to \varphi$ por

$$T(x_1, x_2, \ldots) = (x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \ldots).$$

Es claro que T está bien definida, y es lineal y continua con $\|T\| \leq 1$. Es decir que $T \in \mathcal{B}(\varphi, \varphi)$.

Pero existe T^{-1} y no es continua. En efecto, es inmediato comprobar que

$$T^{-1}(y_1, y_2, \ldots) = (y_1, 2y_2, 3y_3, \ldots),$$

pues para todo $x \in \varphi$

$$T^{-1}T(x) = T^{-1}((x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \ldots)) = x$$

y también para todo $y \in \varphi$

$$TT^{-1}(y) = T((y_1, 2y_2, 3y_3, \ldots)) = y.$$

También es inmediato comprobar que T^{-1} no es continua porque no es acotada en la bola unidad de $(\varphi, \|.\|_\infty)$: Los vectores

$$e^{(n)} := (0, \overset{n-1}{\ldots}, 0, 1, 0 \ldots)$$

pertenecen a esta bola cerrada unidad, pero

$$\|T^{-1}(e^{(n)})\|_\infty = \|(0, \overset{n-1}{\ldots}, 0, n, 0 \ldots)\|_\infty = n.$$

Un ejemplo como el anterior es sólo posible porque $(\varphi, \|.\|_\infty)$ no es un espacio de Banach .

TEOREMA 20 *(Criterio de invertibilidad). Sea $T \in \mathcal{B}(X, X)$.*

Si la serie $\sum_{n=0}^{\infty} T^n$ es convergente, entonces $Id - T$ es invertible, y

$$(Id - T)^{-1} = \sum_{n=0}^{\infty} T^n.$$

Prueba : En este enunciado se usa el convenio de notación $T^0 = Id$.

Como $S = \sum_{n=0}^{\infty} T^n$ existe en $\mathcal{B}(X, X)$, es claro que

$$S = \lim_k \sum_{n=0}^{k} T^n = \lim_k \sum_{n=0}^{k+1} T^n$$

luego

$$0_{\mathcal{B}(X,X)} = \lim_k \sum_{n=0}^{k} T^n - \lim_k \sum_{n=0}^{k+1} T^n = \lim_k (T^{(k+1)}).$$

Entonces, teniendo en cuenta la continuidad de la composición por la izquierda con una aplicación constante,

$$\begin{aligned}
(Id - T) \circ S &= (Id - T) \circ (\textstyle\sum_{n=0}^{\infty} T^n) \\
&= (Id - T) \circ (\lim_k \textstyle\sum_{n=0}^{k} T^n) \\
&= \lim_k ((Id - T) \circ \textstyle\sum_{n=0}^{k} T^n) \\
&= \lim_k (Id \circ \textstyle\sum_{n=0}^{k} T^n - T \circ \sum_{n=0}^{k} T^n) \\
&= \lim_k (\textstyle\sum_{n=0}^{k} T^n - \sum_{n=0}^{k} T^{(n+1)}) \\
&= \lim_k (T^0 - T^{(k+1)}) = Id.
\end{aligned}$$

Por otra parte, teniendo en cuenta la continuidad de la composición por la derecha con una aplicación constante,

$$\begin{aligned} S \circ (Id - T) &= (\textstyle\sum_{n=0}^{\infty} T^n) \circ (Id - T) \\ &= (\lim_k \textstyle\sum_{n=0}^{k} T^n) \circ (Id - T) \\ &= \lim_k ((\textstyle\sum_{n=0}^{k} T^n) \circ (Id - T)) \\ &= \lim_k ((\textstyle\sum_{n=0}^{k} T^n) \circ Id - (\textstyle\sum_{n=0}^{k} T^n) \circ T) \\ &= \lim_k (\textstyle\sum_{n=0}^{k} T^n - \textstyle\sum_{n=0}^{k} T^{(n+1)}) \\ &= \lim_k (T^0 - T^{(k+1)}) = Id. \end{aligned}$$

□

Corolario 7 *Si* $(X, \|\cdot\|)$ *es un espacio de Banach y* $T \in \mathcal{B}(X, X)$*, verifica que* $\|T\| < 1$ *entonces* $Id - T$ *es invertible, con*

$$\|(Id - T)^{-1}\| \leq \frac{1}{1 - \|T\|}.$$

Si $\|I - T\| < 1$ *entonces* T *es invertible, con*

$$\|T^{-1}\| \leq \frac{1}{1 - \|Id - T\|}.$$

Prueba : Como la serie numérica $\sum_{n=0}^{\infty} \|T^n\|$ es geométrica convergente (por ser $\|T\| < 1$ y por ser $\|T^n\| \leq \|T\|^n$), y $\mathcal{B}(X, X)$ es un espacio de Banach , la serie $\sum_{n=0}^{\infty} T^n$ será convergente en este espacio, y podemos aplicar el teorema anterior: Su suma S será el operador inverso de $Id - T$, y

$$\|S\| = \left\| \sum_{n=0}^{\infty} T^n \right\| \leq \sum_{n=0}^{\infty} \|T^n\| \leq \sum_{n=0}^{\infty} \|T\|^n \frac{1}{1 - \|T\|}.$$

Si $\|Id - T\| < 1$, entonces por lo que acabamos de ver $Id - (Id - T) = T$ será invertible, con

$$(Id - (Id - T))^{-1} = T^{-1} = \sum_{n=0}^{\infty} (Id - T)^n,$$

y aplicando el mismo razonamiento que antes,

$$\|T^{-1}\| = \left\| \sum_{n=0}^{\infty} (Id - T)^n \right\| \leq \sum_{n=0}^{\infty} \|(Id - T)^n\| \leq \sum_{n=0}^{\infty} \|Id - T\|^n = \frac{1}{1 - \|Id - T\|}.$$

□

COROLARIO 8 *Sea* $(X, \|\cdot\|)$ *es un espacio de Banach y* $T \in \mathcal{B}(X, X)$*, con* $\|T\| < 1$*. Sea* y *un elemento dado de* X*.*

El problema de encontrar $x \in X$ *tal que*

$$x - Tx = y$$

tiene solución $x \in X$ *dada por*

$$x = \sum_{n=0}^{\infty} T^n(y).$$

Prueba : Si $\|T\| < 1$ el operador $Id - T$ es invertible y sabemos por el teorema anterior que

$$(Id - T)^{-1} = \sum_{n=0}^{\infty} T^n.$$

Si $x - Tx = y$, es decir si $(Id - T)(x) = y$, entonces

$$x = (Id - T)^{-1}(Id - T)(x) = (Id - T)^{-1}(y) = \sum_{n=0}^{\infty} T^n(y).$$

$\square$

Ejemplo 17 Consideremos la ecuación integral en $\mathcal{C}([0,1])$, dada por

$$x(s) - \int_0^1 \frac{1}{2}\cos(st)x(t)dt = y(s). \tag{8}$$

Es decir, fijado $y \in \mathcal{C}([0,1])$, buscamos una función continua $x \in \mathcal{C}([0,1])$ tal que para todo $s \in [0,1]$.

Si $T : \mathcal{C}([0,1]) \to \mathcal{C}([0,1])$ es la aplicación que transforma el elemento $x \in \mathcal{C}([0,1])$ en el elemento $Tx \in \mathcal{C}([0,1])$

$$Tx(s) = \int_0^1 \frac{1}{2}\cos(st)x(t)dt,$$

entonces es claro que T está bien definida y es lineal, y que para cada $s \in [0,1]$

$$|Tx(s)| \leq \int_0^1 |\frac{1}{2}\cos(st)x(t)|dt \leq \int_0^1 \frac{1}{2}|x(t)|dt \leq \frac{1}{2}\|x\|_\infty,$$

luego

$$\|Tx\|_\infty = \sup\{|Tx(s)| : s \in [0,1]\} \leq \frac{1}{2}\|x\|_\infty.$$

De aquí se deduce que T es continua. La ecuación dada 8 puede escribirse

$$(Id - T)(x) = y.$$

Como $\|T\| \leq \frac{1}{2}$ el teorema anterior nos garantiza que existe $(Id-T)^{-1}$ y que $(Id-T)^{-1} = \sum_{n=0}^{\infty} T^n$. Por tanto,

$$x = (Id - T)^{-1}((Id - T)(x)) = (Id - T)^{-1}(y),$$

y formalmente

$$x = \left[\sum_{n=0}^{\infty} T^n\right](y).$$

Es decir, las sumas parciales

$$s_k := \left[\sum_{n=0}^{k} T^n\right](y)$$

forman una sucesión de elementos de $\mathcal{C}([0,1])$ que converge (en $\mathcal{C}([0,1])$, o sea uniformemente) a la solución de la ecuación x. Observar que $s_0 = T^0(y) = Id(x) = y$,

$$s_1 = (T^0 + T)(y) = y + T(y) = s_0 + T(s_0),$$

y si

$$s_k = (T^0 + T + \ldots + T^k)(y) = y + T(s_{k-1}),$$

se tiene que

$$s_{k+1} = (T^0 + T + \ldots + T^k + T^{k+1})(y) = y + T((T^0 + T + \ldots + T^k)(y)) = y + T(s_k).$$

3.1. El conjunto de los operadores invertibles

En $\mathcal{B}(X,X)$ podemos considerar el subconjunto $\mathcal{G}$ formado por los operadores invertibles. Podemos resumir algunas de sus propiedades en el siguiente

Teorema 21 *Sea* $(X, \|\cdot\|)$ *un espacio de Banach . Si* $A \in \mathcal{G}$ *y* B *es cualquier operador de* $\mathcal{B}(X,X)$ *con* $\|A - B\| < \frac{1}{\|A^{-1}\|}$ *entonces:*

a) $B \in \mathcal{G}$.

b) $\|B^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\|A-B\|}$.

c) $\|A^{-1} - B^{-1}\| \leq \frac{\|A^{-1}\|^2\|A-B\|}{1-\|A^{-1}\|\|A-B\|}$.

Prueba : Como $Id - (A^{-1} \circ B) = (A^{-1} \circ A) - (A^{-1} \circ B) = A^{-1} \circ (A - B)$, y $\|A^{-1} \circ (A - B)\| \leq \|A^{-1}\|\|A - B\| < 1$ tenemos que $\|Id - (A^{-1} \circ B)\| < 1$. Podemos aplicar entonces un corolario anterior (7), por el cual sabemos que entonces el operador $A^{-1} \circ B$ es invertible. Si A y $A^{-1} \circ B$ son invertibles también lo será $A \circ (A^{-1} \circ B) = B$, lo que prueba (a). Además el mismo corolario nos da la acotación

$$\|(A^{-1} \circ B)^{-1}\| \leq \frac{1}{1 - \|Id - (A^{-1} \circ B)\|},$$

es decir,

$$\|B^{-1} \circ A\| \leq \frac{1}{1 - \|(A^{-1} \circ A) - (A^{-1} \circ B)\|}.$$

Pero $\|(A^{-1} \circ A) - (A^{-1} \circ B)\| = \|A^{-1} \circ (A - B)\| \leq \|A^{-1}\|\|A - B\|$, luego

$$1 - \|Id - (A^{-1} \circ B)\| \geq 1 - \|A^{-1}\|\|A - B\|,$$

y tomando inversos,

$$\|B^{-1} \circ A\| \leq \frac{1}{1 - \|Id - (A^{-1} \circ B)\|} \leq \frac{1}{1 - \|A^{-1}\|\|A - B\|}.$$

De aquí resulta inmediatamente que

$$\|B^{-1}\| = \|(B^{-1} \circ A) \circ A^{-1}\|$$

$$\leq \|B^{-1} \circ A\|\|A^{-1}\| \leq \frac{1}{1 - \|A^{-1}\|\|A - B\|}\|A^{-1}\|,$$

lo que prueba (b).

Para ver (c) tenemos en cuenta que

$$B^{-1} - A^{-1} = A^{-1} \circ (A - B) \circ B^{-1}.$$

De aquí resulta, aplicando (b) que

$$\|B^{-1} - A^{-1}\| \leq \|A^{-1}\|\|A - B\|\|B^{-1}\| \leq \frac{1}{1 - \|A^{-1}\|\|A - B\|}\|A^{-1}\|\|A^{-1}\|\|A - B\|.$$

lo que prueba (c).

□

COROLARIO 9 *$\mathcal{G}$ es abierto en $(\mathcal{B}(X, X), \|.\|)$.*

3.2. Aplicación: Ecuaciones integrales de Fredhölm

Si $[a, b]$ es un intervalo compacto de la recta real, y $k : [a, b] \times [a, b] \to \mathbb{R}$ es una función suficientemente *buena*, podemos considerar el problema de encontrar una función incógnita $x : [a, b] \to \mathbb{R}$, verificando para cada $s \in [a, b]$

$$\boxed{x(s) - \int_a^b k(s,t)x(t)dt \quad = \quad y(s)}$$

donde $y \in \mathcal{C}([a, b])$ es un dato. Este, como muchos problemas parecidos, es un caso particular de *ecuación integral*[4] (que se llama así porque se plantea a partir de una igualdad que cumple una integral en que interviene la función incógnita).

Si $T : \mathcal{C}([a, b]) \to \mathcal{C}([a, b])$ es el operador que transforma $x \in \mathcal{C}([a, b])$ en $Tx \in \mathcal{C}([a, b])$ dada por

$$Tx(s) := \int_a^b k(s,t)x(t)dt$$

la ecuación integral anterior puede escribirse con notación funcional más condensada como

$$(Id - T)x = y.$$

[4] Concretamente esta se llama **Ecuación integral de Fredhölm de segunda especie**.

Es claro que T $\mathcal{C}([a,b])$ en $\mathcal{C}([a,b])$ porque, si $x \in \mathcal{C}([a,b])$,

$$|Tx(s_1) - Tx(s_2)| = |\int_a^b (k(s_1,t) - k(s_2,t))x(t)dt| \leq \int_a^b |k(s_1,t) - k(s_2,t)||x(t)|dt$$

Entonces como k es uniformemente continua en $[a,b] \times [a,b]$, dado $\varepsilon > 0$ existe $\delta > 0$ tal que si $\|(s_1,t_1) - (s_2,t_2)\| < \delta$ entonces $|k(s_1,t_1) - k(s_2,t_2)| < \varepsilon$. En particular, si $s_1, s_2 \in [a,b]$ con $|s_1 - s_2| < \delta$, para todo $t \in [a,b]$, $\|(s_1,t) - (s_2,t)\| < \delta$, luego

$$|Tx(s_1) - Tx(s_2)| \leq \int_a^b |k(s_1,t) - k(s_2,t)||x(t)|dt \leq \int_a^b \varepsilon\|x\|_\infty dt = \varepsilon(b-a)\|x\|_\infty$$

Esto prueba que la función $s \mapsto Tx(s)$ es (uniformemente) continua en $[a,b]$, y por tanto que T aplica $\mathcal{C}([a,b])$ en $\mathcal{C}([a,b])$.

Desde luego T es lineal.
Siendo k continua en $[a,b] \times [a,b]$, $M := \sup\{|k(s,t)| : (s,t) \in [a,b] \times [a,b]\} < \infty$.
Si $x \in \mathcal{C}([a,b])$ y $s \in [a,b]$,

$$|Tx(s)| \leq \int_a^b |k(s,t)||x(t)|dt \leq M\|x\|_\infty(b-a).$$

De aquí se sigue que

$$\|Tx\|_\infty \leq M(b-a)\|x\|_\infty$$

luego el operador T es continuo con $\|T\| \leq M(b-a)$.
Veremos a continuación un modo más preciso de acotar $\|T\|$:

Proposición 10 *Norma de los operadores de Fredhölm en* $\mathcal{C}([a,b])$
Consideremos el operador $T : \mathcal{C}([a,b]) \to \mathcal{C}([a,b])$ *dado por* $x \mapsto Tx$*, donde*

$$Tx(s) = \int_a^b k(s,t)x(t)dt$$

siendo k *una función continua en* $[a,b] \times [a,b]$*. Entonces*

$$\boxed{\|T\| \;=\; \max_{s\in[a,b]} \int_a^b |k(s,t)|dt.}$$

Prueba : Sea $M := \max_{s\in[a,b]} \int_a^b |k(s,t)|dt$. Para cada $x \in \mathcal{C}([a,b])$,

$$|Tx(s)| = |\int_a^b k(s,t)x(t)dt| \leq \int_a^b |k(s,t)||x(t)|dt \leq \|x\|_\infty M,$$

luego,

$$\|Tx\|_\infty \leq M\|x\|_\infty$$

lo que prueba que $\|T\| \leq M$.

Para ver la desigualdad contraria, tengamos en cuenta que (por un argumento muy parecido al usado para ver que Tx es continua) la función $s \mapsto \int_a^b |k(s,t)|dt$ es continua en el compacto $[a,b]$. Por tanto existe $s_0 \in [a,b]$ tal que

$$M = \int_a^b |k(s_0,t)|dt.$$

Sea ahora el funcional $f : \mathcal{C}([a,b]) \to \mathbb{R}$ dado por

$$f(x) := \int_a^b k(s_0,t)x(t)dt$$

Por lo visto en el ejemplo 9 sabemos que $\|f\| = \int_a^b |k(s_0,t)|dt = M$. Por definición de $\|f\|$, dado cualquier $\varepsilon > 0$ existe $\tilde{x} \in \mathcal{C}([a,b])$ con $\|\tilde{x}\|_\infty \leq 1$ y con

$$f(\tilde{x}) \geq \|f\| - \varepsilon$$

y entonces

$$\|T\| \geq \|T(\tilde{x})\|_\infty \geq T(\tilde{x})(s_0) := \int_a^b k(s_0,t)\tilde{x}(t)dt = f(\tilde{x}) \geq M - \varepsilon.$$

Como ε es arbtitrario, obtenemos que $\|T\| \geq M$, lo que completa la prueba.

$\square$

Ejemplo 18 *Sea T el operador de Fredhölm en $\mathcal{C}([0,1])$ con núcleo $s(1-e^{st})$, es decir*

$$Tx(t) = \int_0^1 s(1-e^{st})x(t)dt.$$

Observar que, para $(s,t) \in [0,1]\times[0,1]$, como $0 \leq st \leq 1$, $1 = e^0 \leq e^{st} \leq e^1$ es decir que $|1-e^{st}| = e^{st}-1$. Además la función $g(s,t) = e^{st}-1$ tiene derivadas parciales que no se anulan en el interior de $[0,1]\times[0,1]$, luego de aquí resulta fácilmente que, para $(s,t) \in [0,1]\times[0,1]$

$$|k(s,t)| \leq |1-e^{st}| = g(s,t) \leq g(1,1) = e-1 =: M,$$

y para esta acotación, $M(b-a) = (e-1)(1-0) = e-1$ no es menor que 1.

No obstante, la norma del operador de Fredhölm $T : \mathcal{C}([0,1]) \to \mathcal{C}([0,1])$ dado por

$$Tx(s) = \int_0^1 s(1-e^{st})x(t)dt$$

es

$$\|T\| = \max\{\int_0^1 |s(1-e^{st})|dt : 0 \leq s \leq 1\}.$$

Como $|1-e^{st}| = e^{st}-1$, queda

$$\begin{aligned} \int_0^1 |s(1-e^{st})|dt &= \int_0^1 s(e^{st}-1)dt \\ &= s\int_0^1 e^{st}dt - s \\ &= s\left[\frac{e^{st}}{s}\right]_0^1 - s \\ &= e^s - e^0 - s. \end{aligned}$$

Es muy fácil ver que la función $s \mapsto e^s - s - 1$ es creciente en $[0,1]$, luego

$$\|T\| = \max\{\int_0^1 |s(1-e^{st})|dt : 0 \leq s \leq 1\} = e^1 - 2 < 1.$$

Por tanto la estimación $\|T\| \leq M(b-a)$ era muy grosera en este caso.

TEOREMA 22 *Si $k : [a,b]\times[a,b] \to \mathbb{R}$ es una función continua verificando que*

$$\max_{s\in[a,b]} \int_a^b |k(s,t)|dt < 1.$$

entonces la ecuación integral de Fredhölm de segunda especie

$$x(s) - \int_a^b k(s,t)x(t)dt = y(s)$$

tiene una solución $x \in \mathcal{C}([a,b]$, dada por

$$x = [\sum_{n=0}^{\infty} T^n](y).$$

Prueba : Si existe x verificando la ecuación $x - T(x) = y$, entonces $(Id-T)(x) = y$. Como entonces $\|T\| = \max_{s\in[a,b]} \int_a^b |k(s,t)|dt < 1$, la aplicación $Id - T$ es invertible y la serie $\sum_{n=0}^{\infty} T^n$ es convergente en $\mathcal{B}(X,X)$ a $(Id-T)^{-1}$, luego

$$x = (Id-T)^{-1}(y) = [\sum_{n=0}^{\infty} T^n](y) = \lim_k [\sum_{n=0}^{k} T^n(y)] = \lim_k x_k.$$

Si tomamos $x_0 = y$, y, en general

$$x_n := \sum_{i=0}^{n} T^i(y)$$

Los vectores x_n podemos entonces considerarlos como aproximaciones sucesivas de la solución x.

$\square$

Además este método iterativo nos permite calcular una cota superior del *error* de la aproximación n-sima:

$$\|x - x_n\| = \|\sum_{k=0}^{\infty} T^k(y) - \sum_{k=0}^{n} T^k(y)\| = \|\sum_{k=n+1}^{\infty} T^k(y)\|$$

$$\leq \sum_{k=n+1}^{\infty} \|T^k(y)\| \leq \sum_{k=n+1}^{\infty} \|T^k\|\|y\| = \frac{\|T\|^{n+1}}{1-\|T\|}\|y\|.$$

No obstante, este teorema no es operativo porque el cálculo efectivo de las aproximaciones x_n puede ser difícil. (Entre otras cosas, requiere conocer las iteradas del operador T).

En general, si $M := \sup\{|K(s,t), (s,t) \in [a,b] \times [a,b]\}$, sabemos que para el operador T que $\|T\| \leq M(b-a)$, y que esta desigualdad a veces puede ser estricta (ver ejemplo 18). Sin embargo, en ocasiones resulta más sencillo calcular $M(b-a)$.

Corolario 10 *Si $k : [a,b] \times [a,b] \to \mathbb{R}$ es una función continua. La ecuación integral de Fredhölm de segunda especie*

$$x(s) - \int_a^b k(s,t)x(t)dt = y(s)$$

tiene una solución $x \in \mathcal{C}([a,b]$, dada por

$$x = [\sum_{n=0}^{\infty} T^n](y)$$

si $M(b-a) < 1$.

Basta ver que $\|T\| \leq M(b-a)$, y si $M(b-a) < 1$ puede aplicarse el teorema anterior. □

Ejemplo 19 *La ecuación integral*

$$x(s) - \int_0^{\frac{1}{2}} \frac{1}{2}(s+t)x(t)dt = 2s$$

en $\mathcal{C}([0,\frac{1}{2}])$ tiene solución.

Su núcleo $k(s,t) = \frac{1}{2}(s+t)$ tiene en $[0,\frac{1}{2}] \times [0,\frac{1}{2}]$ la cota superior $\frac{1}{2}$ pues

$$|\frac{1}{2}(s+t)| \leq \frac{1}{2}[|s| + |t|] \leq \frac{1}{2}$$

y $\frac{1}{2}(b-a) = \frac{1}{4} < 1$. Por tanto, el teorema anterior nos asegura que $\|T\| \leq \frac{1}{4} < 1$ donde, para cada $x \in \mathcal{C}([0,\frac{1}{2}]$

$$Tx(s) := \int_0^{\frac{1}{2}} \frac{1}{2}(s+t)x(t)dt.$$

Observar que el valor exacto de $\|T\|$ es, según la proposición anterior,

$$\|T\| = \max_{s\in[0,1/2]} \int_0^{1/2} |k(s,t)|dt = \max_{s\in[0,1/2]} \int_0^{1/2} \frac{s+t}{2}dt = \max_{s\in[0,1/2]} \frac{4s+1}{16} = \frac{3}{16} < \frac{1}{4}.$$

Pero si $y(s) := 2s$ entonces

$$Ty(s) = \int_0^{\frac{1}{2}} \frac{1}{2}(s+t)2tdt = \frac{s}{8} + \frac{1}{24}.$$

$$T^2y(s) = T[Ty](s) = \int_0^{\frac{1}{2}} \frac{1}{2}(s+t)[\frac{t}{8} + \frac{1}{24}]dt = \frac{7s+2}{384}.$$

$$T^3y(s) = T[T^2y](s) = \int_0^{\frac{1}{2}} \frac{1}{2}(s+t)[\frac{7s+2}{384}]dt = \frac{45s+13}{18432}.$$

Por tanto, si ponemos $x_n := \sum_{i=0}^n T^i y$, tendremos que, en particular,

$$\begin{aligned} x_3(s) &= y(s) + Ty(s) + T^2y(s) + T^3y(s) \\ &= 2s + \frac{3s+1}{24} + \frac{7s+2}{384} + \frac{45s+13}{18432} \end{aligned}$$

es una aproximación de la solución $x(s)$ de esta ecuación.

Podemos además tener una cota del *error* de esta aproximación pues

$$\|x - x_3\|_\infty = \left\| \sum_{i=0}^{\infty} T^i y - \sum_{i=0}^{3} T^i y \right\|_\infty = \left\| \sum_{i=4}^{\infty} T^i y \right\|_\infty$$

$$\leq \sum_{i=4}^{\infty} \|T^i\| \|y\|_\infty \leq \|y\|_\infty \sum_{i=4}^{\infty} \|T\|^i \leq 1 \frac{\|T\|^4}{1-\|T\|} = \frac{(3/16)^4}{1-3/16} = \frac{81}{53248} \approxeq 0{,}001521183.$$

Vamos a ver a continuación un método algo más práctico para resolver de forma explícita estas ecuaciones integrales.

Definición 15 *Si*

$$x(s) - \int_a^b k(s,t)x(t)dt = y(s)$$

es una ecuación de Fredhölm de segunda especie, la sucesión de funciones $k_n : [a,b] \times [a,b] \to \mathbb{R}$ *dada recurrentemente por*

$$k_1(s,t) := k(s,t),\ k_{n+1}(s,t) := \int_a^b k_1(s,u)k_n(u,t)du$$

se llama sucesión de núcleos iterados de la ecuación dada.

Por supuesto que la anterior definición es puramente formal, en el sentido de que sólo tendrá sentido $k_{n+1}(s,t)$ si la función $u \mapsto k_1(s,u)k_n(u,t)$ es integrable en $[a,b]$. Vamos a ver que, en realidad, la sucesión de funciones (k_n) está formada por funciones continuas en $[a,b] \times [a,b]$.

Proposición 11 *La sucesión de núcleos iterados de la ecuación integral anterior está formada por funciones continuas en $[a,b]\times[a,b]$, si k lo es.*

Prueba : Para $k=1$ es evidente. Asumamos que k_n es continua en $[a,b]\times[a,b]$.

Fijados $(s,t)\in[a,b]\times[a,b]$, la función

$$u\mapsto k_1(s,u)k_n(u,t)$$

será también continua en $[a,b]$, luego tiene sentido

$$k_{n+1}(s,t):=\int_a^b k_1(s,u)k_n(u,t)du.$$

Vamos a ver ahora que

$$(s,t)\mapsto\int_a^b k_1(s,u)k_n(u,t)du$$

define una función continua en $[a,b]\times[a,b]$.

La función

$$(s,u,t)\mapsto k_1(s,u)k_n(u,t)$$

es uniformemente continua en el compacto $Q:=[a,b]\times[a,b]\times[a,b]$. Dado pues $\varepsilon>0$ existe $\delta>0$ de manera que si $(s_1,u_1,t_1),(s_2,u_2,t_2)\in Q$ con $\|(s_1,u_1,t_1)-(s_2,u_2,t_2)\|<\delta$ entonces

$$|k_1(s_1,u_1)k_n(u_1,t_1)-k_1(s_2,u_2)k_n(u_2,t_2)|<\frac{\varepsilon}{b-a}.$$

Tomamos ahora $(s_1,t_1),(s_2,t_2)\in[a,b]\times[a,b]$ con

$$\|(s_1,t_1)-(s_2,t_2)\|<\delta.$$

Entonces para todo $u\in[a,b]$, $(s_1,u,t_1),(s_2,u,t_2)\in Q$, con

$$\|(s_1,u,t_1)-(s_2,u,t_2)\|<\delta$$

luego para todo $u\in[a,b]$

$$|k_1(s_1,u)k_n(u,t_1)-k_1(s_2,u)k_n(u,t_2)|<\frac{\varepsilon}{b-a}.$$

por lo que

$$\begin{aligned}|k_{n+1}(s_1,t_1)-k_{n+1}(s_2,t_2)| &= |\textstyle\int_a^b(k_1(s_1,u)k_n(u,t_1)-k_1(s_2,u)k_n(u,t_2))du|\\ &\le \textstyle\int_a^b|k_1(s_1,u)k_n(u,t_1)-k_1(s_2,u)k_n(u,t_2)|du\\ &\le \tfrac{\varepsilon}{b-a}(b-a)=\varepsilon\end{aligned}$$

luego k_{n+1} es (uniformemente) continua en $[a,b]\times[a,b]$.

Por inducción la proposición está probada.

$\square$

La sucesión de núcleos iterados nos permitirá calcular más fácilmente las potencias del operador T.

Proposición 12 *Si k es continuo en $[a,b]\times[a,b]$, y (k_n) es la sucesión de núcleos iterados que genera, y si T es el operador de Fredholm asociado $x \mapsto Tx$, con $Tx(s) := \int_a^b k(s,t)x(t)dt$, entonces se cumple que, para toda función $y \in \mathcal{C}([a,b])$,*

$$T^n y(s) = \int_a^b k_n(s,t)y(t)dt.$$

para cada $s \in [a,b]$.

Prueba : Se prueba también por inducción. Para $n=1$ es evidente por definición de T.

Supongamos que $T^n y(s) = \int_a^b k_n(s,t)y(t)dt$. Entonces una aplicación sencilla del teorema de Fubini nos da

$$\begin{aligned} T_{n+1}y(s) &= T(T^n y)(s) = \textstyle\int_a^b k(s,t)(T^n y)(t)dt \\ &= \textstyle\int_a^b k(s,t)(\int_a^b k_n(t,u)y(u)du)dt \\ &= \textstyle\int_a^b(\int_a^b k(s,t)k_n(t,u)y(u)du)dt \\ &= \textstyle\int_a^b(\int_a^b k(s,t)k_n(t,u)y(u)dt)du \\ &= \textstyle\int_a^b y(u)(\int_a^b k(s,t)k_n(t,u)dt)du \\ &= \textstyle\int_a^b y(u)k_{n+1}(s,u)du \\ &= \textstyle\int_a^b k_{n+1}(s,t)y(t)dt, \end{aligned}$$

lo que completa la prueba.

□

Corolario 11 *La ecuación integral de Fredhölm $x-Tx=y$ con $Tx(s) = \int_a^b k(s,t)x(t)dt$, si el núcleo k es continuo en $[a,b]\times[a,b]$ y si $\|T\| < 1$ tiene por solución la función $x:[a,b]\to\mathbb{R}$ dada por*

$$x(s) = T^0 y(s) + \sum_{n=1}^{\infty}\int_a^b k_n(s,t)y(t)dt = y(s) + \sum_{n=1}^{\infty}\int_a^b k_n(s,t)y(t)dt.$$

Aún veremos una expresión mejor para la solución de la ecuación integral que estamos estudiando.

Proposición 13 *Si $k:[a,b]\times[a,b]\to\mathbb{R}$ es una función continua verificando que*

$$M := \max\{|k(s,t)| : (s,t)\in[a,b]\times[a,b]\} < \frac{1}{b-a}$$

entonces la sucesión de núcleos iterados (k_n) verifica que la serie de Neumann $\sum_{n=1}^{\infty} k_n(s,t)$ *es convergente y la función $(s,t)\mapsto \sum_{n=1}^{\infty} k_n(s,t)$ es continua en $[a,b]\times[a,b]$.*

Prueba : Afirmamos en primer lugar que para todo entero positivo n,

$$|k_n(s,t)| \leq M^n(b-a)^{n-1}.$$

Efectivamente, para $n = 1$ la afirmación no es más que la definición de M.

Asumamos que la afirmación se cumple para $n = 1, 2, \ldots, m$. Entonces, para cualesquiera $s, t \in [a,b] \times [a,b]$ se tendrá

$$\begin{aligned} |k_{n+1}(s,t)| &= |\int_a^b k_1(s,u)k_n(u,t)du| \\ &\leq \int_a^b |k_1(s,u)||k_n(u,t)|du \\ &\leq \int_a^b M|k_n(u,t)|du \\ &\leq \int_a^b MM^n(b-a)^{n-1}du \\ &= M^{n+1}(b-a)^n. \end{aligned}$$

Luego hemos probado la afirmación por inducción sobre n.

Entonces la serie $\sum_{n=1}^{\infty} k_n(s,t)$ es absolutamente convergente pues

$$\sum_{n=1}^{\infty} |k_n(s,t)| \leq \sum_{n=1}^{\infty} M^n(b-a)^{n-1} = M\sum_{n=1}^{\infty}(M(b-a))^{n-1} = M\frac{1}{1-M(b-a)}.$$

Como toda serie de números reales o complejos absolutamente convergente es convergente, está bien definido para cualesquiera $(s,t) \in [a,b] \times [a,b]$,

$$h(s,t) := \sum_{n=1}^{\infty} k_n(s,t).$$

La sucesión de sumas parciales de esta serie es decir

$$\left(\sum_{n=1}^{m} k_n(s,t)\right)_{m\geq 1}$$

está formada por funciones continuas.

Además esta sucesión de sumas parciales converge a la función h uniformemente en $[a,b] \times [a,b]$ porque

$$|h(s,t) - \sum_{n=1}^{m} k_n(s,t)| = |\sum_{n=m+1}^{m} k_n(s,t)| \leq M \sum_{n=m+1}^{m} (M(b-a))^{n-1} = M\frac{(M(b-a))^m}{1-M(b-a)}$$

tiende a 0 cuando m tiende a ∞ independientemente de (s,t). Entonces h es continua por ser límite uniforme (en $[a,b] \times [a,b]$) de una sucesión de funciones continuas, lo que completa la prueba.

$\square$

Definición 16 *La función*

$$h(s,t) := \sum_{n=1}^{\infty} k_n(s,t).$$

definida bajo las condiciones de la proposición anterior se llama núcleo resolvente *de la ecuación integral de Fredholm dada.*

Teorema 23 *Si $k : [a,b] \times [a,b] \to \mathbb{R}$ es una función continua verificando que*

$$M := \text{máx}\{|k(s,t)| : (s,t) \in [a,b] \times [a,b]\} < \frac{1}{b-a}$$

entonces la ecuación integral de Fredhölm de segunda especie

$$x(s) - \int_a^b k(s,t)x(t)dt = y(s)$$

tiene una solución $x \in \mathcal{C}([a,b])$, dada por

$$x(s) = y(s) + \int_a^b h(s,t)y(t)dt.$$

donde h es el núcleo resolvente de la sucesión de núcleos iterados (k_n) generada por el núcleo k.

Prueba : Sabemos por el teorema anterior que bajo las condiciones del enunciado, la ecuación dada tiene una solución $x \in \mathcal{C}([a,b]$, dada por

$$x = [\sum_{n=0}^{\infty} T^n](y).$$

es decir, para cada $s \in [a,b]$,

$$x(s) = y(s) + \sum_{n=1}^{\infty} T^n y(s).$$

También hemos visto que $T^n y(s) = \int_a^b k_n(s,t)y(t)dt$ para $n \geq 1$, luego

$$\begin{aligned} x(s) &= y(s) + \textstyle\sum_{n=1}^{\infty} \int_a^b k_n(s,t)y(t)dt \\ &= y(s) + \text{lím}_m[\textstyle\sum_{n=1}^{m} \int_a^b k_n(s,t)y(t)dt] \\ &= y(s) + \text{lím}_m[\textstyle\int_a^b \sum_{n=1}^{m}(k_n(s,t)y(t))dt] \end{aligned}$$

será otra expresión de esta solución.

Lo que vamos a ver (coloquialmente dicho) es que que el límite anterior conmuta con la integral.

Fijado $s \in [a,b]$, consideremos las sucesión de funciones

$$t \mapsto s_m(s,t) := \sum_{n=1}^{m} (k_n(s,t)y(t))$$

Está claro que según hemos visto antes

$$|s_m(s,t)| \leq \sum_{n=1}^{m} |k_n(s,t)||y(t)| \leq \|y\|_\infty \sum_{n=1}^{m} |k_n(s,t)| \leq \|y\|_\infty \sum_{n=1}^{\infty} |k_n(s,t)| \leq \|y\|_\infty \frac{M}{1-M(b-a)}.$$

Es decir, todas las funciones $s_m(s,.)$ están dominadas en $[a,b]$ por una función constante, que desde luego es integrable en $[a,b]$. Podemos aplicar el teorema de la convergencia dominada, y entonces

$$\int_a^b \lim_m s_m(s,t)dt = \lim_m \int_a^b s_m(s,t)dt$$

o sea,

$$\int_a^b h(s,t)y(t)dt = \lim_m \int_a^b s_m(s,t)dt.$$

Sustituyendo, queda que

$$\begin{aligned} x(s) &= y(s) + \lim_m \left[\int_a^b \left(\sum_{n=1}^m k_n(s,t)y(t)\right) dt\right] \\ &= y(s) + \lim_m \left[\int_a^b s_m(s,t)dt\right] \\ &= y(s) + \int_a^b h(s,t)y(t)dt. \end{aligned}$$

lo que completa la prueba.

$\square$

3.2.1. Caso de los núcleos degenerados

Cuando el núcleo de una ecuación integral de Fredholm tiene la forma

$$k(s,t) := \sum_{i=1}^{m} a(s)b(t)$$

donde a y b son funciones continuas en el intervalo cerrado $[a,b]$ se dice que es *degenerado.*

En este caso, la sucesión de núcleos iterados obtenida a partir de k suele ser de cálculo más sencillo.

En particular, si $k(s,t) = a(s)b(t)$ tendremos

$$k_2(s,t) := \int_a^b k_1(s,u)k_1(u,t)dt = \int_a^b a(s)b(u)a(u)b(t)du = k_1(s,t)\int_a^b k(u,u)du.$$

Vemos pues que k_2 es igual a k_1 multiplicado por una constante, digamos c.

$$k_3(s,t) := \int_a^b k_1(s,u)k_2(u,t)dt = \int_a^b a(s)b(u)ca(u)b(t)du = ck_1(s,t)\int_a^b k(u,u)du = k_1(s,t)c^2.$$

Si $k_n(s,t) = k_1(s,t)c^{n-1}$ entonces

$$k_{n+1}(s,t) = \int_a^b k_1(s,u)k_n(u,t)dt = \int_a^b a(s)b(u)c^{n-1}a(u)b(t)du$$

$$= c^{n-1}k_1(s,t)\int_a^b k(u,u)du = k_1(s,t)c^n.$$

De aquí resulta inmediatamente que la expresión del núcleo resolvente será

$$h(s,t) = a(s)b(t)\sum_{n=1}^{\infty} c^{n-1}.$$

Pero aún hay un método más fácil, de carácter algebraico para resolver este tipo de ecuaciones.

Observar que la ecuación dada es

$$x(s) - \int_a^b a(s)b(t)x(t)dt = y(s).$$

Si ponemos $A := \int_0^1 b(t)x(t)dt$, admite la forma más sencilla

$$x(s) - a(s)A = y(s),$$

y tendremos una solución x cuando conozcamos la constante A. Sustituyendo en la propia definición de A nos queda

$$A = \int_a^b b(t)x(t)dt = \int_a^b b(t)[y(t) + a(t)A]dt,$$

lo que nos permite (si $1 - \int_a^b a(t)b(t)dt \neq 1$) despejar A:

$$(1 - \int_a^b a(t)b(t)dt)A = \int_a^b b(t)y(t)dt \Rightarrow A = \frac{\int_a^b b(t)y(t)dt}{1 - \int_a^b a(t)b(t)dt}.$$

Cualquier ecuación de este tipo con núcleo degenerado admite un tratamiento similar.

Ejemplo 20 *Resolver*

$$x(s) - \int_0^1 \lambda e^{s-t}x(t)dt = f(s).$$

Como $k(s,t) = \lambda e^{s-t}$ y para $s,t \in [0,1]$

$$|k(s,t)| = |\lambda e^{s-t}| \leq |\lambda|e$$

para λ suficientemente pequeño se tendrá que

$$M := \text{máx}\{|k(s,t) : (s,t) \in [0,1] \times [0,1]\} < 1.$$

Formamos la sucesión de núcleos iterados a partir de k.

$$k_2(s,t) := \int_0^1 k_1(s,u)k_1(u,t)du = \int_0^1 \lambda e^{s-u}\lambda e^{u-t}dt =$$

$$= \lambda^2 e^{s-t}.$$

Si suponemos $k_n(s,t) = \lambda^n e^{s-t}$ entonces

$$k_{n+1}(s,t) := \int_0^1 k_1(s,u)k_n(u,t)du = \int_0^1 \lambda e^{s-u}\lambda^n e^{u-t}dt =$$

$$= \lambda^{n+1}e^{s-t}.$$

Por lo tanto,

$$h(s,t) := \sum_{n=1}^{\infty} k_n(s,t) = \sum_{n=1}^{\infty} \lambda^n e^{s-t}$$

tiene sentido para $|\lambda| < 1$, con

$$h(s,t) = e^{s-t}\frac{\lambda}{1-\lambda}.$$

Una solución de la ecuación dada será pues

$$x(s) = f(s) + \int_0^1 \frac{\lambda}{1-\lambda}e^{s-t}f(t)dt,$$

es decir,

$$x(s) = f(s) + \frac{\lambda e^s}{1-\lambda}\int_0^1 e^{-t}f(t)dt.$$

Pero podemos llegar a la misma solución de un modo mucho más simple:

La ecuación dada es

$$x(s) - \lambda e^s \int_0^1 e^{-t}x(t)dt = f(s).$$

Si llamamos $A := \int_0^1 e^{-t}x(t)dt$, se reduce a

$$x(s) - \lambda A e^s = f(s).$$

De la misma definición de la constante A tenemos:

$$A := \int_0^1 e^{-t}x(t)dt = \int_0^1 e^{-t}[f(t) + \lambda A e^t]dt$$

o sea,

$$A = \int_0^1 e^{-t}[f(t) + \lambda A e^t]dt = \int_0^1 e^{-t}f(t)dt + \int_0^1 \lambda A dt$$

que se reduce a

$$(1-\lambda)A = \int_0^1 e^{-t}f(t)dt \Rightarrow A = \frac{1}{1-\lambda}\int_0^1 e^{-t}f(t)dt.$$

Sustituyendo, encontramos

$$x(s) = f(s) + \lambda e^s \frac{1}{1-\lambda}\int_0^1 e^{-t} f(t)dt.$$

Ejemplo 21 Resolver la ecuación integral

$$x(s) = \frac{s}{2} + \frac{1}{2}\sin(s) + \int_0^1 \left(1-\cos(st^2)\right) sx(t)dt.$$

El núcleo será

$$k(s,t) = \left(1-\cos(st^2)\right) s$$

Calculamos

$$\|k\|_\infty = \sup\{|\left(1-\cos(st^2)\right) s| : s,t \in [0,1]\}$$

Al ser k continua en el compacto $[0,1]\times[0,1]$ se tiene que existe $(s_0,t_0) \in [0,1]\times[0,1]$ tal que

$$\|k\|_\infty = k(s_0,t_0) = |\left(1-\cos(s_0t_0^2)\right) s_0|$$

Notemos que las funciónes reales $t \mapsto (1-\cos(st^2))s$ son crecientes en $[0,1]$ para cualquier $s \in [0,1]$.

$$\|k\|_\infty = |\left(1-\cos(s_0t_0^2)\right) s_0| \leq$$
$$\leq (1-\cos(s_0))s_0 \leq 1 - cos(1) = 0,45967$$

Como $\|k\|_\infty < 1 = 1-0$, y en principio podemos aplicar el teorema de los núcleos iterados.

Tomamos pues $k_1(s,t) = k(s,t) = (1-\cos(st^2))\, s$.

$$\begin{aligned} k_2(s,t) = & \int_0^1 k(s,u)k_1(u,t)du \\ = & \int_0^1 (1-\cos(su^2))\, s\, (1-\cos(ut^2))\, udu \end{aligned}$$

Vemos pues que el cálculo efectivo de $k_2(s,t)$ parece muy laborioso, supuesto que fuera posible.

Recurrimos pues a aproximar el n£cleo $k(s,t)$ por un n£cleo degenerado. Usamos la aproximación de Mc Laurin

$$\cos(\alpha) \sim 1 - \frac{\alpha^2}{2}$$

que nos da

$$s(1-\cos(st^2)) \sim s\frac{s^2t^4}{2} =: k_a(s,t).$$

La ecuación integral aproximante será pues

$$x_a(s) = \frac{s}{2} + \frac{1}{2}\sin(s) + \int_0^1 \frac{s^3t^4}{2} x_a(t)dt.$$

Si escribimos

$$A := \int_0^1 \frac{t^4}{2} x_a(t)dt$$

será

$$x_a(s) = \frac{s}{2} + \frac{1}{2}\sin(s) + As^3,$$

o sea,

$$A = \int_0^1 \frac{t^4}{2}\left[\frac{t}{2} + \frac{1}{2}\sin(t) + At^3\right]dt = \int_0^1 \left[\frac{t^5}{4} + \frac{t^4}{4}\sin(t) + A\frac{t^7}{2}\right]dt$$

Operamos en esta igualdad y tenemos:

$$A = \frac{1}{24} + \frac{1}{4}\int_0^1 t^4\sin(t)dt + A\frac{1}{16}$$

Al ser $\int_0^1 t^4\sin(t)dt = 0{,}146647...$, obtenemos

$$\frac{15}{16}A = \frac{1}{24} + \frac{1}{4}0{,}146647,$$

es decir $A = 0{,}0835502$.

Luego la solución aproximada a nuestra ecuación es

$$x_a(s) = \frac{s}{2} + \frac{1}{2}\sin(s) + (0{,}0835502)s^3.$$

Tratamos ahora de estimar el error cometido. En otras palabras, vamos a tratar de conocer la "distancia.$^{\text{en}}$tre la solución $x(s)$ de la ecuación dada y la solución aproximada $x_a(s)$. Para ello calcularemos $\|x - x_a\|_\infty$:

$$|x(s) - x_a(s)| = \left|\frac{s+\sin(s)}{2} + \int_0^1 \left(1-\cos(st^2)\right)sx(t)dt - \frac{s+\sin(s)}{2} + \int_0^1 \frac{s^3t^4}{2}x_a(t)dt\right|$$

$$\leq \int_0^1 \left|s\left(1-\cos(st^2)\right)x(t) - \frac{s^3t^4}{2}x_a(t)\right|dt.$$

$$\leq \int_0^1 \left[|k(s,t)x(t) - k(s,t)x_a(t)| + |k(s,t)x_a(t) - k_a(s,t)x_a(t)|\right]dt$$

$$\leq \int_0^1 \left[|k(s,t)||x(t) - x_a(t)| + |k(s,t) - k_a(s,t)||x_a(t)|\right]dt$$

$$\leq \int_0^1 \left[\|k\|_\infty\|x - x_a\|_\infty + |k(s,t) - k_a(s,t)||x_a(t)|\right]dt$$

$$= \|k\|_\infty\|x - x_a\|_\infty + \int_0^1 |k(s,t) - k_a(s,t)||x_a(t)|dt.$$

De aqu¡ resulta:

$$\|x - x_a\|_\infty = \sup\{\|x(s) - x_a(s)\| : s \in [0,1]\}$$

$$\leq \|k\|_\infty\|x - x_a\|_\infty + \int_0^1 |k(s,t) - k_a(s,t)||x_a(t)|dt.$$

Por tanto,

$$\|x - x_a\| \leq \frac{1}{1 - \|k\|_\infty} \int_0^1 |k(s,t) - k_a(s,t)||x_a(t)|dt.$$

Para acotar el término

$$|k(s,t) - k_a(s,t)| = \left| \left(1 - \cos(st^2)\right) - \frac{s^2t^4}{2} \right| s$$

recordemos que el núcleo degenerado aproximante $k_a(s,t) = s\frac{s^2t^4}{2}$ se obtuvo a partir del desarrollo finito de Mc Laurin:

$$\cos(st^2) = 1 - \frac{1}{2!}(st^2)^2 + \frac{1}{4!}(st^2)^4 - ...$$

Por lo tanto sabemos que

$$|\left(1 - \cos(st^2)\right) - \frac{1}{2!}(st^2)^2| \leq |\frac{1}{4!}(st^2)^4|$$

y como $s \in [0,1]$,

$$s \left| \left(1 - \cos(st^2)\right) - \frac{1}{2!}(st^2)^2 \right| \leq s \left| \frac{1}{4!}(st^2)^4 \right| \leq \frac{1}{4!}t^8$$

Esto nos permite escribir

$$\|x - x_a\| \leq \frac{1}{1 - \|k\|_\infty} \int_0^1 |k(s,t) - k_a(s,t)||x_a(t)|dt$$

$$\leq \frac{1}{1 - \|k\|_\infty} \int_0^1 \frac{1}{4!}t^8|x_a(t)|dt = \frac{1}{1 - \|k\|_\infty} 0{,}00417956$$

$$\frac{1}{1 - 0{,}45967} 0{,}00417956 = 0{,}0077358.$$

También podemos dar otra estimación por un método aparentemente menos directo:

La ecuación dada

$$x(s) = \frac{s}{2} + \frac{1}{2}\sin(s) + \int_0^1 \left(1 - \cos(st^2)\right) sx(t)dt.$$

la escribimos como

$$x(s) - \int_0^1 \left(1 - \cos(st^2)\right) sx(t)dt = \frac{s}{2} + \frac{1}{2}\sin(s)$$

o bien abreviadamente como

$$x(s) - Tx(s) = y(s) \Leftrightarrow (I - T)(x)(s) = y(s)$$

donde

$$Tx(s) := \int_0^1 \left(1 - \cos(st^2)\right) sx(t)dt = \int_0^1 k(s,t)x(t)dt$$

$$y(s) := \frac{s}{2} + \frac{1}{2}\sin(s)$$

Análogamente la ecuación integral aproximante la escribimos como

$$x_a(s) - T_a x_a(s) = y(s) \Leftrightarrow (I - T_a)(x_a)(s) = y(s)$$

donde

$$T_a x_a(s) := \int_0^1 k_a(s,t)x_a(t)dt.$$

Recordemos que si los operadores T y T_a son invertibles entonces $I - T$ $I - T_a$ son invertibles y además

$$\|(I-T)^{-1}\| < \frac{1}{1-\|T\|},\ \|(I-T_a)^{-1}\| < \frac{1}{1-\|T_a\|}$$

Entonces, como

$$\|x_a - x\|_\infty = \|(I-T)^{-1}(I-T)x_a - (I-T)^{-1}(y)\|_\infty$$

$$= \|(I-T)^{-1}[(I-T)x_a - y]\|_\infty \leq \|(I-T)^{-1}\|\|(I-T)x_a - y\|_\infty$$

$$= \|(I-T)^{-1}\|\|(I-T)x_a - (I-T_a)(x_a)\|_\infty$$

$$\leq \|(I-T)^{-1}\|\|T_a - T\|\|x_a\|_\infty < \frac{1}{1-\|T\|}\|T_a - T\|\|x_a\|_\infty,$$

para acotar superiormente $\|x_a - x\|$ basta con conocer cotas superiores de $\frac{1}{1-\|T\|}$, $\|T_a - T\|$, $\|x_a\|$.

Por ejemplo, $T : \mathcal{C}([0,1]) \longrightarrow \mathcal{C}([0,1])$ y viene dado por

$$Tx(s) = \int_0^1 k(s,t)x(t)dt$$

Luego, fijada $x \in \mathcal{C}([0,1])$

$$|Tx(s)| \leq \int_0^1 |k(s,t)||x(t)|dt \leq \|k\|_\infty\|x\|_\infty$$

lo que nos da $\|Tx\|_\infty \leq \|k\|_\infty\|x\|_\infty$ y de aqu¡ resulta

$$\|T\| \leq \|k\|_\infty \Rightarrow 1 - \|T\| \geq 1 - \|k\|_\infty \Rightarrow \frac{1}{1-\|T\|} \leq \frac{1}{1-\|k\|_\infty} \leq \frac{1}{1-0{,}45969}$$

Del mismo modo acotamos $\|T - T_a\|$: Fijada $x \in \mathcal{C}([0,1])$

$$|(T-T_a)x(s)| \leq \int_0^1 |k(s,t) - k_a(s,t)||x(t)|dt \leq \|x\|_\infty \int_0^1 |k(s,t) - k_a(s,t)|dt$$

$$\leq \|x\|_\infty \int_0^1 \frac{1}{4!}t^8 dt = \|x\|_\infty \frac{1}{4!9}$$

lo que nos da $\|Tx\|_\infty \leq \frac{1}{4!9}\|x\|_\infty$ y de aqu¡ resulta

$$\|T - T_a\| \leq \frac{1}{4!9}$$

Por último calculamos $\|x_a\|_\infty$.

$$\|x_a\|_\infty := \text{máx}\{\frac{s}{2} + \frac{1}{2}\sin(s) + (0{,}0835502)s^3 : s \in [0,1]\}$$

$$= \frac{1}{2} + \frac{1}{2}\sin(1) + (0{,}0835502) = 1,00428.$$

Obtenemos finalmente

$$\|x_a - x\|_\infty < \frac{1}{1 - \|T\|}\|T_a - T\|\|x_a\|_\infty$$

$$\leq \frac{1}{1 - 0{,}459697}\frac{1}{4!9}1,00428 = 0{,}00860955.$$

3.3. Aplicación: Ecuaciones integrales de Volterra

Si $[a,b]$ es un intervalo compacto de la recta real, sea

$$\Delta := \{(s,t) \in \mathbb{R}^2 : a \leq t \leq s \leq b\}.$$

En otras palabras, Δ es la envoltura convexa del conjunto $\{(a,a),(b,b),(b,a)\}$, es decir, menor conjunto convexo del plano que contiene a los puntos $(a,a),(b,b),(b,a)$. Sea $k : \Delta \to \mathbb{R}$ una función continua.

Al problema de encontrar una función incógnita $x : [a,b] \to \mathbb{R}$, verificando para cada $s \in [a,b]$

$$\boxed{x(s) - \int_a^s k(s,t)x(t)dt \;\; = \;\; y(s)}$$

donde $y \in \mathcal{C}([a,b])$ es un dato, se le suele llamar *ecuación integral de Volterra de segunda especie*. Puede escribirse con notación funcional más condensada como

$$(Id - K)x = y,$$

donde K es el operador que transforma $x \in \mathcal{C}([a,b])$ en la función Kx dada por

$$Kx(s) := \int_a^s k(s,t)x(t)dt.$$

Desde luego K es lineal. También $k(x,y)$ suele llamarse *nucleo* del operador K o de la ecuación dada.

Vamos a ver que $K : \mathcal{C}([a,b]) \to \mathcal{C}([a,b])$. Es decir, vamos a ver que para cada $x \in \mathcal{C}([a,b])$, Kx es una función contínua en $[a,b]$. Si $x = 0_{\mathcal{C}([a,b])}$ entonces $Kx = 0_{\mathcal{C}([a,b])}$ que es una función continua. Supongamos x distinta de $0_{\mathcal{C}([a,b])}$.

Como k es continua en el compacto Δ, dado $\varepsilon > 0$ existirá $\delta_1 > 0$ tal que si $(s_1,t_1),(s_2,t_2) \in \Delta$ con $\|(s_1,t_1) - (s_2,t_2)\| < \delta_1$ entonces

$$|k(s_1,t_1) - k(s_2,t_2)| < \frac{\varepsilon}{2(b-a)\|x\|_\infty}.$$

Sea $M := \max\{|k(s,t)| : (s,t) \in \Delta\}$. Si $M = 0$ el operador K sería idénticamente nulo, y no habría nada que probar. Sea $0 < \delta_2 := \min\{\frac{\varepsilon}{2M\|x\|_\infty}, \delta_1\}$. Desde luego, $\delta_2 > 0$. Si $s_1, s_2 \in [a,b]$ con $a \leq s_1 < s_2 \leq b$ y con $s_2 - s_1 < \delta_2$, tendremos que los puntos (s_1,t)

(s_2, t) pertenecen al triángulo Δ y $\|(s_1,t)-(s_2,t)\| = s_2 - s_1 < \delta_2 < \delta_1$. Entonces

$$\begin{aligned}
|Kx(s_2) - Kx(s_1)| &= \left|\int_a^{s_2} k(s_2,t)x(t)dt - \int_a^{s_1} k(s_1,t)x(t)dt\right| \\
&= | \int_a^{s_2} k(s_2,t)x(t)dt - \int_a^{s_1} k(s_2,t)x(t)dt \\
&\qquad + \int_a^{s_1} k(s_2,t)x(t)dt - \int_a^{s_1} k(s_1,t)x(t)dt| \\
&= \left|\int_{s_1}^{s_2} k(s_2,t)x(t)dt + \int_a^{s_1}(k(s_2,t)-k(s_1,t))x(t)dt\right| \\
&\leq \int_{s_1}^{s_2} |k(s_2,t)||x(t)|dt + \int_a^{s_1} |k(s_2,t)-k(s_1,t)||x(t)|dt \\
&\leq \int_{s_1}^{s_2} M\|x\|_\infty dt + \int_a^{s_1} \tfrac{\varepsilon}{2(b-a)\|x\|_\infty}\|x\|_\infty dt \\
&\leq M\|x\|_\infty (s_2 - s_1) + \tfrac{\varepsilon}{2(b-a)\|x\|_\infty}\|x\|_\infty (s_1 - a) \\
&< M\|x\|_\infty \tfrac{\varepsilon}{2M\|x\|_\infty} + \tfrac{\varepsilon}{2(b-a)\|x\|_\infty}\|x\|_\infty (b-a) \\
&= \varepsilon.
\end{aligned}$$

Esto demuestra que Kx es una función (uniformemente) continua en $[a,b]$, luego $K : \mathcal{C}([a,b]) \to \mathcal{C}([a,b])$.

Además, para cada $s \in [a,b]$,

$$|Kx(s)| := \left|\int_a^s k(s,t)x(t)\right| \leq M\|x\|_\infty (s-a) \leq M\|x\|_\infty (b-a).$$

Por tanto,

$$\|Kx\|_\infty \leq M(b-a)\|x\|_\infty$$

lo que demuestra que K es continuo con $\|K\| \leq M(b-a)$.

3.3.1. Método de los núcleos iterados

Definición 17 *Si*

$$x(s) - \int_a^s k(s,t)x(t)dt = y(s)$$

es una ecuación de Volterra de segunda clase, la sucesión de funciones $k_n : \Delta \to \mathbb{R}$ *dada recurrentemente por*

$$k_1(s,t) := k(s,t),\ k_{n+1}(s,t) := \int_t^s k_1(s,u)k_n(u,t)du$$

se llama sucesión de núcleos iterados de la ecuación dada.

Si $t \leq u \leq s$ entonces $(s,u),(u,t) \in \Delta$. Por tanto la expresión que figura bajo el signo de integral en la definición anterior es una función (de u) continua en el intervalo $[t,s]$, siempre que k_1 y k_n lo sean en Δ.

La función $k_1 = k$ es continua en Δ por hipótesis. Asumamos que k_n es continua en Δ. Entonces tiene sentido la expresión que define a k_{n+1}. Por hipótesis de inducción existe $A < \infty$, una cota superior en Δ de las funciones continuas $|k_1|, |k_n|, |k_1 k_n|$. Además, como

k_1 y k_n son (como suponemos) uniformemente continuas en Δ, dado $\varepsilon > 0$ existe $\delta > 0$ (que siempre puede tomarse menor que $\frac{\varepsilon}{2A}$) tal que si $(s_1, t_1), (s_2, t_2)$ son puntos de Δ con $\|(s_1, t_1) - (s_2, t_2)\| < \delta$ entonces $|k_i(s_1, t_1) - k_i(s_2, t_2)| < \frac{\varepsilon}{2A(b-a)}$ $(i = 1, n)$.

A partir de aquí notemos que, si $a \le t_1 < t_2 \le s \le b$ con $t_2 - t_1 < \delta$,

$$\begin{aligned}
|k_{n+1}(s,t_1) - k_{n+1}(s,t_2)| &= |\textstyle\int_{t_1}^{s} k_1(s,u)k_n(u,t_1)du - \int_{t_2}^{s} k_1(s,u)k_n(u,t_2)du \\
&= |\textstyle\int_{t_1}^{t_2} k_1(s,u)k_n(u,t_1)du + \int_{t_2}^{s} k_1(s,u)(k_n(u,t_1) - k_n(u,t_2))du| \\
&\le \textstyle\int_{t_1}^{t_2} Adu + \int_{t_2}^{s} A|k_n(u,t_1) - k_n(u,t_2)|du \\
&\le A(t_2 - t_1) + (s - t_2)A\tfrac{\varepsilon}{2A(b-a)} \\
&\le A(t_2 - t_1) + \tfrac{\varepsilon}{2}.
\end{aligned}$$

De igual modo, si $a \le t \le s_1 < s_2 \le b$,

$$\begin{aligned}
|k_{n+1}(s_1,t) - k_{n+1}(s_2,t)| &= |\textstyle\int_{t}^{s_1} k_1(s_1,u)k_n(u,t)du - \int_{t}^{s_2} k_1(s_2,u)k_n(u,t)du \\
&= |\textstyle\int_{t}^{s_1} (k_1(s_1,u) - k_1(s_2,u))k_n(u,t)du + \int_{s_1}^{s_2} k_1(s_2,u)k_n(u,t)du| \\
&\le \textstyle\int_{t}^{s_1} A|k_1(s_1,u) - k_1(s_2,u)|du + \int_{s_1}^{s_2} Adu \\
&\le A(s_1 - t)\tfrac{\varepsilon}{2A(b-a)} + A(s_2 - s_1) \\
&\le \tfrac{\varepsilon}{2} + A(s_2 - s_1).
\end{aligned}$$

Entonces, si $(s_1, t_1), (s_2, t_2)$ son puntos de Δ con $\|(s_1, t_1) - (s_2, t_2)\| < \delta$, tendremos, aplicando estas desigualdades:

$$\begin{aligned}
|k_{n+1}(s_1,t_1) - k_{n+1}(s_2,t_2)| &\le |k_{n+1}(s_1,t_1) - k_{n+1}(s_1,t_2)| + |k_{n+1}(s_1,t_2) - k_{n+1}(s_2,t_2)| \\
&\le \tfrac{\varepsilon}{2} + A|t_2 - t_1| + \tfrac{\varepsilon}{2} + A|s_1 - s_2| \\
&\le \varepsilon + 2A\delta < 2\varepsilon\,,
\end{aligned}$$

luego k_{n+1} es (uniformemente) continua en Δ.

Por inducción hemos probado pues la siguiente

PROPOSICIÓN 14 *La sucesión de núcleos iterados (k_n) de la ecuación integral de Volterra anterior, está formada por funciones continuas en Δ, si k lo es.*

LEMA 3 *Si $k : \Delta \to \mathbb{R}$ es una función continua con*

$$M := \text{máx}\{|k(s,t)| : (s,t) \in \Delta\},$$

entonces la sucesión de núcleos iterados (k_n) verifica para cualquier $(s,t) \in \Delta$ que

$$|k_n(s,t)| \le \frac{M^n(s-t)^{n-1}}{(n-1)!}.$$

Prueba : Efectivamente, para $n = 1$ la afirmación no es más que la definición de M. (Asumiendo el convenio $0! = 1$).

Si que la acotación se cumple para $k_1, k_2, \ldots, k_n$, para cualquier $(s,t) \in \Delta$ se tendrá

$$\begin{aligned}
|k_{n+1}(s,t)| &= |\int_t^s k_1(s,u)k_n(u,t)du| \\
&\leq \int_t^s |k_1(s,u)||k_n(u,t)|du \\
&\leq \int_t^s M|k_n(u,t)|du \\
&\leq \int_a^b M\frac{M^n(u-t)^{n-1}}{(n-1)!}du \\
&= \frac{M^{n+1}}{(n-1)!}\left(\frac{(u-t)^n}{n}\right)_{u=t}^{u=s} \\
&= \frac{M^{n+1}}{(n-1)!}\frac{(s-t)^n}{n} \\
&= \frac{M^{n+1}(s-t)^n}{n!}
\end{aligned}$$

Luego hemos probado el lema por inducción sobre n.

□

Proposición 15 *Bajo las condiciones del lema anterior, para cada $(s,t) \in \Delta$, la* serie de Neumann $\sum_{n=1}^{\infty} k_n(s,t)$ *es convergente y la función $(s,t) \mapsto \sum_{n=1}^{\infty} k_n(s,t)$ es continua en Δ.*

Prueba : La serie $\sum_{n=1}^{\infty} k_n(s,t)$ es absolutamente convergente pues

$$\sum_{n=1}^{\infty} |k_n(s,t)| \leq \sum_{n=1}^{\infty} \frac{M^n(b-a)^{n-1}}{(n-1)!} = M\sum_{n=1}^{\infty} \frac{(M(b-a))^{n-1}}{(n-1)!} = Me^{M(b-a)}.$$

Como toda serie de números reales o complejos absolutamente convergente es convergente, está bien definido para cualesquiera $(s,t) \in \Delta$,

$$h(s,t) := \sum_{n=1}^{\infty} k_n(s,t).$$

La sucesión de sumas parciales de esta serie es decir

$$\left(\sum_{n=1}^{m} k_n(s,t)\right)_{m\geq 1}$$

está formada por funciones continuas.

Además esta sucesión de sumas parciales converge a la función h uniformemente en $[a,b] \times [a,b]$ porque

$$|h(s,t) - \sum_{n=1}^{m} k_n(s,t)| = |\sum_{n=m+1}^{m} k_n(s,t)| \leq M\sum_{n=m+1}^{m} (M(b-a))^{n-1}$$

tiende a 0 cuando m tiende a ∞ independientemente de (s,t). Entonces h es continua por ser límite uniforme (en Δ) de una sucesión de funciones continuas, lo que completa la prueba.

□

Definición 18 *La función*

$$h(s,t) := \sum_{n=1}^{\infty} k_n(s,t).$$

definida bajo las condiciones de la proposición anterior se llama núcleo resolvente *de la ecuación integral de Volterra dada.*

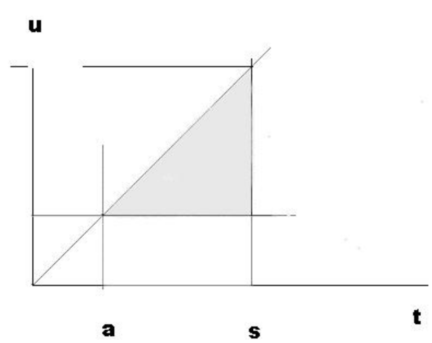
Proposición 16 *En las condiciones de la proposición anterior, si* $Kx(s) := \int_a^s k(s,t)x(t)dt$, *se cumple que, para toda función* $y \in \mathcal{C}([a,b])$, *y todo* $s \in [a,b]$,

$$K^n y(s) = \int_a^s k_n(s,t)y(t)dt.$$

para cada $s \in [a,b]$.

Prueba : Se prueba también por inducción. Para $n = 1$ es evidente por definición de K.

Supongamos que $K^n y(s) = \int_a^s k_n(s,t)y(t)dt$. Entonces una aplicación sencilla del teorema de Fubini en el recinto triangular



$\Delta_s := \{(t,u) : a \le u \le t \le s\}$ nos da

$$\begin{aligned} K_{n+1}y(s) &= K(K^n y)(s) = \textstyle\int_a^s k_1(s,t)(K^n y)(t)dt \\ &= \textstyle\int_{t=a}^{t=s} k_1(s,t)(\int_{u=a}^{u=t} k_n(t,u)y(u)du)dt \\ &= \textstyle\int_a^s(\int_a^t k_1(s,t)k_n(t,u)y(u)du)dt \\ &= \textstyle\iint_{\Delta_s} k(s,t)k_n(t,u)y(u)dtdu \\ &= \textstyle\int_{u=a}^{u=s} y(u)(\int_{t=u}^{t=s} k_1(s,t)k_n(t,u)dt)du \\ &= \textstyle\int_{u=a}^{u=s} y(u)k_{n+1}(s,u)du \\ &= \textstyle\int_a^s k_{n+1}(s,t)y(t)dt, \end{aligned}$$

lo que completa la prueba.

□

TEOREMA 24 *Si $k : \Delta \to \mathbb{R}$ es una función continua entonces la ecuación integral de Volterra de segunda especie*

$$x(s) - \int_a^s k(s,t)x(t)dt = y(s)$$

tiene una solución $x \in \mathcal{C}([a,b]$, dada por

$$x(s) = y(s) + \int_a^s h(s,t)y(t)dt.$$

donde h es el núcleo resolvente suma de la serie de Neumann de núcleos iterados (k_n) generados por el núcleo k.

Prueba : Observemos en primer lugar que la sucesión de operadores lineales y continuos

$$K^n : \mathcal{C}([a,b]) \to \mathcal{C}([a,b]),$$

donde $Kx(s) = \int_a^s k(s,t)x(t)dt$, genera una serie que es absolutamente convergente en el espacio de los operadores de $(\mathcal{C}([a,b]), \|\cdot\|_\infty)$.

En efecto, según la proposición anterior, fijada una función $\varphi \in \mathcal{C}([a,b])$, para todo entero positivo n y para todo $s \in [a,b]$

$$|K^n\varphi(s)| = \left|\int_a^s k_n(s,t)\varphi(t)dt\right| \leq \int_a^s |k_n(s,t)|\|\varphi\|_\infty dt$$

donde k_n es el $n-$ésimo núcleo iterado a partir de k.

Usando la acotación que conocemos para k_n tendremos que

$$\begin{array}{rl} |K^n\varphi(s)| & \leq \int_a^s |k_n(s,t)|\|\varphi\|_\infty dt \\ & \leq \int_a^s \frac{M^n(s-t)^{n-1}}{(n-1)!}\|\varphi\|_\infty dt \\ & = \frac{M^n}{(n-1)!}\left[\frac{-(s-t)^n}{n}\right]_{t=a}^{t=s}\|\varphi\|_\infty \\ & = \frac{M^n}{n!}(s-a)^n\|\varphi\|_\infty \\ & \leq \frac{M^n(b-a)^n}{n!}\|\varphi\|_\infty. \end{array}$$

Por tanto, $\|K^n\varphi\|_\infty \leq \frac{M^n(b-a)^n}{n!}\|\varphi\|_\infty$ lo que, por definición de norma de un operador, nos da la desigualdad

$$\|K^n\| \leq \frac{[M(b-a)]^n}{n!}$$

para cada entero positivo n.

De aquí se sigue que la serie $\sum_{n=0}^\infty \|K^n\|$ es convergente con

$$\sum_{n=0}^\infty \|K^n\| \leq \sum_{n=0}^\infty \frac{[M(b-a)]^n}{n!} = e^{M(b-a)}.$$

Como $\mathcal{B}\left(\mathcal{C}([a,b]),\mathcal{C}([a,b])\right)$ es un espacio de Banach , de ser absolutamente convergente la serie $\sum_{n=0}^{\infty} K^n$, se sigue que es convergente.

Por tanto, según vimos, $Id - K$ es invertible con

$$(Id-K)^{-1} = \sum_{n=0}^{\infty} K^n.$$

Por tanto la ecuación $(Id - K)x = y$ tendrá la solución

$$x = (Id-K)^{-1}(y) = [\sum_{n=0}^{\infty} K^n](y) = y + \sum_{n=1}^{\infty} \int_a^s k_n(s,t)y(t)dt.$$

es decir, para cada $s \in [a,b]$,

$$x(s) = y(s) + \sum_{n=1}^{\infty} \int_a^s k_n(s,t)y(t)dt.$$

Pero fijado $s \in [a,b]$, la sucesión de funciones

$$t \mapsto s_m(s,t) := \sum_{n=1}^{m} (k_n(s,t)y(t))$$

según hemos visto antes verificará

$$|s_m(s,t)| \leq \sum_{n=1}^{m} |k_n(s,t)||y(t)| \leq \|y\|_\infty \sum_{n=1}^{m} |k_n(s,t)|$$

$$\leq \|y\|_\infty \sum_{n=1}^{\infty} \frac{[M(b-a)]^n}{n!} = \|y\|_\infty (e^{M(b-a)} - 1).$$

Es decir, todas las funciones s_m están domidadas en $[a,b]$ por una función constante, que desde luego es integrable en $[a,b]$ (y también en $[a,s]$). Podemos aplicar el teorema de la convergencia dominada para concluir

$$\int_a^s \lim_m s_m(s,t)dt = \lim_m \int_a^s s_m(s,t)dt$$

o sea,

$$\int_a^s h(s,t)y(t)dt = \lim_m \int_a^s s_m(s,t)dt.$$

Sustituyendo, queda que

$$\begin{aligned} x(s) &= y(s) + \lim_m \left[\int_a^s \left(\sum_{n=1}^m k_n(s,t)y(t)\right)dt\right] \\ &= y(s) + \lim_m \left[\int_a^s s_m(s,t)dt\right] \\ &= y(s) + \int_a^s h(s,t)y(t)dt. \end{aligned}$$

lo que completa la prueba.

$\square$

Ejemplo 22 *Resolveremos por el método de los núcleos iterados la ecuación de Volterra de segunda clase*

$$x(s) - \int_0^s \lambda e^{(s-t)} x(t)dt = f(s).$$

En este caso $k(s,t) := \lambda e^{(s-t)} = k_1(s,t)$.

Calculamos directamente

$$k_2(s,t) := \int_t^s k_1(s,u)k_1(u,t)du = \int_t^s \lambda e^{(s-u)} \lambda e^{(u-t)} du = \lambda^2 e^{(s-t)}(s-t)$$

$$\begin{aligned} k_3(s,t) &:= \int_t^s k_1(s,u)k_2(u,t)du = \int_t^s \lambda e^{(s-u)} \lambda^2 e^{(u-t)}(u-t)du \\ &= \lambda^3 e^{(s-t)} \left[\frac{(u-t)^2}{2}\right]_{u=t}^{u=s} = \lambda^3 e^{(s-t)} \frac{(s-t)^2}{2}. \\ k_4(s,t) &:= \int_t^s k_1(s,u)k_3(u,t)du = \int_t^s \lambda e^{(s-u)} \lambda^3 e^{(u-t)} \frac{(u-t)^2}{2!} du \\ &= \lambda^4 e^{(s-t)} \left[\frac{(u-t)^3}{2!3}\right]_{u=t}^{u=s} = \lambda^4 e^{(s-t)} \frac{(s-t)^3}{3!}. \end{aligned}$$

Podemos pensar pues que

$$k_n(s,t) = \lambda^n e^{(s-t)} \frac{(s-t)^{(n-1)}}{(n-1)!}.$$

Supuesto que esta igualdad sea cierta,

$$\begin{aligned} k_{n+1}(s,t) &:= \int_t^s k_1(s,u)k_n(u,t)du = \int_t^s \lambda e^{(s-u)} \lambda^n e^{(u-t)} \frac{(u-t)^(n-1)}{(n-1)!} du \\ &= \lambda^{n+1} e^{(s-t)} \left[\frac{(u-t)^n}{(n-1)!n}\right]_{u=t}^{u=s} = \lambda^{n+1} e^{(s-t)} \frac{(s-t)^n}{n!}. \end{aligned}$$

Luego por inducción hemos visto que, efectivamente, la fórmula pensada para k_n es correcta.

A partir de aquí el núcleo resolvente de esta ecuación será

$$h(s,t) := \sum_{n=1}^{\infty} k_n(s,t) = \lambda \sum_{n=1}^{\infty} e^{(s-t)} \frac{(\lambda(s-t))^{n-1}}{(n-1)!} =$$

$$= \lambda e^{(s-t)} e^{\lambda(s-t)} = \lambda e^{(1+\lambda)(s-t)}.$$

Con el núcleo resolvente encontramos la solución de la ecuación propuesta:

$$x(s) := f(s) + \int_0^s \lambda e^{(1+\lambda)(s-t)} f(t)dt.$$

3.4. Ecuaciones integrales y aplicaciones contractivas (*)

Recordemos el siguiente:

Teorema 25 *Banach, 1922*
Sea (X, d) un espacio métrico completo. Sea $T : X \to X$ una aplicación contractiva*, es decir tal que existe k con $0 < k < 1$ tal que, para cualesquiera $x, y \in X$,*

$$d(T(x), T(y)) \leq kd(x, y).$$

Entonces, T tiene un único punto fijo *$x = T(x) \in X$, y para cada $x_0 \in X$ se tiene que*

$$d\left(T^n(x_0), x\right) \leq \frac{k^n}{1-k} d(x_0, T(x_0)).$$

Prueba

La primera idea básica es que, a partir de cualquier punto $x_0 \in X$, se obtiene una sucesión de iteraciones,

$$(x_n) := (x_0, T(x_0), T^2(x_0), \ldots, T^n(x_0), \ldots)$$

que es de Cauchy.

En efecto,

$$\begin{aligned} d(x_n, x_{n+1}) &= d(T(x_{n-1}), T(x_n)) \leq \\ &\leq kd(x_{n-1}, x_n) \leq k^2 d(x_{n-2}, x_{n-1}) \leq \ldots \\ &\leq k^n d(x_0, x_1). \end{aligned}$$

Por tanto, si n, m son enteros positivos

$$\begin{aligned} d(x_n, x_{n+m}) \leq &\quad d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \ldots + d(x_{n+m-1}, x_{n+m}) \\ \leq &\quad [k^n + k^{n+1} + \ldots + k^{n+m-1}] d(x_0, x_1) \\ \leq &\quad [k^n + k^{n+1} + \ldots + k^{n+m-1} + \ldots] d(x_0, x_1) \\ = &\quad \tfrac{k^n}{1-k} d(x_0, x_1). \end{aligned}$$

Luego, dado $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que si $p \geq n_0$

$$\frac{k^p}{1-k} d(x_0, x_1) < \varepsilon,$$

y entonces, si $p, q \geq n_0$, siempre podemos poner $q = p + m$ y se tendrá

$$d(x_p, x_q) = d(x_p, x_{p+m}) \leq \frac{k^p}{1-k} d(x_0, x_1) < \varepsilon,$$

luego efectivamente la sucesión (x_n) es de Cauchy.

Por ser (X, d) un espacio métrico completo existe

$$x := \lim_{n\to\infty} x_n$$

La aplicación T es contractiva, luego es continua e incluso uniformemente continua en X, por lo que

$$T(x) = T(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} T(x_n) = \lim_{n\to\infty} x_{n+1} = x.$$

Así pues, x es un punto fijo de T, es decir una solución de la ecuación $T(x) = x$ en M.

Esta solución es única, pues si existiera $x' \in M$ con $x \neq x' = T(x')$ se tendría la siguiente contradicción

$$0 \leq d(x, x') = d(T(x), T(x')) \leq kd(x, x') < d(x, x').$$

De la desigualdad

$$d(x_n, x_{n+m}) \leq \frac{k^n}{1-k} d(x_0, x_1)$$

se sigue, (recordando la continuidad de la distancia d), que

$$d(x_n, x) = \lim_{m\to\infty} d(x_n, x_{n+m}) \leq \frac{k^n}{1-k} d(x_0, x_1),$$

la cota de error del enunciado.

□

Este teorema puede aplicarse también a la solución de las ecuaciones integrales que hamos estudiado en este capítulo pues si T es un operador integral en $\mathcal{C}([a,b])$, como los que hemos usado en el apartado anterior, y consideramos la ecuación integral en $\mathcal{C}([a,b])$,

$$x - T(x) = y$$

(donde $y \in \mathcal{C}([a,b])$ es un dato), entonces la aplicación *afín* $U : \mathcal{C}([a,b]) \to \mathcal{C}([a,b])$ dada por

$$U(x) = y + T(x)$$

tiene por constante de Lipschitz $\|T\|$, ya que

$$\|U(x) - U(y)\| = \|T(x) - T(y)\| \leq \|T\| \|x - y\|.$$

Por tanto, si $\|T\| < 1$ entonces U es contractiva. Además es claro que $U(x) = x$ si y sólo si $x - T(x) = y$, es decir un elemento $\tilde{x} \in \mathcal{C}([a,b])$ es solución de la ecuación integral anterior si y sólo si $\tilde{x}$ es un punto fijo de U.

Como $\mathcal{C}([a,b])$ es un espacio métrico completo, por el Teorema de Banach del punto fijo, la condición $\|T\| < 1$ nos garantiza la existencia de una solución de la ecuación integral, y además dado cualquier $x_0 \in \mathcal{C}([a,b])$, la sucesión

$$(x_0, U(x_0), U^2(x_0), U^3(x_0), \ldots)$$

converge en $\mathcal{C}([a,b])$ a la solución x de la ecuación intergral $x - T(x) = y$. Además las cotas de error que proporciona el teorema, a veces son una información muy útil en la práctica.

En general, si $U(x) = y + T(x)$, se tendrá

$$U^2(x) = U(U(x)) = y + T(Ux)) = y + T(y + T(x)) = y + T(y) + T^2(x).$$

Si aceptamos que

$$U^n(x) = y + T(y) + \ldots + T^{n-1}(y) + T^n(x),$$

entonces

$$\begin{aligned} U^{n+1}(x) &= U(U^n(x)) \\ &= y + T(U^n(x)) \\ &= y + T\left(y + T(y) + \ldots + T^{n-1}(y) + T^n(x)\right) \\ &= y + T(y) + \ldots + T^n(y) + T^{n+1}(x). \end{aligned}$$

Hemos demostrado por inducción una fórmula alternativa para calcular la n-sima iterada del operador U, que, como vemos no es sencilla.

Naturalmente, para cada $x_0 \in \mathcal{C}([a, b])$,

$$U^n(x_0) \to \tilde{x},$$

es decir,

$$\lim_{n\to\infty} [y + T(y) + \ldots + T^n(y) + T^{n+1}(x_0)] = \tilde{x}$$

o bien

$$\sum_{n=0}^{\infty} T^n(y) + \lim_{n\to\infty} T^{n+1}(x_0) = \tilde{x}.$$

Si $\|T\| < 1$, entonces como

$$\|T^{n+1}(x)\| \leq \|T^{n+1}\|\|x\| \leq \|T\|^{n+1}\|x\|,$$

se sigue que

$$\lim_{n\to\infty} T^{n+1}(x_0) = 0_{\mathcal{C}([a,b])}$$

por lo que obtenemos de nuevo que

$$\sum_{n=0}^{\infty} T^n(y) = \tilde{x}.$$

Por supuesto, el primer miembro de la igualdad anterior también puede escribirtse

$$(Id - T)^{-1}(y).$$

Ejemplo 23 Sea T el operador de Fredhölm en $\mathcal{C}([0, 1])$ con núcleo $s(1 - e^{st})$, es decir

$$Tx(t) = \int_0^1 s(1 - e^{st})x(t)dt.$$

Vimos en el ejemplo (18) que $\|T\| = e - 2 < 1$. Por tanto, si consideramos la ecuación integral

$$(x - Tx)(s) = e^s - s$$

tenemos que tiene solución, digamos $\widetilde{x} \in \mathcal{C}([0,1])$ y puede ser obtenida como límite de las iteradas de la aplicación afín $U : \mathcal{C}([0,1]) \to \mathcal{C}([0,1])$ dada por

$$Ux(s) = e^s - s - Tx(s).$$

Esta sucesión de iteradas de la aplicación U puede no ser sencilla de construir en la práctica. Por ejemplo, si partimos de la función $x_0 = \mathcal{C}([0,1])$,

$$Ux_0(s) = e^s - s - Tx_0(s) = e^s - s.$$

Igualmente,

$$U^2x_0(s) = U(U(x_0))(s) = e^s - s - T(Ux_0)(s) = e^s - s - \int_0^1 s(1-e^{st})(e^t - t)dt,$$

resulta

$$U^2x_0(s) = e^s - s - \frac{2e^s(s^2(e-1)+1) + s^3(3-2e) + s^2(1-2e) - 2(s+1)}{2s(s+1)},$$

y vemos que su expresión es muy complicada.

4. ESPECTRO DE UN OPERADOR. OPERADORES COMPACTOS

Empezaremos con dos observaciones muy sencillas:

PROPOSICIÓN 17 *En un espacio de Hilbert,* $(X, \langle ., . \rangle)$*, si* $v, w \in X$ *son vectores tales que para todo* $x \in X$

$$\langle x, v \rangle = \langle x, w \rangle$$

se sigue que $w = v$.

Prueba : Por la igualdad anterior para todo $x \in X$ $\langle x, v \rangle - \langle x, w \rangle = \langle x, v - w \rangle = 0$, y basta tomar $x = v - w$ para tener $\langle v - w, v - w \rangle = 0$, es decir $\|v - w\|^2 = 0$, por lo que $v = w$.

□

PROPOSICIÓN 18 *En todo espacio con producto escalar,* X*, para cualquier vector* $x \in X$*,*

$$\|x\| = \sup\{|\langle x, v \rangle|, v \in X, \|v\| \leq 1\}. \tag{9}$$

Prueba : La igualdad es trivial si $x = 0_X$. Si $x \neq 0_X$ y si s es el segundo miembro de 9, por la desigualdad de Cauchy Schwarz, para todo $v \in X$ con $\|v\| \leq 1$ se tiene que $|\langle x, v \rangle| \leq \|x\|\|v\| \leq \|x\|$ luego $s \leq \|x\|$. Como $\left\| \frac{x}{\|x\|} \right\| = 1$,

$$s \geq \left| \left\langle x, \frac{x}{\|x\|} \right\rangle \right| = \|x\|,$$

luego efectivamente, $s = \|x\|$.

□

4.1. Operador adjunto

TEOREMA 26 *Si* X, Y *son espacios de Hilbert complejos (o en general, sobre el mismo cuerpo) y* T *es un operador (aplicación lineal y continua) entre* X *e* Y *entonces existe un único operador* $T^* : Y \to X$ *tal que para cualesquiera* $x \in X$, $y \in Y$,

$$\langle T(x), y \rangle = \langle x, T^*(y) \rangle.$$

Prueba : Fijado y en Y la aplicación

$$x \mapsto \langle T(x), y \rangle =: f_{T,y}(x)$$

toma valores en $\mathbb{C}$, y es lineal (por serlo T y $v \mapsto \langle v, y \rangle$). También $f_{T,y}$ es continua con norma menor o igual que $\|T\|\|y\|$, pues por la desigualdad de Cauchy Schwarz,

$$|f_{T,y}(x)| = |\langle T(x), y \rangle| \leq \|Tx\|\|y\| \leq \|T\|\|x\|\|y\|.$$

Sabemos que, por ser X un espacio de Hilbert, existe un único vector $w \in X$ tal que para todo x de X,

$$f_{T,y}(x) = \langle T(x), y \rangle = \langle x, w \rangle.$$

Como este vector w depende del $y \in Y$ previamente fijado y también del operador T, lo denotaremos por $T^*(y)$.

En resumen podemos afirmar que $T^*(y)$ es el único vector de X que cumple

$$\langle T(x), y \rangle = \langle x, T^*(y) \rangle$$

para todo $x \in X$.

Observar que el producto escalar en el primer miembro de la igualdad anterior es el de Y, mientras que el del segundo miembro es el de X.

Que así se obtiene un operador (aplicación lineal y continua) $T^* : Y \to X$, se prueba del siguiente modo:

Sean $x \in X$, $y_1, y_2 \in Y$, $\alpha, \beta \in \mathbb{C}$:

$$\begin{aligned}
\langle x, T^*(\alpha y_1 + \beta y_2) \rangle &:= \langle T(x), (\alpha y_1 + \beta y_2) \rangle \\
&= \langle T(x), \alpha y_1 \rangle + \langle T(x), \beta y_2 \rangle \\
&= \overline{\alpha} \langle T(x), y_1 \rangle + \overline{\beta} \langle T(x), y_2 \rangle \\
&\overset{\text{def.}}{=} \overline{\alpha} \langle x, T^*(y_1) \rangle + \overline{\beta} \langle x, T^*(y_2) \rangle \\
&= \langle x, \alpha T^*(y_1) \rangle + \langle x, \beta T^*(y_2) \rangle \\
&= \langle x, \alpha T^*(y_1) + \beta T^*(y_2) \rangle
\end{aligned}$$

Como se cumple para todo $x \in X$, que

$$\langle x, T^*(\alpha y_1 + \beta y_2) \rangle = \langle x, \alpha T^*(y_1) + \beta T^*(y_2) \rangle,$$

por la proposición 17 obtenemos:

$$T^*(\alpha y_1 + \beta y_2) = \alpha T^*(y_1) + \beta T^*(y_2).$$

Fijado $y \in Y$, aplicamos la proposición 18 y la definición de T^* para tener

$$\begin{aligned}
\|T^*(y)\| &= \sup\{|\langle T^*(y), v \rangle|, v \in X, \|v\| \leq 1\} \\
&= \sup\{|\langle v, T^*(y) \rangle|, v \in X, \|v\| \leq 1\} \\
&= \sup\{|\langle T(v), y \rangle|, v \in X, \|v\| \leq 1\} \\
&\leq \sup\{\|T(v)\| \|y\|, v \in X, \|v\| \leq 1\} \\
&= \sup\{\|T(v)\|, v \in X, \|v\| \leq 1\} \|y\| \\
&= \|T\| \|y\|,
\end{aligned}$$

luego T^* es continuo, con $\|T^*\| \leq \|T\|$.

$\square$

Este teorema hace posible la siguiente

Definición 19 *Si X, Y son espacios de Hilbert complejos y T es un operador (aplicación lineal y continua) entre X e Y, se define como* operador adjunto *de T al $T^* : Y \to X$ que es el único elemento de $\mathcal{B}(Y, X)$ que verifica, para cualesquiera $x \in X$, $y \in Y$ que*

$$\langle T(x), y\rangle = \langle x, T^*(y)\rangle.$$

Ejemplo 24 En el espacio euclídeo ordinario sea $A : \mathbb{R}^n \to \mathbb{R}^m$ el operador que tiene por matriz asociada (a_{ij}) (respecto de las bases canónicas $(e^{(1)}, \ldots, e^{(n)})$, $(e^{(1)}, \ldots, e^{(m)})$ Sea A^* el operador adjunto de A, y sea (α_{ij}) la matriz asociada este operador A^*. Para cada $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, m\}$,

$$\left\langle A(e^{(i)}), e^{(j)}\right\rangle = \left\langle e^{(i)}, A^*(e^{(j)})\right\rangle,$$

o sea

$$\left\langle a_{1i}e^{(1)} + \ldots + a_{mi}e^{(n)}, e^{(j)}\right\rangle = \left\langle e^{(i)}, \alpha_{1j}e^{(1)} + \ldots + \alpha_{nj}e^{(n)}\right\rangle,$$

es decir

$$a_{ji} = \alpha_{ij}.$$

Por tanto, $\big(\alpha_{ij}\big) = \big(a_{ij}\big)^t$.

Ejemplo 25 *Construiremos el operador adjunto de $S : \ell_2 \to \ell_2$ dado por*

$$S(x_1, x_2, \ldots) := (0, x_1, x_2, \ldots).$$

Se caracterizará $S^* : \ell_2 \to \ell_2$ por

$$\langle S(x), y\rangle = \langle x, S^*(y)\rangle.$$

para cualesquiera $x, y \in \ell_2$. Es decir,

$$\langle (0, x_1, x_2, \ldots), (y_1, y_2, \ldots)\rangle = \langle (x_1, x_2, \ldots), S^*(y_1, y_2, \ldots)\rangle$$

o, lo que es igual

$$x_1y_2 + x_2y_3 + x_3y_4 + \ldots = \langle x, S^*(y)\rangle.$$

Pero entonces

$$\langle (x_1, x_2, \ldots), (y_2, y_3, \ldots)\rangle = x_1y_2 + x_2y_3 + x_3y_4 + \ldots = \langle x, S^*(y)\rangle.$$

es decir, para todo $x \in \ell_2$

$$\langle x, (y_2, y_3, \ldots)\rangle = x_1y_2 + x_2y_3 + x_3y_4 + \ldots = \langle x, S^*(y)\rangle.$$

Esto nos da inmediatamente que

$$S^*(y) = (y_2, y_3, \ldots).$$

Ejemplo 26 Sea X el espacio de Hilbert complejo ℓ_2, y sea $T : X \to X$ dado por

$$T\big((z_n)\big) = (i^n z_n).$$

Calculemos $T^* : X \to X$.

Pongamos $(w_n) = T(y_n)$ para $(y_n) \in X$.
Para cualesquiera $z = (z_n), y = (y_n) \in X$,

$$\langle T(z), y\rangle = \langle z, T^*(y)\rangle,$$

o sea

$$\langle (i^n z_n), (y_n)\rangle = \langle (z_n), (w_n)\rangle,$$

que se traduce en

$$\sum_{n=1}^{\infty} i^n z_n \overline{y_n} = \sum_{n=1}^{\infty} z_n \overline{w_n}.$$

Si escribimos $\overline{a_n + b_n i} = i^n$, la igualdad anterior quedará

$$\sum_{n=1}^{\infty} z_n \overline{(a_n + ib_n) y_n} = \sum_{n=1}^{\infty} z_n \overline{w_n},$$

que puede escribirse como

$$\langle (z_n), ((a_n + ib_n) y_n)\rangle = \langle (z_n), (w_n)\rangle.$$

Por la proposición 17,

$$\big((a_n + ib_n) y_n\big) = (w_n).$$

Calculamos a_n y b_n:

$$\overline{a_n + b_n i} = i^n \Rightarrow a_n + b_n i = \overline{i^n} = (\bar{i})^n = (-i)^n = (-1)^n i^n.$$

Por tanto,

$$(w_n) = T^*\big((y_n)\big) = \big((a_n + ib_n) y_n\big) = \big((-1)^n i^n y_n\big).$$

El lema siguiente es una forma alternativa para calcular normas de operadores entre espacios de Hilbert.

Lema 4 *Sean X, Y espacios de Hilbert. Sea $T \in \mathcal{B}(X, Y)$. Entonces*

$$\|T\| = \sup\{|\langle T(x), y\rangle| : \|x\| \leq 1,\ \|y\| \leq 1\}.$$

Prueba : Por la desigualdad de Cauchy Schwarz si $y \in B_Y$, y $x \in B_X$

$$|\langle T(x), y\rangle| \leq \|T(x)\|\|y\| \leq \|T(x)\| \leq \|T\|\|x\| \leq \|T\|$$

y tomando supremos,

$$s := \sup\{|\langle T(x), y\rangle| : \|x\| \leq 1,\ \|y\| \leq 1\} \leq \|T\|.$$

Sea ahora $x \in B_X$ cualquier vector con $T(x) \neq 0_Y$. Entonces $\frac{1}{\|T(x)\|}T(x) \in B_Y$, luego por definición de s,

$$s \geq \left|\left\langle T(x), \frac{1}{\|T(x)\|}T(x)\right\rangle\right| = \|T(x)\|,$$

y esta desigualdad se cumple trivialmente si $T(x) = 0_Y$.

Por tanto,

$$s \geq \sup\{\|Tx\| : x \in B_X\} =: \|T\|.$$

□

Proposición 19 *Si X, Y son espacios de Hilbert sobre el mismo cuerpo y T es un operador (aplicación lineal y continua) entre X e Y, entonces $(T^*)^* = T$, y $\|T\| = \|T^*\|$.*

Prueba : Sean $x \in X$, $y \in Y$. Por definición de $(T^*)^*$,

$$\begin{aligned} \langle x, (T^*)^*(y)\rangle &:= \langle T^*(x), y\rangle \\ &= \overline{\langle y, T^*(x)\rangle} \\ &= \overline{\langle T(y), x\rangle} \\ &= \langle x, T(y)\rangle \end{aligned}$$

Una vez más, como la igualdad se cumple para todo $x \in X$, resulta $(T^*)^*(y) = T(y)$.

Por el lema anterior

$$\begin{aligned} \|T^*\| &= \sup\{|\langle T^*(y), x\rangle| : \|y\| \leq 1,\ \|x\| \leq 1\} \\ &= \sup\{|\langle x, T^*(y)\rangle| : \|y\| \leq 1,\ \|x\| \leq 1\} \\ &= \sup\{|\langle T(x), y\rangle| : \|y\| \leq 1,\ \|x\| \leq 1\} \\ &= \|T\|. \end{aligned}$$

□

Proposición 20 *Si X, Y son espacios de Hilbert sobre el mismo cuerpo, para cualesquiera $S, T \in \mathcal{B}(X, Y)$ y escalares α, β,*

$$(\alpha S + \beta T)^* = \overline{\alpha} S^* + \overline{\beta} T^*.$$

Prueba : Sean $x \in X$, $y \in Y$.

$$\begin{aligned} \langle x, (\alpha S + \beta T)^*(y)\rangle &:= \langle(\alpha S + \beta T)(x), y\rangle \\ &= \langle \alpha S(x), y\rangle + \langle \beta T(x), y\rangle \\ &= \alpha\langle x, S^*(y)\rangle + \beta\langle x, T^*(y)\rangle \\ &= \langle x, \overline{\alpha} S^*(y)\rangle + \langle x, \overline{\beta} T^*(y)\rangle \\ &= \langle x, (\overline{\alpha} S^* + \overline{\beta} T^*)(y)\rangle \end{aligned}$$

Como esta igualdad se cumple para todo $x \in X$, ha de ser

$$(\alpha S + \beta T)^*(y) = (\overline{\alpha} S^* + \overline{\beta} T^*)(y),$$

lo que completa la prueba.

□

4.1.1. Otra expresión del operador adjunto

Operador conjugado Recordemos que si X, Y son espacios de Banach sobre el mismo cuerpo y T es un operador (aplicación lineal y continua) entre X y Y se define como operador conjugado de T al $T^t : Y^* \to X^*$ al dado por $T^t(g) = g \circ T$.

Es inmediato que T^t es lineal. También es continua, pues para cada $g \in X_2^*$

$$\begin{aligned} \|T^t(g)\| &:= \sup\{|T^t(g)(x_1) : x_1 \in X_1, \|x_1\| \leq 1\} \\ &= \sup\{|(g \circ T)(x_1) : x_1 \in X_1, \|x_1\| \leq 1\} \\ &= \sup\{|g(T(x_1)) : x_1 \in X_1, \|x_1\| \leq 1\} \\ &\leq \sup\{\|g\| \|T(x_1)\| : x_1 \in X_1, \|x_1\| \leq 1\} \\ &\leq \sup\{\|g\| \|T\| \|x_1\| : x_1 \in X_1, \|x_1\| \leq 1\} = \|T\| \|g\|, \end{aligned}$$

luego T^t es continua con $\|T^t\| \leq \|T\|$.

El caso de los espacios de Hilbert Los espacios de Hilbert se pueden identificar con sus respectivos duales. Entonces el concepto de aplicación conjugada en ciento modo se hace más sencillo:

Proposición 21 *Si X, Y son espacios de Hilbert sobre el mismo cuerpo y T es un operador (aplicación lineal y continua) entre X e Y,*

$$T^* := J_1^{-1} \circ T^t \circ J_2.$$

donde:

J_1, J_2 son las isometrías de X_1, X_2 en sus respectivos duales.

$T^t : X_2^* \to X_1^*$ es la aplicación conjugada de T.

Prueba : Recordemos que $J_1^{-1}(f)$, para $f \in X_1^*$ es el elemento x de X_1 tal que $f(v) = \langle v, x \rangle$ para todo $v \in X_1$. Entonces para todo $v \in X_1$, $f(v) = \langle v, J_1^{-1}(f) \rangle$.

De igual modo, $J_2(y)$ para un vector y de X_2 es la forma lineal continua $w \mapsto \langle w, y \rangle$.

Sean $x \in X$ y $y \in Y$.

$$\begin{aligned} \langle x, J_1^{-1}(T^t(J_2(y))) \rangle &= (T^t(J_2(y)))(x) \\ &= (J_2(y) \circ T)(x) \\ &= J_2(y)(T(x)) \\ &= \langle T(x), y \rangle \\ &= \langle x, T^*(y) \rangle \end{aligned}$$

□

4.1.2. Caso particular de $\mathcal{B}(X)$.

La teoría de adjuntos de operadores en espacios de Hilbert es más rica en resultados cuando $X = Y$, como también lo es la teoría de operadores en general.

Proposición 22 *Sean $S, T \in \mathcal{B}(X)$. Se verifica:*

a) $(Id_X)^* = Id_X$.

b) $(S \circ T)^* = T^* \circ S^*$.

c) T es invertible si y sólo si T^* es invertible, y en ese caso

$$(T^*)^{-1} = (T^{-1})^*.$$

Prueba : (a) Si $x, y \in X$, Por definición de operador adjunto,

$$\langle x, y\rangle = \langle Id_X(x), y\rangle = \langle x, (Id_X)^*(y)\rangle$$

luego $\langle x, y\rangle = \langle x, (Id_X)^*(y)\rangle$ para todo x de X, luego $y = Id_X(y) = (Id_X)^*(y)$ para todo y de X.

(b) Si $x, y \in X$, teniendo en cuenta que $(T^*)^* = T$ y $(S^*)^* = S$,

$$\begin{aligned}
\langle x, (T^* \circ S^*)(y)\rangle &= \langle x, T^*(S^*(y))\rangle \\
&= \langle T(x), S^*(y)\rangle \\
&= \langle S(T(x)), y\rangle \\
&= \langle (S \circ T)(x), y\rangle \\
&= \langle x, (S \circ T)^*(y)\rangle
\end{aligned}$$

Como esta igualdad se cumple para todo $x \in X$,

$$(T^* \circ S^*)(y) = (S \circ T)^*(y).$$

luego $(T^* \circ S^*) = (S \circ T)^*$.

(c) Supongamos que exista $T^{-1} \in \mathcal{B}(X, X)$. Fijado $x \in X$, para todo $y \in X$,

$$\begin{aligned}
\langle x, y\rangle &= \langle T(T^{-1}(x)), y\rangle \\
&= \langle T^{-1}(x), T^*(y)\rangle \\
&= \langle x, (T^{-1})^*(T^*(y))\rangle \\
&= \langle x, ((T^{-1})^* \circ T^*)(y)\rangle.
\end{aligned}$$

Como esta igualdad se cumple para todo $x \in Y$, de nuevo por la proposición 17,

$$Id_X(y) = y = ((T^{-1})^* \circ T^*)(y),$$

y dado que esta igualdad se cumple para cualquier y de X concluimos que el operador $(T^{-1})^*$ es el inverso de T^* por la izquierda.

Repitiendo un argumento parecido, como

$$\begin{aligned}\langle x,y\rangle &= \langle T^{-1}(T(x)),y\rangle \\ &= \langle T(x),(T^{-1})^*(y)\rangle \\ &= \langle x,T^*((T^{-1})^*(y))\rangle \\ &= \langle x,(T^*\circ(T^{-1})^*)(y)\rangle\end{aligned}$$

también para todo $y \in X$

$$Id_X(y) = y = (T^*\circ(T^{-1})^*)(y)$$

luego el operador $(T^{-1})^*$ es el inverso de T^* por la derecha. Por tanto,

$$(T^*)^{-1} = (T^{-1})^*.$$

□

Ejemplo 27 Si M es un subespacio cerrado de un espacio de Hilbert, X, el operador P_M (proyección ortogonal sobre M) cumple la siguiente propiedad:

$$\langle x,P_M(y)\rangle = \langle P_M(x),y\rangle$$

para cualesquiera $x,y \in X$.

Por lo tanto,

$$\langle x,P_M(y)\rangle = \langle P_M(x),y\rangle = \langle x,(P_M)^*(y)$$

Esto nos da que $P_M = (P_M)^*$, es decir que P_M coincide con su adjunto.

El ejemplo anterior motiva la siguiente definición:

Definición 20 *Un operador $T \in \mathcal{B}(X)$ se llama autoadjunto si $T^* = T$.*

Ejemplo 28 *Si $T \in \mathcal{B}(X)$, entonces el operador $T\circ T^*$ es autoadjunto.*

Basta tener en cuenta que

$$(T\circ T^*)^* = (T^*)^*\circ T^* = T\circ T^*.$$

Ejemplo 29 Consideremos el espacio de Hilbert real $X = L^2([a,b])$ y el operador de Fredholm $T : X \to X$, dado por

$$T(x)(s) := \int_a^b k(s,t)x(t)dt$$

donde el núcleo k es una función continua en $[a, b] \times [a.b]$ Sean $x, y \in X$. Una aplicación sencilla del teorema de Fubini nos da

$$\begin{aligned}
\langle x, T^*(y)\rangle &:= \langle T(x), y\rangle \\
&= \int_a^b Tx(s)y(s)ds \\
&= \int_a^b \left(\int_a^b k(s,t)x(t)dt\right) y(s)ds \\
&= \int_a^b \left(\int_a^b k(s,t)x(t)y(s)dt\right) ds \\
&= \int_a^b \left(\int_a^b k(s,t)x(t)y(s)ds\right) dt \\
&= \int_a^b x(t) \left(\int_a^b k(s,t)y(s)ds\right) dt \\
&= \int_a^b x(s) \left(\int_a^b k(t,s)y(t)dt\right) ds \\
&= \langle x, h\rangle,
\end{aligned}$$

con $h(s) = \int_a^b k(t, s)y(t)dt$. Como $x \in X$ es arbitrario, obtenemos que $T^*(y) = h$, es decir,

$$T^*(y)(s) = \int_a^b k(t, s)y(t)dt.$$

Por tanto si k es simétrico, o sea si $k(s, t) = k(t, s)$ entonces $T = T^*$, es decir que T es autoadjunto.

4.2. Espectro, valores y vectores propios

A lo largo de esta sección se supone dado un espacio de Banach complejo $(X, \|\cdot\|)$.

Hemos visto en temas anteriores que algunas ecuaciones integrales pueden expresarse de modo abstracto como $(Id - T)(x) = y$, donde y es un dato, x la función incógnita, y K cierto operador en un espacio de funciones (usualmente $\mathcal{C}([a, b])$). También hemos visto que, si existe la aplicación lineal $(Id-T)^{-1}$, entonces podrá resolverse formalmente la ecuación $(Id - T)(x) = y$ haciendo $(Id - T)^{-1}(Id - T)(x) = (Id - T)^{-1}(y)$, es decir, $x = (Id-T)^{-1}(y)$. Esto sugiere que es importante que el operador $Id-T$ admita inverso.

Definición 21 *El* espectro *de un operador $T \in \mathcal{B}(X)$ es el subconjunto $sp(T)$ del plano complejo definido por*

$$sp(T) := \{\lambda \in \mathbb{C} : \lambda Id - T \;\; \textit{es no invertible}\}(= \{\lambda \in \mathbb{C} : T - \lambda Id \;\; \textit{es no invertible}\}).$$

Cada elemento de $sp(T)$ suele llamarse valor espectral *de T. Por tanto, para todo número complejo λ no perteneciente al espectro de T, llamado* valor resolvente (o valor regular) de *T existe el operador $(\lambda Id - T)^{-1}$. El conjunto de los valores resolventes de T, o sea $\mathbb{C}\backslash sp(T)$ se llama* conjunto resolvente *de T, denotándose por $\rho(T)$.* [5]

[5] Existe y es continuo $(\lambda Id - T)^{-1} : X \to X$ si y sólo si existe y es continuo $(T - \lambda Id)^{-1} : X \to X$.

También es corriente la notación $\sigma(T)$ en lugar de $sp(T)$. Otra notación frecuentemente usada es $R_\lambda(T)$ o simplemente R_λ en lugar de $(\lambda Id - T)^{-1}$.

Naturalmente, para el caso de operadores entre espacios normados reales, suele definirse el espectro como el conjunto de aquellos números reales que no son valores resolventes de T. En todo caso debemos entender que un operador $A \in \mathcal{B}(X)$ es invertible cuando A^{-1} existe como aplicación lineal y además es continua, es decir $A^{-1} \in \mathcal{B}(X)$.

Ejemplo 30 *Si $T \in \mathcal{B}(X)$ es no invertible, entonces $0 \in \sigma(T)$.*

Aunque la noción de espectro es herramienta importante en mucho campos del Análisis Matemático, su origen está en el problema de la diagonalización de matrices.

Recordemos que si X es de dimensión n y T es un operador en X tal que su matriz respecto de la base $\left(e^1, \ldots, e^n\right)$ tiene nulos todos sus elementos fuera de la diagonal principal, que son $(\lambda_1, \ldots, \lambda_n)$, entonces $T(e^j) = \lambda_j e^j$ para cada $j = 1, \ldots, n$. Recíprocamente, si se quiere encontrar una base respecto de la cual T tiene una matriz diagonal, basta encontrar n vectores linealmente independientes $(v^1, \ldots, v^n)$ tales que $T(v^i) = \lambda_i v^i$, $(i = 1, \ldots, n)$.

Definición 22 *Decimos que el número complejo λ es un* valor propio *o* autovalor *de T si existe un vector no nulo $e \in X$ tal que $T(e) = \lambda e$. En este caso también decimos que e es un* vector propio *(o* autovector*) de T correspondiente al valor propio λ.*

Suele usarse la notación $\sigma_p(T)$ para el conjunto de los valores propios de T. Si para algún vector no nulo $T(e) = \lambda e$ y $T(e) = \mu e$ entonces $\lambda = \mu$. Pero existen infinitos vectores propios correspondientes un valor propio dado de T, pues si e es no nulo y $T(e) = \lambda e$ para todo escalar no nulo a se cumple que $T(ae) = aT(e) = a\lambda e = \lambda(ae)$, es decir que ae es un vector propio también asociado al valor propio λ. Además las suma de dos vectores propios asociados a un mismo valor propio es también un vector propio asociado a ese mismo valor propio. En definitiva tenemos que el conjunto V_λ de los valores propios asociados a un determinado valor propio λ es un subespacio vectorial cerrado de X. La dimensión algebraica de V_λ se llama orden de multiplicidad del valor propio λ. Un valor propio de multiplicidad mayor que 1 se llama múltiple o degenerado.

Si para algún vector $e \neq 0_X$, $T(e) = \lambda e$, esto es equivalente a $(\lambda Id - T)(e) = 0_X$. Por tanto decir que λ es un valor propio de T es tanto como decir que $\lambda Id - T$ tiene un vector no nulo en su núcleo y por tanto $\lambda Id - T$ es no inyectiva. Esto implica que $\lambda Id - T$ es no invertible. En suma, todo valor propio de un operador es un elemento de su espectro: $\sigma_P(T) \subset \sigma(T)$. Por esta razón, el conjunto $\sigma_P(T)$ se le llama *espectro puntual* del operador T. En el caso de ser X de dimensión finita, sabemos que un operador $\lambda Id - T : X \to X$ es no invertible si y sólo si su núcleo es no trivial, es decir si y sólo si existen vectores no nulos $e \in X$ tales que $(\lambda Id - T)(e) = 0_X$, o sea tales que $T(e) = \lambda e$. En resumen, cuando X es de dimensión finita, $\sigma(T) = \sigma_P(T)$.

Si X es de dimensión infinita puede haber operadores uno-a-uno que no sean sobre, y por tanto no tengan aplicación inversa, es decir que puede haber casos en que un valor espectral no sea un valor propio. Veremos a continuación dos ejemplos:

Ejemplo 31 Sea $T : \ell_2 \to \ell_2$ el operador dado por

$$T(x_1, x_2, \ldots) := (0, x_1, 0, x_2, \ldots, 0, x_n, \ldots).$$

T es inyectivo, (y con $\|T\| = 1$), el núcleo de T es $\{0_{\ell_2}\}$. Es imposible que para algún $x \neq 0_X$, $T(x) = 0_X$. Luego 0 no es un valor propio de T. En cambio T no es invertible, luego

$$0 \in \sigma(T) \backslash \sigma_P(T).$$

Ejemplo 32 Sea T el operador $T : \ell_2 \to \ell_2$ dado por

$$T(x) = T(x_1, x_2, \ldots) = (x_2, x_3, \ldots),$$

Si $0\langle|\lambda| < 1$ y tomamos el vector (no nulo) $v := (1, \lambda, \lambda^2, \ldots) \in \ell_2$, tenemos

$$T(v) = (\lambda, \lambda^2, \ldots) = \lambda(1, \lambda, \lambda^2, \ldots) = \lambda v,$$

luego λ es un valor propio si $0 < |\lambda| < 1$. Por otra parte, $T(1, 0, \ldots) = 0_{\ell_2} = 0(1, 0, \ldots)$, es decir que 0 también es un valor propio de T.

Por otra parte, si $|\lambda| \geq 1$ y fuera $T(v) = \lambda v$ para algún $v = (v_1, v_2, \ldots) \in \ell_2$ no nulo, entonces

$$(v_2, \ldots) = \lambda(v_1, v_2, \ldots),$$

de donde

$$\left.\begin{array}{ll} v_2 & = \lambda v_1 \\ v_3 & = \lambda v_2 = \lambda^2 v_1 \\ & \ldots \\ v_{n+1} & = \lambda v_n = \cdots = \lambda^n v_1 \\ & \ldots \end{array}\right\}$$

y se tendría:

$$v = (v_1, \lambda v_1, \lambda^2 v_1, \ldots)$$

por lo que llegaríamos a la siguiente contradicción:

$$0 \neq \|v\|^2 = |v_1|^2 + |\lambda||v_1|^2 + |\lambda|^2|v_1|^2 + \ldots \geq |v_1|^2 + |v_1|^2 + \ldots = +\infty$$

Por lo tanto, si $|\lambda| \geq 1$, λ no puede ser valor propio de T.

Como

$$(T - \lambda Id)(x_1, x_2, \ldots) = (x_2 - \lambda x_1, x_3 - \lambda x_2, \ldots).$$

entonces,

$$\ker(T - \lambda Id) = \{x \in \ell_2 : x_2 = \lambda x_1, x_3 = \lambda x_2, \ldots\} = \{x_1(1, \lambda, \lambda^2, \ldots)\}$$

Si $|\lambda| \geq 1$ el vector $x_1(1, \lambda, \lambda^2, \ldots)$ no puede pertenecer a ℓ_2 a menos que $x_1 = 0$. Por tanto, si $|\lambda| \geq 1$ $\ker(T - \lambda Id) = \{0_{\ell_2}\}$, lo que implica que $T - \lambda Id$ es inyectiva.

Sin embargo, vamos a ver que $(T - Id)(\ell_2) \subsetneq \ell_2$, y por tanto $T - Id$ no es invertible.

Caso de ser $y = (y_1, y_2, \ldots) = (T - Id)(x)$,

$$\left.\begin{array}{ll} y_1 & = x_2 - \lambda x_1 \\ y_2 & = x_3 - \lambda x_2 \\ & \ldots \\ y_{n+1} & = x_n - \lambda x_{n+1} \\ & \ldots \end{array}\right\}$$

y entonces

$$\left.\begin{array}{ll} x_2 & = y_1 + \lambda x_1 \\ x_3 & = y_2 + \lambda x_2 = y_2 + \lambda y_1 + \lambda^2 x_1 \\ x_4 & = y_3 + \lambda x_3 = y_3 + \lambda y_2 + \lambda^2 y_1 + \lambda^3 x_1 \\ & \ldots \\ x_{n+1} & = y_n + \lambda y_{n-1} + \lambda^2 y_{n-2} + \ldots + \lambda^{n-2} y_1 + \lambda^{n-1} x_1 \\ & \ldots \end{array}\right\}$$

Por ejemplo, si $\lambda = 1$ e

$$y = (1, \frac{1}{2}, \ldots \frac{1}{n}, \ldots)$$

fácilmente se llega a una contradicción. Vemos pues que $T - 1Id$ no es sobre, y por tanto el operador $T - Id$ no es invertible, resultando entonces que 1 es un valor espectral de T que no es un valor propio de T.

TEOREMA 27 *Sean $x_1, \ldots, x_n$ vectores propios del operador $T \in \mathcal{L}(X; X)$ correspondientes a valores propios distintos entre sí $\lambda_1, \ldots, \lambda_n$. Entonces el conjunto $\{x_1, \ldots, x_n\}$ es linealmente independiente.*

Prueba : Supondremos, para llegar a una contradicción que $\{x_1, \ldots, x_n\}$ es linealmente dependiente. Sea x_m el primero de estos vectores que sea combinación lineal de los que le precedan. (Lo cual implica que $\{x_1, \ldots, x_{m-1}\}$ es linealmente independiente, y que $x_m = \alpha_1 x_1 + \ldots + \alpha_{m-1} x_{m-1}$). Entonces

$$\begin{aligned} 0_X = (\lambda_m Id - T)(x_m) &= \lambda_m(\alpha_1 x_1 + \ldots + \alpha_{m-1} x_{m-1}) - T(\alpha_1 x_1 + \ldots + \alpha_{m-1} x_{m-1}) \\ &= \lambda_m(\alpha_1 x_1 + \ldots + \alpha_{m-1} x_{m-1}) - \alpha_1 \lambda_1 x_1 + \ldots + \alpha_{m-1} \lambda_{m-1} x_{m-1} \\ &= (\lambda_m - \lambda_1)\alpha_1 x_1 + \ldots + (\lambda_m - \lambda_{m-1})\alpha_{m-1} x_{m-1}. \end{aligned}$$

Como $\{x_1, \ldots, x_{m-1}\}$ es linealmente independiente, ha de ser

$$0 = (\lambda_m - \lambda_1)\alpha_1 = \ldots) = (\lambda_m - \lambda_{m-1})\alpha_{m-1}.$$

y como los valores propios $\lambda_1, \ldots, \lambda_n$ son distintos dos a dos, $\alpha_1 = \ldots = \alpha_{m-1} = 0$, lo que implica que $x_m = \alpha_1 x_1 + \ldots + \alpha_{m-1} x_{m-1} = 0_X$, una contradicción.

$\square$

Teorema 28 *Si $T \in \mathcal{B}(X)$ entonces el espectro $\sigma(T)$ es un conjunto compacto contenido en el disco cerrado de $\mathbb{C}$,*

$$\{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|\}.$$

Prueba : Si $|\lambda| > \|T\|$ entonces $\|\lambda^{-1}T\| < 1$. Vimos que si un operador K tiene norma estrictamente menor que 1 entonces $Id - K$ es invertible. Luego el operador

$$Id - \lambda^{-1}T = -\lambda^{-1}(T - \lambda Id)$$

es invertible también, y entonces $\lambda \in \rho(T)$, es decir λ no es valor espectral de T.

Por otra parte vamos a ver que el complementario de $sp(T)$ es un abierto. Si es vacío nada hay que probar. Si no es vacío y $\rho \in \mathbb{C}$ es un valor resolvente de T tomemos $s \in \mathbb{C}$ con $|\rho - s| < \frac{1}{\|(T-\rho I)^{-1}\|}$.

$$\|(T - \rho I) - (T - sI)\| = |\rho - s| < \frac{1}{\|(T - \rho I)^{-1}\|}.$$

Cuando vimos que $\mathcal{G}$ el conjunto de los operadores invertibles de $\mathcal{B}(X)$ es un abierto, también vimos que si $A \in \mathcal{G}$ y $B \in \mathcal{B}(X)$ con $\|B - A\| < \frac{1}{\|A^{-1}\|}$ entonces B es invertible.

Aplicando este resultado, $T - sI$ es invertible, y entonces s es también un valor resolvente de T. En definitiva, el conjunto de los valores resolventes de T es abierto, y $sp(T)$ es un cerrado contenido en la bola cerrada de $\mathbb{C}$ de centro el origen y radio $\|T\|$, luego es un compacto.

□

En principio el espectro de un operador podría ser vacío, y esto puede ocurrir[6]en un espacio de Banach real:

Ejemplo 33 *Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ dado por*

$$T(x, y) := (-y, x).$$

En este caso

$$(T - \lambda Id)(x, y) = (-y, x) - (\lambda x, \lambda y) = (-\lambda x - y, x - \lambda y)$$

tiene determinante

$$\begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = 1 + \lambda^2$$

distinto de cero para todo $\lambda \in \mathbb{R}$, luego ningún número real λ hace no invertible al operador $T - \lambda Id$. En otras palabras, $\sigma(T) = \emptyset$.

[6]Solamente.

4.3. Operadores compactos

Sean X, Y espacios normados sobre el mismo cuerpo. Recordemos la siguiente propiedad de los elementos de $\mathcal{B}(X;Y)$.

PROPOSICIÓN 23 *Si $T : X \to Y$ es una aplicación lineal y continua, y $A \subset X$ un subconjunto acotado no vacío de X, entonces el conjunto $T(A)$ es acotado. Por tanto, una aplicación lineal $T : X \to Y$ es continua si y sólo si transforma conjuntos acotados en conjuntos acotados.*

Prueba : Supongamos que $T \in \mathcal{B}(X;Y)$ y que A es un conjunto acotado en X. Si $T(A)$ es no acotado existiría una sucesión (y_n) en $T(A)$ con $\|y_n\| \geq n$ $(n = 1, 2, \ldots)$. Pero entonces existe (x_n) en A con $T(x_n) = y_n$. Como A es acotado existe una constante M con $\|x_n\| \leq M$ para todo entero positivo n,, y se da la siguiente contradicción: Para todo $n \in \mathbb{N}$:

$$n \leq \|y_n\| = \|T(x_n)\| \leq \|T\|\|x_n\| \leq \|T\|M$$

Así pues si $A \subset X$ es acotado y T lineal y continua $T(A)$ siempre es acotado.

Recíprocamente, si $T : X \to Y$ es lineal y transforma conjuntos acotados en conjuntos acotados, en particular T transforma la bola unidad de X en un conjunto acotado de Y, es decir, T es acotada en la bola unidad de X, y por tanto continua.

$\square$

DEFINICIÓN 23 *Una aplicación lineal $T : X \to Y$ se dice que es un* operador compacto *si transforma conjuntos acotados en conjuntos relativamente compactos.*

Un operador lineal compacto es siempre una aplicación lineal continua (es decir es, propiamente, lo que muchas veces se llama operador) porque la bola cerrada unidad B_X es siempre un conjunto acotado y entonces $T(B_X)$ es un conjunto relativamente compacto y por consiguiente acotado.

PROPOSICIÓN 24 *Caracterización sucesional*

Una aplicación lineal y continua $T : X \to Y$ es compacta si y sólo si de toda sucesión acotada (x_m) en X puede extraerse una subsucesión $\big(x_{m_k}\big)_{k=1}^{\infty}$ tal que $\big(T(x_{m_k})\big)_{k=1}^{\infty}$ es convergente.

Prueba : Si T es compacto y (x_m) es acotada, el conjunto $B := \{T(x_m) : m = 1, 2, ...\}$ es relativamente compacto. Entonces, como $\big(T(x_m)\big)$ es una sucesión en el compacto $C := cl(B)$, admitirá una subsucesión $\big(T(x_{m_k})\big)$ convergente (a un punto de C).

Recíprocamente, si se cumple la condición y $A \subset X$ es acotado, dada cualquier sucesión (y_n) en $T(A)$, existe $x_n \in A$ $(n = 1, 2, ...)$ con $y_n = T(x_n)$. Como A es acotado, (x_n) es acotada, luego $\big(T(x_n)\big)$, o sea (y_n) admitirá una subsucesión $\big(T(x_{n_k})\big) \equiv (y_k)$ convergente. Luego $T(A)$ es relativamente compacto.

$\square$

Ejemplo 34 Sea X un espacio normado. El operador identidad $Id : X \to X$ es compacto si y sólo si $\dim(X) < \infty$, porque la bola cerrada unidad B_X es siempre un conjunto acotado, luego si Id es compacto

$$Id(B_X) = B_X$$

será relativamente compacto. Esto en espacios de dimensión infinita no puede ocurrir (por el teorema de Riesz), luego si $\dim(X)$ es no finita $Id : X \to X$ nunca es un operador compacto.

En cambio si X es de dimensión finita y A es acotado, entonces $Id(A) = A$ es relativamante compacto, pues por ser A acotado existe una bola cerrada (y por tanto compacta) B con $A \subset B$, luego $cl(A) \subset B$ es compacta pues es un conjunto cerrado contenido en el compacto B.

Si el operador (lineal y continuo) T es no compacto, existirá $A_0 \subset X$ que es acotado y con $cl(T(A_0) =: M$ no compacto. Como acabamos de ver que $T(A_0)$ es siempre acotado (y entonces también lo es $cl(T(A_0))$, sucederá que M es un cerrado acotado de Y que no es compacto. Si el espacio Y es de dimensión finita, esto es imposible. El siguiente resultado mejora un poco estas observaciones.

PROPOSICIÓN 25 *Si $T \in \mathcal{B}(X,Y)$ es de rango finito-dimensional, es decir si el subespacio vectorial $T(X)$ es de dimensión finita, entonces T es un operador compacto.*

Prueba : Si A es cualquier conjunto acotado de X entonces $T(A)$ es acotado. Como $T(X)$ es un subespacio vectorial de dimensión finita de Y, entonces es cerrado.

$$T(X) = cl(T(X)) \supset cl(T(A)).$$

El conjunto $cl(T(A))$ es cerrado y acotado[7] contenido en $T(X)$ que es de dimensión finita. Luego $cl(T(A))$ es compacto, es decir $T(A)$ es relativamente compacto.

□

Ejemplo 35 *Sea $K : \mathcal{C}([a,b]) \to \mathcal{C}([a,b])$ un operador de la forma*

$$Kx(s) := \int_a^b [\phi_1(s)\psi_1(t) + \phi_2(s)\psi_2(t)]x(t)dt$$

Toda función $Kx \in K\big(\mathcal{C}([a,b])\big)$ puede escribirse

$$Kx = \big(\int_a^b \psi_1(t)x(t)dt\big)\phi_1 + \big(\int_a^b \psi_1(t)x(t)dt\big)\phi_2$$

luego $\dim(K(\mathcal{C}([a,b]))) \leq 2$. Por tanto K es compacto, según la proposición anterior.

[7]Si M es acotado en un espacio métrico, está contenido en una bola cerrada, digamos B. Entonces $cl(M) \subset cl(B) = B$, luego $cl(M)$ también es acotado.

Ejemplo 36 *Sea $T : \mathcal{C}([a,b]) \to \mathcal{C}([a,b])$ el operador integral $x \mapsto Tx$ definido por*

$$Tx(s) := \int_a^b k(s,t)x(t)dt$$

donde $k : [a,b] \times [a,b] \to \mathbb{R}$ es una función continua.

En primer lugar podemos ver que la función Tx así definida es continua en $[a,b]$: Si $\|x\|_\infty = 0$, entonces Tx es la función constante 0 en $[a,b]$ que, desde luego es continua. Supongamos pues que $\|x\|_\infty \neq 0$. Si damos $\varepsilon > 0$, como k es uniformemente continua en $[a,b] \times [a,b]$, existe un $\delta > 0$ de modo que para cualesquiera $(\alpha,\beta),(\lambda,\mu) \in [a,b] \times [a,b]$ con $\|(\alpha,\beta) - (\lambda,\mu)\| < \delta$ se tiene que

$$|k(\alpha,\beta) - k(\lambda,\mu)| < \frac{\varepsilon}{2\|x\|_\infty(b-a)}.$$

Si tomamos $s_1, s_2 \in [a,b]$ con $|s_1 - s_2| < \delta$, para cualquier $t \in [a,b]$ se tiene que

$$\|(s_1,t) - (s_2,t)\| < \delta$$

y entonces

$$|k(s_1,t) - k(s_2,t)| < \frac{\varepsilon}{2\|x\|_\infty(b-a)}.$$

(Observemos que $\delta > 0$ depende de k, y también de (la norma de) la función x elegida en $\mathcal{C}([a,b])$).

Teniendo en cuenta esta desigualdad, si $s_1, s_2 \in [a,b]$ con $|s_1 - s_2| < \delta$

$$\begin{aligned}
|Tx(s_1) - Tx(s_2)| &= \left|\int_a^b k(s_1,t)x(t)dt - \int_a^b k(s_2,t)x(t)dt\right| \\
&= \left|\int_a^b (k(s_1,t) - k(s_2,t))x(t)dt\right| \\
&\leq \int_a^b |(k(s_1,t) - k(s_2,t))||x(t)|dt \\
&\leq \|x\|_\infty \int_a^b |k(s_1,t) - k(s_2,t)|dt \\
&\leq \|x\|_\infty \int_a^b \frac{\varepsilon}{2\|x\|_\infty(b-a)}dt \\
&= \frac{\varepsilon}{2} < \varepsilon
\end{aligned}$$

Con esto hemos probado que la función Tx es (uniformemente) continua en el intervalo $[a,b]$.

Vamos a ver que el operador T es compacto (cuando se considera en $\mathcal{C}([a,b])$ la norma del supremo).

Sea $\mathcal{S}$ un conjunto acotado no vacío de $\mathcal{C}([a,b])$. Existe entonces $M < \infty$ tal que $\|\varphi\| \leq M$ para cualquier $\varphi \in S$. Además, como k es continua en el conjunto compacto $[a,b] \times [a,b]$ existe $K < \infty$ cota superior de $|k|$ en $[a,b] \times [a,b]$. Entonces,

$$|T\varphi(s)| \leq \int_a^b |k(s,t)||\varphi(t)|dt \leq \|\varphi\|_\infty K(b-a) \leq MK(b-a)$$

luego $\|T\varphi\|_\infty \leq MK(b-a)$, y esta cota no depende de la función φ elegida en $\mathcal{S}$. En otras palabras, podemos decir que el conjunto de funciones $T(\mathcal{S})$ es equiacotado (o también uniformemente acotado).

Por otra parte, si $s_1, s_2 \in [a,b]$, y $x \in \mathcal{S}$,

$$\begin{aligned} |Tx(s_1) - Tx(s_2)| &= \left|\int_a^b k(s_1,t)x(t)dt - \int_a^b k(s_2,t)x(t)dt\right| \\ &\leq \|x\|_\infty \int_a^b |(k(s_1,t) - k(s_2,t))|dt \\ &\leq M \int_a^b |(k(s_1,t) - k(s_2,t))|dt \end{aligned}$$

Si damos $\varepsilon > 0$, utilizando de nuevo que k es uniformemente continua en $[a,b] \times [a,b]$, sabemos que existe un $\tilde{\delta} > 0$ de modo que para cualesquiera $(\alpha,\beta), (\lambda,\mu) \in [a,b] \times [a,b]$ con $\|(\alpha,\beta) - (\lambda,\mu)\| < \tilde{\delta}$ se tiene que

$$|k(\alpha,\beta) - k(\lambda,\mu)| < \frac{\varepsilon}{2M(b-a)}.$$

(Observemos que $\tilde{\delta}$ sólo depende de ε (y del núcleo k y de M).

En particular, si tomamos $s_1, s_2 \in [a,b]$ con $|s_1 - s_2| < \tilde{\delta}$, para cualquier $t \in [a,b]$,

$$\|(s_1,t) - (s_2,t)\| < \tilde{\delta}$$

y entonces

$$|k(s_1,t) - k(s_2,t)| < \frac{\varepsilon}{2M(b-a)}.$$

De aquí que

$$\int_a^b |k(s_1,t) - k(s_2,t)|dt < \frac{\varepsilon}{2(b-a)}(b-a) = \frac{\varepsilon}{2M}$$

y entonces,

$$|Tx(s_1) - Tx(s_2)| \leq M\frac{\varepsilon}{2M} = \frac{\varepsilon}{2} < \varepsilon.$$

Con esto hemos probado que, dado $\varepsilon > 0$ el valor de $\tilde{\delta} > 0$ obtenido para ver que la función Tx es continua en $[a,b]$, **no depende de la función x elegida en $\mathcal{S}$**. Esto suele reflejarse diciendo que el conjunto de funciones $\mathcal{S}$ es *equicontinuo*.

Usamos ahora el siguiente resultado (bien conocido):

TEOREMA 29 *(Arzelà-Ascoli). En $(\mathcal{C}([a,b]), \|.\|_\infty)$ todo conjunto (no vacío) S que sea acotado y equicontínuo es relativamente compacto.*

Como consecuencia de este teorema de Arzelà-Ascoli, el conjunto $T(\mathcal{S})$ es relativamente compacto. El operador T transforma pues conjuntos acotados en conjuntos relativamente compactos, y por lo tanto T es un operador compacto.

4.4. Propiedades de los operadores compactos

Proposición 26 *Si T_1, T_2 son operadores compactos entre los espacios normados X, Y, entonces para cualesquiera escalares α, β el operador $\alpha T_1 + \beta T_2$ es compacto. En consecuencia el conjunto $\mathcal{K}(X, Y)$ de los operadores compactos entre los espacios X, Y tiene estructura de espacio vectorial (sobre el mismo cuerpo de escalares en que están definidos X, Y).*

Prueba : Sea (x_n) una sucesión acotada en X. Como T_1 es compacto, obtenemos una subsucesión $\left(x_{n_k}\right)$ de (x_n) con $T(x_{n_k}) \to y_1 \in Y$. Pero $\left(x_{n_k}\right)$ es también una sucesión acotada en X, y por ser compacto T_2 admitirá una subsucesión $\left(x_{n_{k_j}}\right)$ con $T(x_{n_{k_j}}) \to y_2 \in Y$. Entonces $\left(x_{n_{k_j}}\right)$ es una subsucesión de (x_n) y

$$(\alpha T_1 + \beta T_2)(x_{n_{k_j}}) \to \alpha y_1 + \beta y_2,$$

lo que prueba que $\alpha T_1 + \beta T_2$ es compacto.

□

Proposición 27 *Sean X, Y, Z espacios normados sobre el mismo cuerpo. Sean $T \in \mathcal{B}(X, Y)$, $S \in \mathcal{B}(Y, Z)$. Entonces*
a) Si T es compacto, $S \circ T$ es compacto.
b) Si S es compacto, $S \circ T$ es compacto.

Prueba : Si A es acotado en X y T es compacto, el conjunto $T(A)$ será relativamente compacto en Y. Como S es continuo se tiene que $S(cl(T(A))$ es compacto en Z por ser imagen continua de un conjunto compacto. Pero como S es continua $cl(S(T(A)) = S(cl(T(A))$, $cl(S \circ T)(A)$ es compacto, lo que prueba (a).

Igualmente, sabemos que si A es acotado, por ser T lineal y continuo, $T(A)$ es acotado, y si S es opertador compacto, $(S \circ T)(A) = S(T(A))$ será relativamente compacto, lo que prueba (b).

□

Teorema 30 *$\mathcal{K}(X, Y)$ es cerrado en $\mathcal{B}(X, Y)$.*

Prueba : Sea (T_n) una sucesión en $\mathcal{K}(X, Y)$ con $\|T_n - T\| \to 0$ para cierto operador $T \in \mathcal{B}(X, Y)$. Hemos de ver que T es compacto, o sea que $T \in \mathcal{K}(X, Y)$.

Sea (x_m) una sucesión acotada en X, es decir tal que $\|x_m\| \leq C < \infty$ para $m = 1, 2, \ldots$. Aplicaremos un procedimiento diagonal: Por ser T_1 compacto, $(T_1(x_m))$ admitirá una subsucesión convergente (y por tanto de Cauchy) digamos $(T_1(x_{1,m})_{m=1}^{\infty})$. De igual modo, $(x_{1,m})$ es acotada y admitirá una subsucesión $(x_{2,m})$ tal que $(T_2(x_{2,m}))$ es sucesión de Cauchy en Y. Continuando de este modo, por inducción formamos las sucesiones $(x_{p,m})$ para $p = 1, 2, \ldots$, de manera que $(x_{p,m})$ es subsucesión de $(x_{p-1,m})$ para $p = 2, \ldots$, y además $(T_p(x_{p,m}))$ es una sucesión de Cauchy.

La sucesión diagonal $(x_{m,m}) =: (v_m)$ es subsucesión de la (x_m) y también de cada una de las $(x_{p,m})$ para $p = 1, 2, \ldots$. Por tanto $(T_p(x_{m,m}))$ es una sucesión de Cauchy pues es subsucesión de la sucesión (convergente) $(T_p(x_{p,m}))$.

Dado $\varepsilon > 0$ existe un natural p para el que $\|T_p - T\| < \frac{\varepsilon}{3C}$. Como $(T_p(x_{m,m})) \equiv (T_p(v_m))$ es sucesión de Cauchy, existe un natural q tal que si $j, k \geq q$

$$\|T_p(v_j) - T_p(v_k)\| < \frac{\varepsilon}{3}.$$

Entonces, si $j, k \geq q$,

$$\begin{aligned}
\|T(v_j) - T(v_k)\| &\leq \|T(v_j) - T_p(v_j)\| + \|T_p(v_j) - T_p(v_k)\| + \|T_p(v_k) - T(v_k)\| \\
&\leq \|T - T_p\|\|v_j\| + \|T_p(v_j) - T_p(v_k)\| + \|T_p - T\|\|v_k\| \\
&\leq \|T - T_p\|C + \|T_p(v_j) - T_p(v_k)\| + \|T_p - T\|C \\
&< \tfrac{\varepsilon}{3C}C + \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3C}C = \varepsilon
\end{aligned}$$

lo que prueba que $(T(v_m))$ es una sucesión de Cauchy en Y, y por tanto convergente. Dado que (v_m) es subsucesión de (x_m) hemos probado que T es compacto.

□

Ejemplo 37 Consideremos los operadores $T_m : \ell_2 \to \ell_2$ definidos por

$$T_m(x_1, x_2, \ldots) = (x_1, x_2, \ldots, x_m, 0, 0, ...).$$

Como para cada entero positivo m T_m es de rango finito, (T_m) es una sucesión de operadores compactos. Fijado $x = (x_1, x_2, \ldots) \in \ell_2$ está claro que $T_m(x) \to Id(x) = x$. Sin embargo, no puede ser $\|T_m - Id\| \to 0$ porque, según el teorema anterior, el operador Id habría de ser compacto, y sabemos que no lo es.

Ejemplo 38 Consideremos los operadores $T_m : \ell_2 \to \ell_2$ definidos por

$$T_m(x_1, x_2, \ldots) = \left(x_1, \frac{x_2}{2}, \ldots, \frac{x_m}{m}, 0, 0, ...\right).$$

Como en el ejemplo anterior, cada T_m es un operador de rango finito, y por tanto compacto.

Sea $T : \ell_2 \to \ell_2$ dado por

$$T(x_1, x_2, \ldots) = \left(x_1, \frac{x_2}{2}, \ldots, \frac{x_n}{n}, ...\right).$$

Es inmediato ver que T es lineal y para cada $x \in \ell_2$,

$$\|T(x)\|_2^2 = \sum_{n=1}^{\infty} |\frac{x_n}{n}|^2 \leq \sum_{n=1}^{\infty} |x_n|^2 = \|x\|_2^2$$

luego T es continuo con $\|T\| \leq 1$. Además, para cada $x \in \ell_2$

$$\begin{aligned}
\|(T - T_n)(x)\|_2^2 &= \textstyle\sum_{p=n+1}^{\infty} |\frac{x_p}{p}|^2 \\
&= \tfrac{|x_{n+1}|^2}{(n+1)^2} + \tfrac{|x_{n+2}|^2}{(n+2)^2} + \tfrac{|x_{n+3}|^2}{(n+3)^2} + \ldots \\
&\leq \tfrac{|x_{n+1}|^2}{(n+1)^2} + \tfrac{|x_{n+2}|^2}{(n+1)^2} + \tfrac{|x_{n+3}|^2}{(n+1)^2} + \ldots \\
&= \tfrac{1}{(n+1)^2}(|x_{n+1}|^2 + |x_{n+2}|^2 + |x_{n+3}|^2 + \ldots) \\
&\leq \tfrac{1}{(n+1)^2}\|x\|_2^2.
\end{aligned}$$

Por definición de norma de un operador se sigue que

$$\|T - T_n\| \leq \frac{1}{n+1},$$

luego $\|T - T_n\| \to 0$, por lo que el operador T será compacto según el teorema anterior.

PROPOSICIÓN 28 *Sean X, Y espacios de Banach sobre el mismo cuerpo. Sea T un operador compacto entre X e Y. Entonces, dado $\varepsilon > 0$ existe un subespacio M, (dependiendo de ε), tal que*

a) M es de dimensión finita,
b) $M \subset T(X)$,
c) Para cada $x \in X$,

$$d(T(x), M) = \inf\{\|T(x) - y\| : y \in M\} \leq \varepsilon\|x\|.$$

Prueba : Sea $\{B(Tx, \frac{\varepsilon}{2}) : x \in B_X\}$. Esta colección de bolas abiertas de radio $\frac{\varepsilon}{2}$ recubre al conjunto $T(B_X)$, que tiene clausura compacta, por ser T un operador compacto. Como

$$cl[T(B_X)] \subset cl\left[\bigcup_{x\in B_X} B(T(x), \frac{\varepsilon}{2})\right] \subset \bigcup_{x\in B_X} cl[B(T(x), \frac{\varepsilon}{2})] \subset \bigcup_{x\in B_X} B(T(x), \varepsilon)$$

la colección de bolas abiertas $\{B(T(x), \varepsilon) : x \in B_X\}$ es un recubrimiento del compacto $cl(T(B_X))$, que admitirá un subrecubrimiento finito, $\{B(T(x_1), \varepsilon), \ldots, B(T(x_n), \varepsilon)\}$, de $cl(T(B_X))$ y a la fuerza de $T(B_X)$.

Sea M el subespacio que generan los vectores $\{T(x_1), \ldots, T(x_n)\}$. Obviamente $M \subset T(X)$ y M es de dimensión finita.

Para todo $x \in B_X$, Tx pertenece a alguna de estas bolas, digamos $B(T(x_i), \varepsilon)$, luego

$$d(T(x), M) \leq d(T(x), T(x_i)) < \varepsilon.$$

Si $x \in X$ con $x \neq 0_X$ tenemos que $\frac{1}{\|x\|}x \in B_X$, por lo que

$$d\left(T\left(\frac{1}{\|x\|}x\right), M\right) < \varepsilon.$$

Si $m \in M$,

$$d(T(x), M) \leq d(T(x) - \|x\|m) = \|x\| \left\|T\left(\frac{1}{\|x\|}x\right) - m\right\|.$$

Tomando ínfimos al variar m en M,

$$\begin{aligned} d(T(x), M) &\leq \inf\{\|x\| \left\|T\left(\tfrac{1}{\|x\|}x\right) - m\right\| : m \in M\} \\ &= \|x\| \inf\{\left\|T\left(\tfrac{1}{\|x\|}x\right) - m\right\| : m \in M\} \\ &= \|x\| d\left(T\left(\tfrac{1}{\|x\|}x\right), M\right) \\ &< \|x\|\varepsilon. \end{aligned}$$

Para $x = 0_X$ trivialmente $0 = d(T(x), M) = \inf\{\|T(x) - y\| : y \in M\} \leq \varepsilon\|x\| = 0$, luego para todo $x \in X$ se cumple (c).

$\square$

4.5. Espectro de los operadores compactos. Generalidades.

Empezamos recordando un resultado que en realidad ya se vio en la primera parte de estas notas:

LEMA 5 *(Riesz). Sean $M \subset Y$ con $M \neq Y$, ambos subespacios del espacio normado X. Si M es cerrado, dado cualquier $\theta \in (0,1)$ existe $y \in Y$ con $\|y\| = 1$ y con $\|m - y\| \geq \theta$ para todo $m \in M$.*

Prueba : Dado $v \in Y \backslash M$, $d(v, M) > 0$ porque M es cerrado. Por definición de $d(v, M)$ existe $m_\theta \in M$ tal que

$$0 < d(v, M) \leq \|v - m_\theta\| < \frac{d(v, M)}{\theta}.$$

Sea $y = \frac{1}{\|v - m_\theta\|}(v - m_\theta) \in Y$. Entonces, para cualquier $m \in M$,

$$\begin{aligned} \|m - y\| &= \tfrac{1}{\|v-m_\theta\|}\|m\|v - m_\theta\| - (v - m_\theta)\| \\ &= \tfrac{1}{\|v-m_\theta\|}\|m\|v - m_\theta\| + m_\theta - v\| \\ &\geq \tfrac{1}{\|v-m_\theta\|} d(v, M) \\ &> \tfrac{1}{\frac{d(v,M)}{\theta}} d(v, M) \\ &= \theta. \end{aligned}$$

$\square$

TEOREMA 31 *Sea $T \in \mathcal{K}(X)$. Entonces $\sigma_P(T)$ es a lo sumo numerable, y su único posible punto de acumulación es el 0.*

Prueba : Veremos que para cada número real $r > 0$ el conjunto

$$\{\lambda \in \sigma_P(T) : |\lambda| \geq r\}$$

es finito (si no es vacío). A partir de aquí es inmediato que

$$\sigma_P(T) \backslash \{0\} = \bigcup_n \{\lambda \in \sigma_P(T) : |\lambda| \geq \frac{1}{n}\}$$

será a lo sumo numerable.

Por reducción al absurdo, supongamos que para algún r_0 el conjunto $\{\lambda \in \sigma_P(T) : |\lambda| \geq r_0\}$ es no finito. Podemos encontrar entonces una sucesión (λ_n) de valores propios de T, distintos dos a dos y con $|\lambda_n| \geq r_0$ para todo n.

Sean (x_n) los correspondiente vectores propios asociados (es decir, con $T(x_n) = \lambda_n x_n$). El conjunto $\{x_1, \ldots, x_n, \ldots\}$ es linealmente independiente.

Sea M_n el subespacio de X generado por $\{x_1, \ldots, x_n\}$. Observar que $T(M_n) \subset M_n$ para cada $n \in \mathbb{N}$. Observar también que si $x \in M_n$ puede escribirse de manera única $x = \alpha_1 x_1 + \ldots + \alpha_n x_n$, y entonces

$$(\lambda_n Id - T)(x) = \lambda_n x - (\alpha_1 \lambda_1 x_1 + \ldots + \alpha_n \lambda_n x_n)$$

$$= \alpha_1(\lambda_n - \lambda_1)x_1 + \ldots + \alpha_{n-1}(\lambda_n - \lambda_{n-1})x_{n-1} \in M_{n-1}.$$

Los subespacios M_n son cerrados por ser de dimensión finita. Por el Lema de Riesz, dado $\theta = \frac{1}{2}$ existe $y_n \in M_n$ con $\|y_n\| = 1$ y tal que para todo $x \in M_{n-1}$, $\|y_n - x\| \geq \frac{1}{2}$. Sean m, n con $m < n$. Pongamos

$$T(y_n) - T(y_m) = \lambda_n y_n - (\lambda_n y_n + T(y_m) - T(y_n)) =: \lambda_n y_n - x'$$

con $x' := \lambda_n y_n - T(y_n) + T(y_m) = (\lambda_n Id - T)(y_n) + T(y_m)$.

Como $y_m \in M_m \subset M_{n-1}$, $T(y_m) \in T(M_{n-1}) \subset M_{n-1}$. Por otra parte,

$$(\lambda_n Id - T)(y_n) \in (\lambda_n Id - T)(M_n) \subset M_{n-1}$$

según acabamos de ver. Luego x' es suma de dos vectores de M_{n-1} y por tanto pertenece a M_{n-1}. Tal como se ha elegido la sucesión (y_n), y dado que $\frac{1}{\lambda_n}x' \in M_{n-1}$

$$\begin{aligned} \|T(y_n) - T(y_m)\| &= \|\lambda_n y_n - x'\| \\ &= |\lambda_n| \|y_n - \tfrac{1}{\lambda_n}x'\| \\ &\geq |\lambda_n| \tfrac{1}{2} \\ &\geq \tfrac{r_0}{2}. \end{aligned}$$

Siendo m, n arbitrarios (con $m < n$) hemos probado que la sucesión $(T(y_n))$ no admite subsucesiones de Cauchy. Pero esto es una contradicción porque (y_n) es una sucesión acotada en X y T es un operador compacto, con lo cual $(T(y_n))$ ha de admitir una subsucesión convergente (y por tanto de Cauchy).

□

Teorema 32 *Sean X un espacio normado, $T \in \mathcal{B}(X)$ un operador compacto y M un subespacio cerrado de X tal que*

$$M \cap \ker(Id - T) = \{0_X\}$$

Entonces la restricción a M del operador $Id - T$, es decir $(Id - T)\big|_M$, tiene inversa continua y $(Id - T)(M)$ es cerrado en X.

Prueba : Desde luego si $M \cap \ker(Id - T) = \{0_X\}$ la aplicación lineal $(Id - T)\big|_M$ es inyectiva y entonces está definida como aplicación lineal su inversa $S := \left((Id - T)\big|_M\right)^{-1}$: $(Id - T)(M) \to M$.

Supongamos, para llegar a una contradicción, que S no es continua. Entonces para todo natural n no se cumple la condición $[\forall y \in (Id - T)(M),\ \|S(y)\| \leq n\|y\|]$, es decir que existe $y_n \in (Id - T)(M)$ tal que $\|S(y_n)\| > n\|y_n\|$ para cada entero positivo n. En otras palabras, existe $x_n \in M$ tal que $\|S(Id - T)(x_n)\| > n\|(Id - T)(x_n)\|$. Como $\|x_n\| = \|S(Id - T)(x_n)\|$, de la desigualdad anterior obtenemos $\|x_n\| > n\|(Id - T)(x_n)\|$, o sea,

$$\frac{1}{n} > \|(Id - T)(\frac{1}{\|x_n\|}x_n)\|$$

para cada entero positivo n. Tomando $v_n := \frac{1}{\|x_n\|} x_n$, tenemos que (v_n) es una sucesión de vectores de la esfera unidad de M con $\|(Id-T)(v_n)\| < \frac{1}{n}$ para cada natural n.

Como el operador T es compacto la sucesión $(T(v_n))$ admite una subsucesión $\big(T(v_{n_k})\big)$ convergente a cierto vector, digamos $y \in X$. Pero entonces como para cada $k \in \mathbb{N}$,

$$\begin{aligned} \|v_{n_k} - y\| &\leq \|v_{n_k} - T(v_{n_k})\| + \|T(v_{n_k}) - y\| \\ &= \|(Id-T)(v_{n_k})\| + \|T(v_{n_k}) - y\| \\ &< \|T(v_{n_k}) - y\| + \tfrac{1}{n_k} \end{aligned}$$

la sucesión (v_{n_k}) converge a y (lo que en particular implica que $\|y\| = 1$) y como M es cerrado que $y \in M$. Como T es continuo, $T(v_{n_k}) \to T(y)$, y también $T(v_{n_k}) \to y$, luego $y = T(y)$, es decir $(Id-T)(y) = 0_X$. Luego y es un vector de M que está en el núcleo de $Id-T$. Pero la restricción de $Id-T$ a M es inyectiva, es decir su núcleo se reduce a $\{0_X\}$, luego ha de ser $y = 0_X$. Pero esto es una contradicción, pues $\|y\| = 1$.

Veamos que el subespacio $(Id-T)(M)$ es cerrado en X. Sea (y_n) sea una sucesión en $(Id-T)(M)$ convergente a cierto vector y de X. Para cada natural n sea con $x_n \in M$ con $y_n = (Id-T)(x_n)$. Para ver que $(Id-T)(M)$ es cerrado nos bastará ver que $y \in (Id-T)(M)$. Al ser $x_n = S(y_n)$ y ser S continua según hemos demostrado ya, (x_n) convergerá hacia $S(y)$. Entonces la sucesión (x_n) es acotada (por ser convergente) y como el operador T es compacto la sucesión $(T(x_n))$ admitirá una subsucesión $(T(x_{n_k}))$ convergente a cierto vector $z \in X$. Entonces, como

$$x_{n_k} = T(x_{n_k}) + (Id-T)(x_{n_k})$$

tendremos que (x_{n_k}) converge a $z+y$ que pertenecerá a M porque este conjunto es cerrado y los vectores x_{n_k} están en él. Por continuidad de $Id-T$,

$$y = \lim_k y_{n_k} = \lim_k (Id-T)(x_{n_k}) = (Id-T)(z+y) \in (Id-T)(M),$$

lo que completa la prueba.

□

Tomando $M = X$ en el teorema anterior, obtenemos:

COROLARIO 12 *Sean X un espacio normado, y $T \in \mathcal{B}(X)$ un operador compacto con*

$$\ker(Id-T) = \{0_X\}$$

Entonces $Y := (Id-T)(X)$ es cerrado en X y $(Id-T)$ tiene inversa continua $(Id-T)^{-1} : Y \to X$.

TEOREMA 33 *Sea $(X, \|.\|)$ un espacio normado. Sea $T \in \mathcal{B}(X)$ un operador compacto. Entonces $\ker(Id-T)$ es de dimensión finita.*

Prueba:

Si $M := \ker(Id-T) = \{0_X\}$ nada hay que probar. Desde luego, por ser T continua, M es cerrado. Si $x \in \ker(Id-T)$ es un vector no nulo, entonces $T(x) = x$. Luego la restricción de $Id-T$ al núcleo $\ker(Id-T)$ es la aplicación identidad sobre M. Como T es un operador compacto, si $B_M = M \cap B_X$ es la bola cerrada unidad de M (conjunto acotado), tenemos que $T(B_M)$ es relativamente compacto, es decir $T(B_M) = Id(B_M) = B_M$ es relativamente compacto en $(X, \|.\|)$. Luego la clausura (respecto de X) del conjunto B_M es un conjunto compacto (en X). (Toda sucesión (x_n) de elementos de B_M tendrá una subsucesión (x_{n_k}) convergente a un punto x de B_M, es decir con $\|x_{n_k} - x\|_X \to 0$). Pero la convergencia en $(X, \|.\|_X)$ implica la convergencia en M, es decir $\|x_{n_k} - x\|_M = \|x_{n_k} - x\|_X \to 0$, luego B_M es un conjunto compacto en el espacio normado $(M, \|.\|_M)$. Por tanto según el teorema de F. Riesz, M es de dimensión finita.

□

Corolario 13 *Sea X un espacio normado. Si $T \in \mathcal{B}(X)$ es compacto, todo valor propio no nulo de T tiene orden de multiplicidad finita.*

Prueba:

Si λ es un valor propio no nulo de T entonces existe $e \in X$ no nulo tal que $T(e) = \lambda e$, o sea $(T - \lambda Id)(e) = 0_X$. De aquí se sigue que $(\lambda^{-1}T - Id)(e) = 0_X$ o sea que $e \in \ker(Id - \lambda^{-1}T)$. Este argumento es válido para cualquier vector propio asociado al valor propio λ, luego el subespacio $V_\lambda(T)$ de los vectores propios asociados al valor propio λ está contenido en el núcleo $\ker(Id - \lambda^{-1}T)$. Pero $\lambda^{-1}T$ es un operador compacto, y por el teorema anterior este núcleo $\ker(Id - \lambda^{-1}T)$ es de dimensión finita. Así pues $V_\lambda(T)$ es de dimensión finita.

□

Teorema 34 *Sea X un espacio normado. Si $T \in \mathcal{B}(X)$ es compacto, entonces para cada $\lambda \neq 0$, $(\lambda Id - T)(X)$ es un subespacio cerrado de X.*

Prueba : Por comodidad escribiremos $T_\lambda := \lambda Id - T$. Suponemos que $Y := T_\lambda(X)$ es no cerrado. Existe una sucesión (x_n) en X tal que $y_n := T_\lambda(x_n) \to y \in X \backslash Y$. Como Y es subespacio vectorial de X, de ser $y \notin Y$ se sigue que $y \neq 0_X$, y entonces $y_n = T_\lambda(x_n) \neq 0_X$ para n suficientemente grande, o sea que $x_n \notin N_\lambda := \ker(T_\lambda)$ para n suficientemente grande. Como N_λ es cerrado y $x_n \notin N_\lambda$ se tiene que para $n \geq n_0 \in \mathbb{N}$

$$d_n := d(x_n, N_\lambda) > 0.$$

Por definición de ínfimo, existe $z_n \in N_\lambda$ tal que

$$0 < d_n \leq a_n := \|x_n - z_n\| < 2d_n.$$

Afirmamos que $a_n \to \infty$.

Caso contrario, $(x_n - z_n)$ tiene una subsucesión acotada, y como T es compacto a su vez esa subsucesión admite otra subsucesión, $(x_{n_k} - z_{n_k})$ tal que $(T(x_{n_k}))$ es convergente.

Pero como $Id = \lambda^{-1}(T + (\lambda Id - T)) = \lambda^{-1}(T + T_\lambda)$,

$$\begin{aligned} x_{n_k} - z_{n_k} &= \lambda^{-1}(T(x_{n_k} - z_{n_k}) + T_\lambda(x_{n_k} - z_{n_k})) \\ &= \lambda^{-1}(T(x_{n_k} - z_{n_k}) + T_\lambda(x_{n_k})) \\ &= \lambda^{-1}(T(x_{n_k} - z_{n_k}) + y_{n_k}) \end{aligned}$$

y las sucesiónes $(T(x_{n_k} - z_{n_k}))$, (y_{n_k}) son convergentes, entonces $(x_{n_k} - z_{n_k})$ es también convergente, es decir que existe $v \in X$ con $x_{n_k} - z_{n_k} \to v$.

Pero en tal caso $T_\lambda(x_{n_k} - z_{n_k}) \to T_\lambda(v)$, y como $z_{n_k} \in N_\lambda$, en realidad vemos que $T_\lambda(x_{n_k}) \to T_\lambda(v)$ y también $T_\lambda(x_{n_k}) \to y$, luego $y = T_\lambda(v) \in T_\lambda(X) =: Y$, lo cual es absurdo porque estamos suponiendo $y \notin Y$.

En definitiva, $a_n \to \infty$. Sea ahora $w_n := \frac{1}{a_n}(x_n - z_n)$. Como la sucesión $(T_\lambda(x_n))$ es convergente y por tanto acotada, y $a_n \to \infty$,

$$T_\lambda(w_n) = \frac{1}{a_n}(T_\lambda(x_n) - T_\lambda(z_n)) = \frac{1}{a_n}(T_\lambda(x_n)) \longrightarrow 0_X.$$

Además

$$w_n = \frac{1}{\lambda}(T + T_\lambda)(w_n) = \frac{1}{\lambda}(T(w_n) + T_\lambda(w_n))$$

y la sucesión $(T(w_n))$ admite una subsucesión convergente por ser T compacto, entonces (w_n) admitirá una subsucesión (w_{n_j}) con $(T(w_{n_j}))$ convergente a cierto $u \in X$. Pero, para $j = 1, 2, \ldots$,

$$w_{n_j} = \frac{1}{\lambda}(T + T_\lambda)(w_{n_j}) = \frac{1}{\lambda}(T(w_{n_j}) + T_\lambda(w_{n_j})) \longrightarrow \frac{1}{\lambda}(u + 0_X)$$

luego (w_{n_j}) converge a cierto $w \in X$, y entonces tomando límites en la primera igualdad anterior

$$w = \frac{1}{\lambda}(T(w) + 0_X)$$

es decir $(\lambda Id - T)(w) = 0_X$, o, en otras palabras $w \in N_\lambda$.

Pero en es caso para todo entero positivo,

$$u_n := z_n + a_n w \in N_\lambda$$

de donde

$$d_n \leq \|x_n - u_n\| = a_n w_n - a_n w\| = a_n\|w_n - w\| < 2d_n\|w_n - w\|,$$

y dividiendo por $d_n > 0$ obtenemos que para todo entero positivo es $\frac{1}{2} \leq \|w_n - w\|$ una contradiccioón pues $w_n \to w$.

□

Observación Si X, Y son espacios vectoriales sobre el mismo cuerpo, y $J : X \to Y$ es lineal e inyectiva, entonces de ser X de dimensión no finita, se sigue que Y es de dimensión no finita.

En efecto, basta tener en cuenta que si $v^1, \ldots, v^k$ son vectores linealmente independientes de X, entonces $J(v^1), \ldots, J(v^k)$ son vectores linealmente independientes en Y ya que de ser $\beta_1 J(v^1) + \ldots + \beta_k J(v^k) = 0_Y$ se sigue que $J(\beta_1 v^1 + \ldots + \beta_k v^k) = 0_Y$ o sea que $\beta_1 v^1 + \ldots + \beta_k v^k \in \ker(J) = \{0_X\}$. Por tanto $\beta_1 v^1 + \ldots + \beta_k v^k = 0_X$ y se obtiene que $0 = \beta_1 = \ldots = \beta_k$. Por tanto vemos que, ciertamente, $J(v^1), \ldots, J(v^k)$ son vectores linealmente independientes en Y.

TEOREMA 35 *Sean X un espacio de Banach y $T \in \mathcal{B}(X)$ un operador compacto. Si $Id - T$ es inyectiva, entonces $(Id - T)(X) = X$.*

Prueba.

Si X es de dimensión finita el resultado se cumple, como es conocido de Algebra elemental. Supondremos pues que X es de dimensión infinita. Si $(Id - T)(X) \neq X$, sea $X_1 := (Id-T)(X)$. Estamos suponiendo, para llegar a una contradicción que $X_1 \neq X$. Por ser T compacto y por el teorema anterior, X_1 es cerrado. Además, siendo $Id-T$ inyectiva X_1 ha de ser de dimensión infinita, por la observación anterior. Sea T_1 la restricción del operador T al subespacio vectorial X_1.

Observemos que $T(X_1) \subset X_1$. En efecto, si $x \in X_1$ existe $z \in X$ con $x = (Id-T)(z) = z - T(z)$. Pongamos $y = T(z)$.

$$y - T(y) = T(z) - T^2(z) = T(z - T(z)) = T(x)$$

luego $T_1(x) = T(x) \in X_1$, es decir que X_1 es T_1-invariante.

Como X_1 es cerrado y T compacto, T_1 es un operador compacto de X_1.

Sea ahora $X_2 := (Id - T_1)(X_1) = (Id - T_1)(Id - T)(X) = (Id - T)^2(X)$.

Razonando como antes vemos que T_2 es un operador compacto del espacio de Banach X_2. Afirmamos que $X_2 \neq X_1$. Caso contrario, para cada $x \in X$, $(Id - T)(x) \in X_1 = X_2$ luego existe $x' \in X_1$ tal que

$$(Id - T)(x) = (Id - T_1)(x') = x' - T_1(x') = x' - T(x') = (Id - T)(x')$$

pero como, por hipótesis $Id - T$ es inyectivo concluimos que $x = x'$ es decir que $x \in X_1$. Como el razonamiento vale para cada $x \in X$ tenemos que $X = X_1$, lo cual es absurdo pues estamos suponiendo lo contrario.

Además ha de ser X_2 de dimensión infinita.

Razonando por recurrencia, supuesto que hemos construido los subespacios cerrados de dimensión infinita $X_1, \ldots, X_n$ con

$$X \supsetneq X_1 \supsetneq X_2 \supsetneq \ldots \supsetneq X_n$$

y con $X_n := (Id - T)^n(X)$, y los operadores $T_n : X_n \to X_n$ con T_n la restrucción de T a X_n, pondremos $X_{n+1} = T_n(X_n)$. Siendo T_n compacto, por el teorema anterior X_{n+1} es cerrado, y caso de ser $X_{n+1} = X_n$ obtendríamos $X_n = X_{n-1}$, una contradicción. Así pues, por inducción hemos formado una sucesión de subespacios cerrados de dimensión no finita (X_n), con $X_n \supsetneq X_{n+1}$ para todo entero positivo n. Aplicamos ahora el lema de Riesz [8] y obtenemos $x_n \in X_n$ con $\|x_n\| = 1$ y con $d(x_n, X_{n+1}) \geq \frac{1}{2}$.

[8] Sea X un espacio normado y sea Y un subespacio de X. Si $Y \neq X$, para todo $\varepsilon > 0$ existe $x \in X$, con $\|x\| = 1$ y con $d(x, Y) > 1 - \varepsilon$.

Sean x_n, x_m con $n > m$.

$$T(x_n) - T(x_m) = -(x_n - T(x_n)) - (x_m - T(x_m)) + x_n - x_m$$

pero $x_n - T(x_n) = (Id - T_n)(x_n) \in X_{n+1} \subset X_{m+1}$ y también $(x_m - T(x_m)) + x_n \in X_{m+1}$, luego

$$\|T(x_n) - T(x_m)\| = \| - (x_n - T(x_n)) - (x_m - T(x_m)) + x_n - x_m\| \geq d(X_{m+1}, x_m) \geq \frac{1}{2}.$$

Pero esto es una contradicción ya que T es compacto ya (x_n) una sucesión acotada, por lo que $\big(T(x_n)\big)$ admitirá alguna subsucesión de Cauchy.

□

TEOREMA 36 *Sean X un espacio de Banach y $T \in \mathcal{B}(X)$ un operador compacto. Si $Id - T$ es inyectiva, entonces $(Id - T)(X) = X$ y el operador $Id - T$ es invertible.*

Prueba:
Como $Id - T$ es inyectivo y por el teorema anterior $(Id - T)(X) = X$, para ver que es invertible sólo necesitamos comprobar que $(Id - T)^{-1}$ es continuo. Pero esto es consecuencia directa del corolario 12.

□

Podemos hacer una demostración directa de que $(Id - T)^{-1}$ es continuo. Para ello basta probar, (según vimos al estudiar isomorfirmos topológicos) que existe una constante positiva m tal que para todo vector $x \in X$ con norma 1, se cumpla que $\|(Id - T)(x)\|\| \geq m$. [9]

Supongamos pues que la desigualdad anterior no se cumpla. Entonces para todo entero positivo n existe un vector y_n de norma 1 y tal que $\|(Id - T)(y_n)\|\| < \frac{1}{n}$. (Esto en paerticular nos dice que $y_n - T(y_n) \to 0_X$). Como la sucesi'on (y_n) es acotada y el operador T es compacto, la sucesión $(T(y_n))$ admitirá una subsucesión $(T(y_{n_k}))$ que será convergente a cierto vector $w \in X$. Pero como

$$y_{n_k} = y_{n_k} - T(y_{n_k}) + T(y_{n_k})$$

$0_X = \lim_k [y_{n_k} - T(y_{n_k})]$, resulta que existe $\lim_k y_{n_k} = 0_X + \lim_k T(y_{n_k}) = w$. Esto, en particular, nos dice que $\|w\| = 1$. Pero como T es continuo, de ser $\lim_k y_{n_k} = w$ se sigue que $w = \lim_k T(y_{n_k}) = T(w)$, y entonces $w \in \ker(Id - T)$. Pero siendo inyectiva $(Id - T)$, $\ker(Id - T) = \{0_X\}$, lo que implica $w = \{0_X\}$, cosa absurda pues $\|w\| = 1$.

Ejemplo 39 Para el operador $T : \ell_2 \to \ell_2$ dado por

$$T(x_1, x_2, \ldots) = (0, 0, x_1, x_2, \ldots)$$

[9] Pues entonces para todo vector de la esfera unidad de X con $\|(Id - T)^{-1}(x)\| \neq 0$

$$\|(Id - T)\left(\frac{(Id - T)^{-1}(x)}{\|(Id - T)^{-1}(x)\|}\right)\| \geq m$$

es decir, $\frac{1}{m} \geq \|(Id - T)^{-1}(x)\|$, lo que pueba que $(Id - T)^{-1}$ es acotada en la esfera unidad de X, o sea continua.

se tiene que

$$(Id - T)(x_1, x_2, \ldots) = (x_1, x_2, x_3 - x_1, x_4 - x_2, \ldots)$$

y entonces

$$\ker(Id - T) = \{0_{\ell_2}\}.$$

Por tanto $Id{-}T$ es inyectivo. Sin embargo de ser $(Id{-}T)(x) = v$ con $v = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots\right)$,

$$\left.\begin{array}{l} x_1 = 1 \\ x_2 = \frac{1}{2} \\ x_3 - x_1 = \frac{1}{3} \\ \vdots \\ x_n - x_{n-2} = \frac{1}{n} \\ \vdots \end{array}\right\}$$

de donde resulta

$$\left.\begin{array}{l} x_1 = 1 \\ x_2 = \frac{1}{2} \\ x_3 = 1 + \frac{1}{3} \\ x_4 = \frac{1}{2} + \frac{1}{4} \\ x_5 = 1 + \frac{1}{3} + \frac{1}{5} \\ x_6 = \frac{1}{2} + \frac{1}{4} + \frac{1}{6} \\ \vdots \end{array}\right\}.$$

Al ser $x_{2n+1} \geq 1$ se llega a una contradicción, pues entonces $x \notin \ell_2$. Luego $Id - T$ no es sobre, y por tanto no puede ser invertible. Según el teorema anterior, T no puede ser compacto.

Teorema 37 *Sean X un espacio de Banach y $T \in \mathcal{B}(X)$ un operador compacto. Entonces todo valor espectral no nulo de T es un valor propio (no nulo) de T*

Prueba:

Sea λ un valor espectral no nulo de T, es decir tal que $\lambda Id - T$ es no invertible. Se tiene entonces que $Id - \lambda^{-1}T$ es también no invertible. Dado que el operador $\lambda^{-1}T$ es compacto, por el teorema anterior, $Id - \lambda^{-1}T$ es no inyectiva (pues si lo fuera sería invertible). En ese caso $\lambda(Id - \lambda^{-1}T) = \lambda Id - T$ es también no inyectiva, y tiene algún vector no nulo, digamos e, en su núcleo. Luego $(\lambda Id - T)(e) = 0_X$, o sea que $T(e) = \lambda e$ y λ es un valor propio de T. Como todo valor propio (no nulo) es un valor espectral siempre, se ha completado la prueba.

□

4.6. Espectro de operadores autoadjuntos en espacios de Hilbert

En adelante X será un espacio de Hilbert.

Proposición 29 *Sea X un espacio de Hilbert y sea $T \in \mathcal{B}(X)$ un operador autoadjunto. Entonces:*
1) Para todo $x \in X$ se cumple que $\langle T(x), x\rangle$ es un número real.
2) $\|T\| := \sup\{|\langle T(x), x\rangle| : \|x\| \leq 1\}$.
3) Los valores propios de T son reales y a valores propios distintos corresponden subespacios ortogonales.

Prueba:
Sea $x \in X$. Se tiene, por ser T autoadjunto,

$$\langle T(x), x\rangle = \langle x, T(x)\rangle = \overline{\langle T(x), x\rangle}$$

lo que prueba (1).

Para ver (2), notemos que si T es el operador nulo nada hay que probar. Sea pues T no nulo y sea $s := \sup\{|\langle T(x), x >| : \|x\| \leq 1\}$. Por la desigualdad de Cauchy-Schwartz, para todo $x \in X$ con norma 1,

$$|\langle T(x), x\rangle| \leq \|T(x)\|\|x\| \leq \|T\|\|x\|\|x\| = \|T\|,$$

luego $s \leq \|T\|$.
Sea ahora $z \in X$ tal que $\|z\| = 1$ y $\|T(z)\| \neq 0$. Como T es no nulo existen vectores z en estas condiciones. Consideremos los vectores $v := \|T(z)\|^{\frac{1}{2}} z$, $w := \|T(z)\|^{\frac{-1}{2}} T(z)$. Es inmediato que $\|v\|^2 = \|T(z)\|\|z\| = \|T(z)\|$ y $\|w\|^2 = \|T(z)\|^{-1}\|T(z)\|^2 = \|T(z)\|$.
Sean $y_1 := v + w$, $y_2 := v - w$. Haciendo cálculos elementales, y teniendo en cuenta que T es autoadjunto, tenemos:

$$\begin{aligned} \langle T(y_1), y_1\rangle - \langle T(y_2), y_2\rangle &= 2(\langle T(v), w\rangle + \langle T(w), v\rangle) \\ &= 2(\langle T(z), T(z)\rangle + \langle T^2(z), z\rangle) \\ &= 2(\langle T(z), T(z)\rangle + \langle T(z), T(z)\rangle) \\ &= 4\|T(z)\|^2 \end{aligned} \tag{10}$$

Por otra parte, si $y \neq 0_X$ y $x := \frac{1}{\|y\|} y$, entonces $\|x\| = 1$ y

$$|\langle T(y), y\rangle| = |\langle T(\|y\|x), \|y\|x\rangle| = \|y\|^2 |\langle T(x), x\rangle| \leq \|y\|^2 s.$$

Teniendo en cuenta esta igualdad, y también la ley del paralelogramo queda

$$\begin{aligned} |\langle T(y_1), y_1\rangle - \langle T(y_2), y_2\rangle| &\leq |\langle T(y_1), y_1\rangle| + |\langle T(y_2), y_2\rangle| \\ &= \|y_1\|^2 s + \|y_2\|^2 s \\ &= s(\|v + w\|^2 + \|v - w\|^2) \\ &= 2s(\|v\|^2 + \|w\|^2) \\ &= 4s\|T(z)\|. \end{aligned} \tag{11}$$

Por tanto, combinando las relaciones (10,11):

$$4\|T(z)\|^2 = \langle T(y_1), y_1\rangle - \langle T(y_2), y_2\rangle \leq |\langle T(y_1), y_1\rangle - \langle T(y_2), y_2\rangle| \leq 4s\|T(z)\|$$

de donde resulta que $\|T(z)\| \leq s$.

Finalmente, de esta desigualdad se sigue que

$$\|T\| = \sup\{\|T(x)\| : x \in X, \|x\| = 1\} = \sup\{\|T(z)\| : z \in X, \|z\| = 1, \|T(z)\| \neq 0\} \leq s$$

lo que completa la prueba de que $\|T\| = s$.

Sean $\lambda \in \mathbb{C}$ un valor propio no nulo de T y $e \in X$ uno de sus vectores propios asociados. Se tiene

$$\lambda\langle e, e\rangle = \langle \lambda e, e\rangle = \langle T(e), e\rangle = \langle e, T(e)\rangle = \langle e, \lambda e\rangle = \overline{\lambda}\langle e, e\rangle$$

lo que prueba que $\lambda = \overline{\lambda}$ es decir que λ es un número real. Sean ahora λ, μ valores propios distintos de T, $x \in V_\lambda$, $y \in V_\mu$:

$$\lambda\langle x, y\rangle = \langle \lambda x, y\rangle = \langle T(x), y\rangle = \langle x, T(y)\rangle = \langle x, \mu y\rangle = \mu\langle x, y\rangle.$$

Como $\lambda \neq \mu$ ha de ser $\langle x, y\rangle = 0$.

$\square$

TEOREMA 38 *Sea X un espacio de Hilbert y sea $T \in \mathcal{B}(X)$ un operador autoadjunto y compacto. Entonces T tiene un valor propio de módulo igual a $\|T\|$.*

Prueba:

Si T es el operador nulo el único valor propio que admite es el 0, luego el resultado se cumple trivialmente. Supongamos pues que $\|T\| > 0$. Como T es autoadjunto, por la proposición anterior, $\|T\| := \sup\{|\langle T(x), x > | : \|x\| \leq 1\}$, y por definición de supremo existe x_n en B_X tal que

$$\|T\| - \frac{1}{n}\langle\big|\langle T(x_n), x_n\rangle\big| \leq \|T\|,$$

es decir que podemos dar una sucesión (x_n) en B_X tal que

$$\lim_n |\langle T(x_n), x_n\rangle| = \|T\|.$$

La sucesión de números reales $(\langle T(x_n), x_n\rangle)$ es acotada (pues es convergente en módulo), luego admite una subsucesión $(\langle T(x_{n_k}), x_{n_k}\rangle)$ que es convergente a cierto número real λ. Desde luego, $|\lambda| = \|T\|$, con lo que podemos asegurar que $\lambda \neq 0$. Veremos que λ es un valor propio de T, lo que completará la prueba.

Como T es un operador compacto y la sucesión (x_{n_k}) es acotada, la sucesión $(T(x_{n_k}))$ admitirá una subsucesión $(T(x_{n_{k_j}}))$ convergente a cierto vector $z \in X$ Como para cada natural j,

$$\begin{aligned}
\|T(x_{n_{k_j}}) - \lambda x_{n_{k_j}}\|^2 &= \left\langle T(x_{n_{k_j}}) - \lambda x_{n_{k_j}}, T(x_{n_{k_j}}) - \lambda x_{n_{k_j}} \right\rangle \\
&= \|T(x_{n_{k_j}})\|^2 - 2\lambda\langle T(x_{n_{k_j}}), x_{n_{k_j}}\rangle + \lambda^2\|x_{n_{k_j}}\|^2 \\
&\leq \|T(x_{n_{k_j}})\|^2 - 2\lambda\langle T(x_{n_{k_j}}), x_{n_{k_j}}\rangle + \lambda^2 \\
&\leq \|T\|^2 - 2\lambda\langle T(x_{n_k}), x_{n_j}\rangle + \lambda^2 \\
&= 2\lambda^2 - 2\lambda\langle T(x_{n_{k_j}}), x_{n_{k_j}}\rangle
\end{aligned}$$

y además $\lim_j \langle T(x_{n_{k_j}}), x_{n_{k_j}}\rangle = \lambda$, se sigue inmediatamente que

$$\lim_j \|T(x_{n_{k_j}}) - \lambda x_{n_{k_j}}\| = 0.$$

Por tanto $T(x_{n_{k_j}}) - \lambda x_{n_{k_j}} \to 0_X$ y como $T(x_{n_{k_j}}) \to z$,

$$\lambda x_{n_{k_j}} = T(x_{n_{k_j}}) - [T(x_{n_{k_j}}) - \lambda x_{n_{k_j}}] \to z.$$

Sea $y = \frac{1}{\lambda} z$. Naturalmente $\lim_j x_{n_{k_j}} = y$, luego por la continuidad de T y del producto escalar

$$\langle T(y), y\rangle = \lim_j \langle T(x_{n_{k_j}}), x_{n_{k_j}}\rangle = \lambda \neq 0$$

luego $y \neq 0_X$ y además, de nuevo por la continuidad de T,

$$T(y) = \lim_j T(x_{n_{k_j}}) = z = \lambda y$$

lo que prueba que λ es un valor propio de T y que y es un vector propio asociado al valor propio λ.

□

Recordemos que si X es un espacio de Hilbert y $M \subset X$ es no vacío se cumple:
a) $M^{\perp} := \{x \in X : \forall y \in M, \langle x, y\rangle = 0\}$ es un subespacio vectorial cerrado de X.
b) $(M^{\perp})^{\perp} \supset M$.
c) Si M es un subespacio cerrado de X, $X = M \bigoplus M^{\perp}$.
d) $M^{\perp\perp}$ es el menor subespacio cerrado de X que contiene a M.
e) $M^{\perp\perp\perp} = M^{\perp}$.

TEOREMA 39 *(Teorema espectral de Hilbert-Schmidt.)*
Sea X un espacio de Hilbert y sea $T \in \mathcal{B}(X)$ un operador autoadjunto y compacto. Entonces

1) Si T es de rango finito entonces los valores propios de T forman un conjunto finito $\{\lambda_1, \ldots, \lambda_k\}$ y existe un conjunto ortonormal $\{x_1, \ldots x_k\}$ con $x_i \in V_{\lambda_i}$ $(i = 1, \ldots, k)$, de manera que cada $x \in X$ se puede escribir como

$$x = \sum_{i=1}^{k} \langle x, x_i\rangle x_i + y$$

con $y \in \ker(T)$. (Por lo que $T(x) = \sum_{i=1}^{k} \lambda_i \langle x, x_i\rangle x_i$).

2) Si el rango de T es de dimensión infinita entonces los valores propios no nulos forman una sucesión infinita de números reales que converge a 0, y existe un conjunto ortonormal numerable $\{x_1, \ldots x_k, \ldots\}$ con $x_i \in V_{\lambda_i}$ $(i = 1, 2, \ldots,)$, de manera que cada $x \in X$ se puede escribir como

$$x = \sum_{i=1}^{\infty} \langle x, x_i\rangle x_i + y$$

con $y \in \ker(T)$. (Por lo que $T(x) = \sum_{i=1}^{\infty} \lambda_i \langle x, x_i\rangle x_i$). Además la multiplicidad de cada λ_i es finita e igual al número de veces que aparece repetido en la sucesión.

Prueba:

Supongamos X en el caso (2). Sea λ_1 un valor propio de T con $|\lambda_1| = \|T\|$. Tal valor propio existe según hemos visto en el teorema anterior. Sea $x_1 \in V_{\lambda_1}$ con $\|x_1\| = 1$ y sean $X_1 = X$, $T_1 = T$. Definimos $X_2 := \{x_1\}^{\perp}$, que será un subespacio cerrado de X_1, y por tanto será un espacio de Hilbert. Vamos a ver que X_2 es T_1-invariante, es decir $T(X_2) \subset X_2$. En efecto, si $x \in X_2$, por ser T autoadjunto,

$$\langle T_1(x), x_1\rangle = \langle x, T(x_1)\rangle = \langle x, \lambda_1 x_1\rangle = 0,$$

luego $T_1(x)$ es ortogonal a $\{x_1\}$, es decir, pertenece a X_2. Definiremos el operador $T_2 : X_2 \to X_2$ por $T\big|_{X_2}$ que evidentemente será autoadjunto y compacto.

Dado $x \in X$, sea $y(x) := x - \langle x, x_1\rangle x_1$. Observemos que,

$$\langle y(x), x_1\rangle = \langle x, x_1\rangle - \langle x, x_1\rangle = 0$$

lo que prueba que $y(x) \in X_2$. Si fuera $y(x) = 0_X$ para todo x entonces $x = \langle x, x_1\rangle x_1$. Luego, salvo en el caso de que X sea de dimensión 1, el subespacio X_2 es no trivial, es decir no se reduce al $\{0_X\}$. Por otra parte, de ser $y(x) \in \ker(T)$ para todo $x \in X$, se seguiría que

$$T(x) = \langle x, x_1\rangle T(x_1) = \langle x, x_1\rangle \lambda_1 x_1$$

es decir que todo vector $T(x)$ pertenecería al subespacio generado por $\{x_1\}$, lo que contradice al hecho de tener T rango de dimensión infinita. Luego podemos garantizar que en X_2 hay vectores no nulos que no pertenecen al núcleo de T (ni al de T_2), es decir que T_2 no es idénticamente nulo. Sabemos entonces que T_2 admite un valor propio λ_2 que verifica $|\lambda_2| = \|T_2\|$ por lo que λ_2 será distinto de 0. Por definición de norma de un operador es claro que $\|T_2\| \leq \|T_1\|$. Sabemos también que el correspondiente vector propio x_2 (que elegimos normalizado), es ortogonal a x_1, precisamente por pertenecer a $X_2 := \{x_1\}^{\perp}$.

Reiterando este proceso, supongamos construidos para $i = 2, \ldots, n-1$ subespacios cerrados $X_2, \ldots, X_{n-1}$, T-invariantes con $X_i \subset X_{i-1}$; $\lambda_1, \lambda_2, \ldots, \lambda_{n-1}$ valores propios de T tales que $\|T\big|_{X_i}\| = |\lambda_i|$ y los correspondientes vectores propios $\{x_1, x_2, \ldots, x_{n-1}\}$ ortonormales con $x_i \in V_{\lambda_i} \cap X_i$ $(i = 2, \ldots, n-1)$.

Pondremos entonces para $(n \geq 2)$, $X_n := \{x_1, \ldots, x_{n-1}\}^{\perp}$ que es un subespacio cerrado de X, y por tanto un espacio de Hilbert. En este caso, todo vector de X_n es ortogonal en particular a $x_2, \ldots, x_{n-2}$ luego pertenece a X_{n-1}, luego $X_n \subset X_{n-1}$. Es claro que si $x \in X_n$, para $i = 1, \ldots, n-1$ se tiene por ser T autoadjunto, que

$$\langle T(x), x_i\rangle = \langle x, T(x_i)\rangle = \langle x, \lambda_i x_i\rangle = 0$$

luego $T(x)$ es ortogonal a $x_1, \ldots, x_{n-1}$ por lo que pertenece a X_n, es decir que X_n es también T invariante. Sea $T_n := T\big|_{X_n}$, que evidentemente será un operador compacto y autoadjunto del espacio de Hibert X_n.

Dado $x \in X$, sea $y(x) := x - \sum_{i=1}^{n-1} \langle x, x_i\rangle x_i$. Observemos que, para $i = 2, \ldots, n-1$ se tiene que

$$\langle y(x), x_i\rangle = \langle x, x_i\rangle - \langle x, x_i\rangle = 0$$

lo que prueba que $y(x) \in X_n$. Si fuera $y(x) = 0_X$ para todo x entonces $x = \sum_{i=1}^{n-1} \langle x, x_i\rangle x_i$. Luego, salvo en el caso de que X sea de dimensión $n-1$, el subespacio X_n es no trivial, es decir no se reduce al $\{0_X\}$. Por otra parte, de ser $T(y(x)) = 0_X$ para todo $x \in X$, se seguiría que

$$T(x) = \sum_{i=1}^{n-1} \langle x, x_i \rangle T(x_i) = \sum_{i=1}^{n-1} \langle x, x_i \rangle \lambda_i (x_i)$$

es decir que $T(x)$ pertenecería al subespacio generado por $\{x_1, \ldots, x_n\}$, lo que contradice al hecho de ser T de rango de dimensión infinita. Luego podemos garantizar que en X_n hay además vectores que no pertenecen al núcleo de T (ni al de T_n), es decir que T_n no es idénticamente nulo. Tendrá pues T_n un valor propio λ_n con $|\lambda_n| = \|T_n\|$ lo que en particular nos dice que $\lambda_n \neq 0$. Elegiremos el correspondiente vector propio normalizado x_n, que por pertenecer a X_n será ortogonal a $\{x_1, \ldots, x_{n-1}\}$.

Para la sucesión de valores propios (λ_n) que hemos construido por inducción se cumple que $|\lambda_{n+1}| \leq |\lambda_n|$. Vamos a ver que $\lambda_n \to 0$. Caso contrario, como $(|\lambda_n|)$ es monótona decreciente, existiría $\varepsilon > 0$ con $|\lambda_n| > \varepsilon$ para todo natural n. En ese caso la sucesión de vectores $\left(\frac{1}{\lambda_n} x_n\right)$ sería acotada, y verifica que

$$T(\frac{1}{\lambda_n} x_n) = x_n.$$

Por ser T compacto (x_n) debería de admitir una subsucesión convergente, pero esto es absurdo pues el conjunto $\{x_1, \ldots\}$ es ortonormal (y dos cualesquiera de sus miembros distan $\sqrt{2}$ lo que imposibilita que admita ninguna subsucesión de Cauchy).

Una propiedad importante de la sucesión ortonormal (x_n) así construida es que

$$\{x_1, x_2, \ldots\}^{\perp} = \bigcap_{n=1}^{\infty} X_n = \ker(T) \tag{12}$$

En efecto es claro que $x \in \bigcap_{n=1}^{\infty} X_n$ si y sólo si pertenece a $\{x_1, x_2, \ldots, x_n\}^{\perp}$ para todo entero positivo n, lo que equivale a que $x \in \{x_1, x_2, \ldots\}^{\perp}$, lo que prueba la primera igualdad. Además si $x \in \bigcap_{n=1}^{\infty} X_n$ entonces para todo entero positivo n,

$$\|T(x)\| = \|T_n(x)\| \leq \|T_n\| \|x\| = |\lambda_n| \|x\|$$

y como $\lim_n \lambda_n = 0$, se sigue que $T(x) = 0_X$ es decir que $x \in \ker(T)$. Recíprocamente, si $x \in \ker(T)$ entonces $\langle T(x), x_n \rangle = 0$ para todo $n \in \mathbb{N}$, pero en ese caso,

$$0 = \langle T(x), x_n \rangle = \langle x, T(x_n) \rangle = \langle x, \lambda_n x_n \rangle = \lambda_n \langle x, x_n \rangle$$

y como ninguno de los λ_n es nulo se sigue que x es ortogonal a todos y cada uno de los x_n, es decir $x \in \{x_1, x_2, \ldots\}^{\perp}$, lo que completa la prueba de la segunda igualdad en (12).

Si λ es un valor propio de T distinto de los λ_n y $x \neq 0_X$ es cualquiera de sus vectores propios asociados, al ser para cada n, $\lambda \neq \lambda_n$ por la Proposición (29), $x \perp x_n$. Luego $x \in \{x_1, x_2, \ldots\}^{\perp} = \ker(T)$, y por tanto $0_X = T(x) = \lambda x$, luego $\lambda = 0$.

Como $\lim_n \lambda_n = 0$ y los λ_n son no nulos, ningún número distinto de cero puede figurar un número infinito de veces en la sucesión. Por otra parte, si un determinado $\lambda_k \in \{\lambda_n : n \in \mathbb{N}\}$ aparece repetido p-veces en la sucesión (λ_n), es decir si $\lambda = \lambda_{n_1} = \lambda_{n_2} = \ldots = \lambda_{n_p}$ $(n_1 \langle n_2 < \ldots < n_p)$ entonces el subespacio propio asociado V_λ (tal como ha sido contruida la sucesión (x_n)) contiene p vectores propios ortonormales, digamos $\{x_{n_1}, x_{n_2}, \ldots, x_{n_p}\}$ $(n_1 < n_2 < \ldots < n_p)$, luego la dimensión de V_λ es mayor o igual que p. Pero no puede ser mayor que p pues en ese caso podríamos completar $\{x_{n_1}, x_{n_2}, \ldots, x_{n_p}\}$ hasta formar una base algebraica de V_{λ_k}, y luego aplicarle el proceso de ortonormalización de Gram-Schmidt, obteniendo de esta manera una nueva base ortonormal de V_{λ_k} que, por supuesto, contendría a los $\{x_{n_1}, x_{n_2}, \ldots, x_{n_p}\}$ (vectores ortonormales) y, al menos, a otro vector v ortogonal a los anteriores y con norma 1. Como v es también un vector propio del valor propio repetido λ, es ortogonal a cada uno de los vectores

$$\{x_n : n \in \mathbb{N}, n \notin \{n_1, \ldots, n_p\}\}$$

(que son vectores propios de valores propios *diferentes* del de v). En suma v es ortogonal al conjunto $\{x_1, x_2, \ldots\}$, es decir $v \in \{x_1, x_2, \ldots\}^\perp = \ker(T)$, lo que implica $\lambda v = T(v) = 0_X$, una contradicción.

Por otra parte, sea S el subespacio de X que es clausura del generado por los vectores $\{x_1, x_2, \ldots\}$. Por (12) y por el teorema de la proyección

$$\begin{aligned} X &= (\ker(T)^\perp \bigoplus \ker(T) \\ &= ((\{x_1, x_2, \ldots\})^\perp)^\perp \bigoplus \ker(T) \\ &= S \bigoplus \ker(T). \end{aligned}$$

Luego todo vector x de X puede descomponerse en forma única como $x = z + y$ con $z \in S$, $y \in \ker(T)$. Todo elemento z de S puede escribirse en forma única como $z = \sum_{i=1}^\infty \langle z, x_n\rangle x_n$, pues el conjunto $\{x_1, x_2, \ldots\}$ es linealmente independiente y además es ortonormal. Luego para todo vector $z \in S$, podemos escribir

$$x = z + y = \sum_{i=1}^\infty \langle z, x_n\rangle x_n + y = \sum_{i=1}^\infty \langle z + y, x_n\rangle x_n + y = \sum_{i=1}^\infty \langle x, x_n\rangle x_n + y$$

ya que $\langle y, x_n\rangle = 0$ para todo $n \in \mathbb{N}$. De la expresión anterior y teniendo en cuenta que T es continua se obtiene

$$T(x) = \sum_{i=1}^\infty \langle x, x_n\rangle T(x_n) + T(y) = \sum_{i=1}^\infty \langle x, x_n\rangle \lambda_n x_n\,.$$

Pasamos ahora al caso en que T tenga rango finito. En primero lugar afirmamos que T sólo puede tener un número finito de valores propios no nulos distintos. Caso contrario, es decir, si existiera una sucesión (λ_n) de valores propios de T con $\lambda_i \neq \lambda_j$ para $i \neq j$. Sea y_n un vector propio normalizado correspondiente al valor propio λ_n. Es claro que a valores propios distintos corresponden subespacios propios ortogonales, según vimos más arriba, luego $\{y_n : n \in \mathbb{N}\}$ es un conjunto ortonormal en X (y por tanto linealmente independiente). Pero como $y_n = T(\frac{1}{\lambda_n} y_n)$ entonces $\{y_n : n \in \mathbb{N}\} \subset T(X)$ luego la dimensión de $T(X)$ no puede ser finita, contra lo que estamos suponiendo.

Supongamos que para algún natural p, $dim(T(X)) = p$. Sea, como hemos construido hasta ahora, $X_{p+1} = \{x_1, \ldots, x_p\}^\perp$. Los vectores propios $\{x_1, \ldots, x_p\}$ forman un conjunto ortonormal y están contenidos en $T(X)$, luego para todo $x \in X$ existen escalares $\alpha_1, \ldots, \alpha_n$ con

$$T(x) = \sum_{i=1}^{p} \alpha_i x_i.$$

En particular, si $x \in X_{p+1}$ se tiene, según hemos visto, que $T(x) \in X_{p+1}$ y entonces $\langle T(x), x_i \rangle = 0$ para $i = 1, \ldots, p$, de donde para todo $j \in \{1, \ldots, p\}$

$$0 = \langle T(x), x_j \rangle = \left\langle \sum_{i=1}^{p} \alpha_i x_i, x_j \right\rangle = \alpha_j$$

Por tanto $T(x) = 0_X$. Siendo x arbitrario en X_{p+1} hemos probado que $X_{p+1} \subset \ker(T)$.

De aquí que el algoritmo anterior de construcción de la sucesión de los subespacios $X_1 \supset X_2 \supset \ldots$ se llegará a un natural k para el que X_k está contenido en $\ker(T)$, y el correspondiente operador T_k será idénticamente nulo. A partir de este punto razonamos como en el caso anterior: Si $\{x_1, ..., x_{k-1}\}$ son los vectores propios normalizados producidos por el algoritmo, con $x_i \in X_i$ $(i = 1, \ldots, k-1)$ entonces el vector $x - \sum_{i=1}^{k-1} \langle x, x_i \rangle x_i$ pertenece a X_k, o sea, es nulo, luego $0_X = T(x - \sum_{i=1}^{k-1} \langle x, x_i \rangle x_i)$ para todo $x \in X$, lo que nos da $T(x)$ como combinación lineal de $x_1, \ldots, x_{k-1}$, es decir $T(x) = \sum_{i=1}^{k-1} \langle x, x_i \rangle \lambda_i x_i$ y $x = \sum_{i=1}^{k-1} \langle x, x_i \rangle x_i + z$ con $z \in \ker(T)$.

$\square$

4.6.1. Solución de ecuaciones

Consideremos la ecuación

$$T(x) - \mu x = y \tag{13}$$

donde x es la incógnita en un espacio de Hilbert X, y un elemento fijo de X, μ una constante no nula y $T : X \to X$ un operador autoadjunto y compacto en dicho espacio.

TEOREMA 40 *Respecto de las soluciones de la ecuación (13) se tiene:*

1) Si μ no es un valor propio de T entonces existe una única solución de (13) que viene dada por

$$x = \frac{1}{\mu} \left(\sum_{n=1}^{\infty} \frac{\lambda_n \langle y, x_n \rangle}{\lambda_n - \mu} x_n - y \right)$$

donde $\{\lambda_n : n \in \mathbb{N}\}$ es el conjunto (numerable) de los valores propios no nulos de T, y para cada y, para cada $n \in \mathbb{N}$, x_n es un vector propio asociado al valor propio λ_n.

2) Si μ es un valor propio de T pero $y \in V_\mu(T)^\perp$, entonces existe una solución de (13) que viene dada por

$$x = \frac{1}{\mu} \left(\sum_{n=1}^{\infty} \sigma_n x_n - y \right)$$

donde $\{\lambda_n : n \in \mathbb{N}\}$ es el conjunto (numerable) de los valores propios no nulos de T, y para cada y, para cada $n \in \mathbb{N}$, x_n es un vector (de norma 1) propio asociado al valor propio λ_n, y $\sigma_n = \frac{\lambda_n \langle y, x_n \rangle}{\lambda_n - \mu}$ si $\mu \neq \lambda_n$, y σ_n es arbitrario si $\mu = \lambda_n$.

Prueba:

por el Teorema Espectral de Hilbert Schmidt el operador T admite la expresión

$$T(x) = \sum_{i=1}^{\infty} \lambda_i \langle x, x_i \rangle x_i$$

donde los escalares λ_i y los vectores (propios) x_i son como en el enunciado (salvo que T sea de rango finito, en cuyo caso se expresará del mismo modo pero con una suma finita). Si $x \in X$ es solución de la ecuación (13) ha de ser $x = \frac{1}{\mu}(T(x) - y)$, luego

$$x = \frac{1}{\mu}\left(\sum_{i=1}^{\infty} \lambda_i \langle x, x_i \rangle x_i - y\right)$$

Entonces, para cada $n \in \mathbb{N}$

$$\langle x, x_n \rangle = \frac{1}{\mu}\left(\sum_{i=1}^{\infty} \lambda_i \langle x, x_i \rangle \langle x_i, x_n \rangle - \langle y, x_n \rangle\right) = \frac{1}{\mu}(\lambda_n \langle x, x_n \rangle - \langle y, x_n \rangle).$$

De aquí resulta

$$\langle x, x_n \rangle = \frac{\frac{-\langle y, x_n \rangle}{\mu}}{1 - \frac{\lambda_n}{\mu}} = \frac{\langle y, x_n \rangle}{\lambda_n - \mu}, \tag{14}$$

y sustituyendo tenemos un único valor posible para x:

$$x = \frac{1}{\mu}\left(\sum_{n=1}^{\infty} \lambda_n \frac{\langle y, x_n \rangle}{\lambda_n - \mu} x_n - y\right)$$

que existirá como elemento de X si la serie del segundo miembro converge en X.

Para ver que una serie de la forma $\sum_{n=1}^{\infty} \beta_n x_n$ (con $\beta_n \in \mathbb{C}$, $x_n \in X$ para cada n) es convergente en X podemos aplicar el teorema de Riesz-Fischer.[10] Para ver que la serie numérica $\sum_{n=1}^{\infty} \left|\lambda_n \frac{\langle y, x_n \rangle}{\lambda_n - \mu}\right|^2$ es convergente tengamos en cuenta que, dado que $\lambda_n \to 0$ y que $\mu \neq 0$ existirá $\alpha > 0$ tal que $|\lambda_n - \mu| > \alpha$ para cada n; También $|\lambda_n| \leq \|T\|^2$ para todo n, luego tenemos

$$\sum_{n=1}^{\infty} \left|\lambda_n \frac{\langle y, x_n \rangle}{\lambda_n - \mu}\right|^2 \leq \|T\|^2 \sum_{n=1}^{\infty} \left|\frac{\langle y, x_n \rangle}{\lambda_n - \mu}\right|^2 \leq \frac{\|T\|^2}{\alpha^2} \sum_{n=1}^{\infty} |\langle y, x_n \rangle|^2$$

[10] Si $\mathcal{S} = \{\varphi^0, \varphi^1, \ldots\}$ es un sistema ortonormal en el espacio de Hilbert X. Si (c_n) es una sucesión de números complejos con $\sum_{n=0}^{\infty} |c_n|^2 < \infty$. (Es decir, sea (c_n) un elemento de ℓ_2). Entonces existe el elemento $y = \sum_{n=0}^{\infty} c_n \varphi^n$ en X. Además

a) Los coeficientes de Fourier de y respecto de $\mathcal{S}$ son los c_k, es decir, para cada $k \geq 0$,

$$\langle y, \varphi^k \rangle = c_k.$$

b)

$$\|y\|^2 = \sum_{n=0}^{\infty} |c_n|^2 = \|(c_n)\|_{\ell_2}$$

y por la desigualdad de Bessel,

$$\sum_{n=1}^{\infty} \left| \lambda_n \frac{\langle y, x_n \rangle}{\lambda_n - \mu} \right|^2 \leq \frac{\|T\|^2}{\alpha^2} \sum_{n=1}^{\infty} |\langle y, x_n \rangle|^2 \leq \frac{\|T\|^2}{\alpha^2} \|y\|^2 .$$

Así pues queda probado que existe la única solución de la ecuación (13) y es

$$x = \frac{1}{\mu} \left(\sum_{n=1}^{\infty} \lambda_n \frac{\langle y, x_n \rangle}{\lambda_n - \mu} x_n - y \right).$$

Supongamos ahora que para algún entero positivo k es $\mu = \lambda_k$. El cálculo hecho en (14) es decir

$$\langle x, x_n \rangle = \frac{\langle y, x_n \rangle}{\lambda_n - \mu},$$

tiene sentido para aquellos valores de n para los que $\lambda_n \neq \mu$. A lo sumo hay un número finito de enteros positivos k para los que $\lambda_k = \mu$. La expresión

$$x = \frac{1}{\mu} \left(\sum_{i=1}^{\infty} \lambda_i \langle x, x_i \rangle x_i - y \right)$$

sigue siendo válida, aunque no permite calcular el coeficiente de x_k. Pero si y es ortogonal al vector propio x_k, (lo que ocurre con seguridad si y es ortogonal al subespacio de vectores propios $V_\mu(T)$) entonces de la expresión anterior se sigue que

$$\langle x, x_k \rangle = \left\langle \frac{1}{\mu} \left(\sum_{i=1}^{\infty} \lambda_i \langle x, x_i \rangle x_i - y \right), x_k \right\rangle = \frac{1}{\mu} \lambda_k \langle x, x_k \rangle - 0 = \langle x, x_k \rangle$$

luego cualquier escalar es admisible como valor de $\langle x, x_k \rangle$, lo que permite tomar arbitrarios los coeficientes de los x_k para los que $\lambda_k = \mu$,

$\square$

TEOREMA 41 *Sea X un espacio de Hilbert y sea $T \in \mathcal{B}(X)$ un operador compacto y autoadjunto. Para $\mu \neq 0$ consideremos la ecuación*

$$T(x) - \mu x = y \tag{15}$$

y la correspondiente ecuación homogénea

$$T(x) - \mu x = 0_X. \tag{16}$$

Entonces se verifica:

1) Para todo $y \in X$ la ecuación (15) tiene solución única si y sólo si la ecuación homogénea correspondiente (16) no tiene más solución que la trivial 0_X.

2) La ecuación (15) tiene solución si y sólo si su término independiente y es ortogonal a toda solución de la ecuación homogénea.

Prueba: Si para cada término independiente y la ecuación completa tiene solución única, en particular tomando $y = 0_X$ tendrá una solución $x_0 \in X$, que evidentemente es solución única de la ecuación homogénea (16).

Recíprocamente, si la ecuación homogénea (16) sólo tiene la solución trivial 0_X, μ no es valor propio de T, luego por el teorema anterior la ecuación completa (15) tiene solución única.

Para ver (2) supongamos que $y \in X$ es ortogonal a toda solución (trivial o no) de la ecuación homogénea (16). Si esta ecuación homogégena admite únicamente la solución trivial, por (1) la ecuación completa (15) tiene solución única. Si la ecuación homogénea tuviera alguna solución no trivial, entonces μ es un valor propio de T y el conjunto de soluciones de la ecuación homogénea es justamente el espacio de vectores propios $V_\mu(T)$. Si $y \in X$ es ortogonal a toda solución de la ecuación homogénea, ha de ser $y \in (V_\mu(T))^\perp$, y nuevamente por el teorema anterior vemos que la ecuación completa (15) tiene solución.

Recíprocamente, supongamos que $x \in X$ es una solución de la ecuación completa (15). Para cada solución no trivial z de la ecuación homogénea (16), se tiene, por ser T autoadjunto,

$$\begin{aligned} \langle y, z\rangle &= \langle T(x) - \mu x, z\rangle \\ &= \langle T(x), z\rangle - \mu\langle x, z\rangle \\ &= \langle x, T(z)\rangle - \mu\langle x, z\rangle \\ &= \langle x, \mu z\rangle - \mu\langle x, z\rangle = 0. \end{aligned}$$

(pues μ es real, ya que es un valor propio de T). Luego el término independiente $y \in X$ es ortogonal a toda solución (trivial o no) de la ecuación homogénea.

□

Corolario 14 *En las condiciones del teorema anterior, o bien para cada $y \in X$ la ecuación no homogénea tiene solución (única), o bien la ecuación homogénea tiene solución no trivial.*

Prueba : La ecuación no homogénea admite solución no trivial o no la admite. Si no la admite, μ no es valor propio de T, y entonces por el teorema (13) sabemos que para cada y la ecuación completa tiene solución única.

□

Si para un operador $A \in \mathcal{B}(X)$ se cumple la disyuntiva o anterior se dice que cumple la *Alternativa de Fredholm*.

4.7. Aplicación a ecuaciones integrales de Fredholm

Consideraremos la ecuación integral tipo Fredholm

$$x(s) = y(s) + \lambda \int_a^b k(s,t)x(t)dt \tag{17}$$

planteada en el espacio de Hilbert $X = L^2([a,b])$, en la cual $\lambda \in \mathbb{C}$, $y \in X$, $k \in L^2([a,b] \times [a,b])$ son datos. Si $k(s,t) = \overline{k(t,s)}$ el operador de Fredholm $x \mapsto K(x)$ con

$K(x)(s) := \int_a^b k(s,t)x(t)dt$ es compacto y autoadjunto. Vamos a estudiar para qué valores de λ la ecuación (17) tiene solución. Es claro que la correspondiente ecuación homogenea

$$x(s) = \lambda \int_a^b k(s,t)x(t)dt \tag{18}$$

admite siempre la solución trivial $x = 0_X$. Se dice que el número complejo $\lambda \neq 0$ es una raiz característica del núcleo k si λ^{-1} es un valor propio de K, es decir si existe $x \in X$ con $x \neq 0_X$ y con $K(x)(s) = \int_a^b k(s,t)x(t)dt = \lambda^{-1}x(s)$ o lo que es lo mismo, si la ecuación (18) admite una solución no trivial. En ese caso cada una de estas soluciones no nulas se llama *función propia* correspondiente a λ, es decir, una función propia es un vector propio asociado al valor propio λ^{-1} del operador K. Se llama *rango de* λ al orden de multiplicidad del valor propio λ^{-1}.

TEOREMA 42 *Las raices características del núcleo k tienen rango finito, y forman una sucesión (r_n) de manera que cada r_n tiene una función propia asociada x_n verificando:*

1) Si λ no es un raiz característica, la ecuación no homogénea (17) tiene una y sólo una solución que viene dada por

$$x(s) = y(s) + \lambda \sum_{n=1}^{\infty} \frac{\langle y, x_n \rangle}{\lambda - r_n} x_n(s).$$

2) Si λ es un raiz característica de rango q, entonces la ecuación no homogénea (17) tiene solución si y sólo si el término independiente $y \in X$ es ortogonal a todas las funciones propias correspondientes a λ que sean linealmente independientes. En ese caso la ecuación tiene un número infinito de soluciones que vienen dadas por

$$x(s) = y(s) - \lambda \sum_{n=1}^{\infty} s_n x_n(s).$$

donde $s_n := \frac{\langle y,x_n \rangle}{\lambda - r_n}$ si $\lambda \neq r_n$, y s_n es una constante arbitraria si $\lambda = r_n$.

Prueba: Si ponemos $\mu := \frac{1}{\lambda}$, $z = -\mu y$, la ecuación dada puede escribirso como $K(x) - \mu x = z$. Como K es operador compacto y autoadjunto, al aplicar el teorema espectral obtenemos que los valores propios de K tienen orden de multiplicidad finito y que:

1) Si μ no es un valor propio de K, la única solución de $K(x) - \mu x = z$ viene dada por

$$x = \frac{1}{\mu} \left(\sum_{n=1}^{\infty} \frac{\langle y, x_n \rangle}{\lambda - r_n} x_n - z \right)$$

donde (λ_n) es la sucesión de valores propios no nulos del operador K, y para cada n x_n es un vector propio normalizado correspondiente a λ_n . De aquí resulta inmediatamente :

$$x = \lambda \left(\sum_{n=1}^{\infty} \frac{\langle y, x_n \rangle}{\lambda - r_n} x_n + \frac{1}{\lambda} y \right) = y - \lambda \sum_{n=1}^{\infty} \frac{\langle y, x_n \rangle}{\lambda - r_n} x_n \,.$$

2) En el caso en que λ coincida con alguna raiz característica, según el teorema espectral la solución tendrá una expresión análoga a la anterior pero sustituyendo el coeficiente del término con cuya raiz coincida por una constante arbitraria.

□

5. TRES TEOREMAS CLÁSICOS EN ANÁLISIS FUNCIONAL

(...)*The open mapping theorem, the Banach Steinhaus theorem and the Hahn-Banach theorem. (...) These three results are fundamental theorems of functional analysis, and it may be argued that any book purporting to be a functional analysis text must include them.*[11]

5.1. Conjuntos convexos

Sea X un espacio vectorial. Si v, w son vectores de X usaremos la notación

$$[v,w] := \{x \in X : x = \lambda v + (1-\lambda)w, 0 \leq \lambda \leq 1\},$$

y el conjunto del segundo miembro se llamará segmento (cerrado, no orientado) de extremos v, w. Naturalmente que $[v,w] = [w,v]$, luego hay que cuidar cualquier confusión con la notación ordinaria de intervalos en la recta real. De una forma análoga podemos usar las notaciones $[v,w[,]v,w]$ y $]v,w[$.

Definición 24 *Un conjunto $A \subset X$ se llama convexo si para cualesquiera $v, w \in A$ se tiene que $[v,w] \subset A$.*

Por convenio (o por "vacuidad") el conjunto vacío es convexo.

Ejemplo 40 Son convexos:

a) Todo subespacio vectorial de X.

b) Todo subespacio afín de X, es decir, todo conjunto de la forma

$$v + S := \{v + w \in X : w \in S\}$$

con S subespacio vectorial de X. En particular, todo conjunto reducido a un sólo punto, $\{x\} = x + \{0_X\}$ con $x \in X$.

Ejemplo 41 En un espacio normado toda bola (abierta o cerrada) es un conjunto convexo.

Ejemplo 42 En el espacio vectorial $X := \mathcal{C}([0,1])$ los conjuntos
$C_1 := \{y \in X : \forall t \in [0,1] \text{ existe } y'(t) \text{ y } y'(t) + 2y(t) = \text{sen}(t)\}$.
$C_2 := \{f \in X : \forall t \in [0,1], |\int_0^t f(s)ds| \langle t\}$.
$C_3 := \{f \in X : \forall s,t \in [0,1] |f(s) - f(t)| < |s-t|\}$.
$K := \{f \in X : \forall t \in [0,1], 0 = f(0) \leq f(t) \leq f(1) = 1\}$.

Basta tratar de comprobarlo en los ejemplos anteriores para darse cuenta que la convexidad suele ser una propiedad muy fácil de verificar. Su significado intuitivo es claro: Los conjuntos convexos contienen a los segmentos que unes a dos cualesquiera de sus puntos.[12]

[11] K. Saxe, Beginning Functional Analysis, p. 151

[12] Aunque la aplicación a conjuntos de funciones como los del ejemplo anterior de puntos de vista *geométricos* es esencial en Análisis Funcional, hay que cuidar de que la intuición geométrica no vaya más

5.1.1. Propiedades inmediatas de los conjuntos convexos

Se comprueban por aplicación directa de la definición que si A, B son convexos no vacíos en el espacio vectorial X, $v \in X$, y λ es un escalar real

a) $A + B$ es convexo.

b) $v + A$ es convexo.

c) λA es convexo.

Además

d) La intersección arbitraria de una familia de conjuntos convexos es un conjunto convexo.

e) Si L es una aplicación lineal entre X y el espacio vectorial Y, entonces $L(A)$ es convexo, y si C es un conjunto convexo en Y, $L^{-1}(C)$ es convexo.

Como consecuencia de que la convexidad se mantiene bajo intersecciones arbitrarias, (propiedad (d)) , dado un conjunto cualquiera $M \subset X$ existe un conjunto convexo ue contiene al M y que está contenido en cualquier otro convexo que contega al M, (es decir que es el convexo "más pequeño"que contiene al M) . Ese conjunto se llama envoltura convexa de M:

DEFINICIÓN 25 *Se llama envoltura convexa de M a la intersección de la familia de los conjuntos convexos que contienen al M, (familia que es no vacía pues X seguro que está en ella). Para referirnos a la envoltura convexa de M usaremos la notación $co(M)$.*

Así como el subespacio vectorial generado por M, $lin(M)$, es el "más pequeño"subespacio vectorial que contiene al conjunto M y puede describirse como el conjunto de las combinaciones lineales (¡y por tanto de un número finito!) de vectores de M, también podemos encontrar una descripción parecida para la envoltura convexa de M, $co(M)$. (Notar que $lin(\emptyset) = \{0_X\}$, y $co\{\emptyset\} = \emptyset$).

DEFINICIÓN 26 *Sean $x_1, ..., x_n$ vectores de X. Una combinación lineal de la forma $\sum_{i=1}^{n} a_i x_i$ se llama convexa si los escalares a_i son reales no negativos y con suma igual a* 1.

Con este vocabulario podrá decirse que un conjunto es convexo si contiene las combinaciones lineales convexas de dos cualesquiera de sus vectores. Veremos que puede afirmarse mucho más:

PROPOSICIÓN 30 *El conjunto $M \subset X$ es convexo si y sólo si contiene todas las combinaciones convexas de vectores de M .*

---

allá de donde debe llegar. Este último conjunto K es un ejemplo de cómo la intuición puede confundir: Si A es un conjunto acotado no vacío en un espacio normado $(X, \|.\|)$, un punto $p \in A$ se dice que es diametral en A si

$$r(p, A) := sup\{\|p - x\| : x \in A\} = \text{diam}(A).$$

Los conjuntos convexos con más de un punto y que constan únicamente de puntos diametrales son algo bastante molesto a la intuición geométrica y podría parecer que no puede haberlos. Sin embargo en el espacio normado $(\mathcal{C}([0,1]), \|.\|_\infty)$, todo punto del anterior conjunto K es diametral en K, ya que si $f \in K$ y f_n es la función de K dada por $f_n(t) = t^n$, se comprueba fácilmente que $\sup\{\|f - f_n\|_\infty : n = 1, 2, ...\} = 1 = \text{diam}(K)$.

Prueba: Si M es vacío o se reduce a un punto, nada hay que probar. En otro caso, desde luego si M contiene todas las combinaciones convexas de sus vectores, en particular contendrá las de todos los pares $\{v, w\} \subset M$, luego M es convexo. Recíprocamente, si M es convexo contiene todas las combinaciones convexas de dos cualesquiera de sus vectores. Supongamos que M contiene las combinaciones convexas de $k-1$ cualesquiera de sus vectores, y sea

$$x = \lambda_1 x_1 + ... + \lambda_k x_k$$

con $x_1, ..., x_k \in M$, $\lambda_i \geq 0$, $\sum_{i=1}^{k} \lambda_i = 1$. Si $\lambda_k = 1$, entonces $x = x_k \in M$. Si $\lambda_k < 1$ sean

$$\mu := \lambda_1 + ... + \lambda_{k-1}, v := \lambda_1 x_1 + ... + \lambda_{k-1} x_{k-1}.$$

Entonces v es una combinación convexa de $k-1$ vectores de M, y por tanto pertenece a M (según estamos suponiendo) . Pero

$$x = \mu v + \lambda_k x_k$$

será un elemento del convexo M porque v, x_k pertenecen a M y $\lambda_k = 1 - \mu$.

Así pues M contendrá también a las combinaciones convexas de k de sus vectores. Por inducción hemos probado la proposición.

□

Esto nos permite ya describir la envoltura convexa de un conjunto cualquiera:

PROPOSICIÓN 31 *Si $M \subset X$, $co(M)$ está formado por las combinaciones convexas de elementos de M.*

Prueba:

Sea C el conjunto de las combinaciones convexas de elementos de M. Cualquier conjunto convexo B conteniendo al M contendrá a todas las combinaciones convexas de elementos de B, y por tanto de elementos de M, luego contendrá al conjunto C.

Por otra parte, si $\lambda \in [0, 1]$ y x, y pertenecen a C, podrán escribirse como combinaciones convexas de elementos de M, por ejemplo

$$x = \sum_{i=1}^{n} a_i x_i, \quad y = \sum_{j=1}^{m} b_j y_j.$$

Pero entonces $\lambda x + (1-\lambda)y$ es una combinación convexa de elementos de M ya que los escalares λa_i , $(1-\lambda)b_j$ son no negativos y suman 1. Luego $\lambda x + (1-\lambda)y \in C$, por lo que C es convexo. Como además, según hemos visto, está contenido en cualquier otro convexo que contenga al M, resulta $C = co(M)$.

□

De forma bastante parecida se demuestra la siguiente propiedad de la envoltura convexa

PROPOSICIÓN 32 *Si $A_1, ..., A_n$ son subconjuntos convexos de un espacio vectorial real X, entonces*

$$co(A_1 \cup ... \cup A_n)$$

es el conjunto de todas las combinaciones convexas de la forma $\lambda_1 a_1 + ...\lambda_n a_n$ *, donde* $a_i \in A_i$*,* $(i = 1, ..., n)$*.* [13] .

5.2. Convexidad en espacios normados

Aunque la convexidad por sí misma es una propiedad de caracter algebraico, tiene una conexión muy estrecha con ciertas propiedades topológicas si se considera en un marco en que haya estructura vectorial y topología. También otros marcos más generales, pero los espacios normados tienen ambas cosas. Suponemos dado, pues un espacio normado real $(X, \|.\|)$. La primera propiedad que enunciamos se demuestra directamente:

PROPOSICIÓN 33 *Si* $A \subset X$ *es convexo, entonces su clausura* $cl(A)$ *es un conjunto convexo.*

Prueba:
Sean $\lambda \in [0, 1]$, $x, y \in cl(A)$. Existen sucesiones $(x_n), (y_n)$ de elementos de A tales que $x_n \to x$, $y_n \to y$. Entonces para cada entero positivo n, $\lambda x_n + (1 - \lambda)y_n \in A$ por ser A convexo. Como

$$\lambda x_n + (1 - \lambda)y_n \to \lambda x + (1 - \lambda)y$$

se tiene que $\lambda x + (1 - \lambda)y \in cl(A)$, luego $cl(A)$ es convexo.

□

La siguiente propiedad puede servir de ejemplo de "buenasïntuiciones geométricas:

PROPOSICIÓN 34 *Sea* a *un punto de un conjunto convexo* A*, y sea* x *un punto interior de* A*. Entonces todos los puntos de* $[x, a[$ *son interiores de* A*.*

Prueba:
Sea $z := a + \lambda(x - a)$, donde $0 < \lambda < 1$, un punto cualquiera de $[x, a[$. Por ser x punto interior de A, existe $r > 0$ tal que $B(x, r) \subset A$. Veremos que $B(z, \lambda r) \subset A$, lo que implica que z es punto interior de A.

Sea pues $y \in B(z, \lambda r)$. Entonces $\|\lambda^{-1}(y - z)\| < r$, de donde se sigue que

$$y' := x + \lambda^{-1}(y - z)$$

pertenece a $B(x, r)$ y por tanto al conjunto A, que es convexo. Tenemos entonces que $\lambda y' + (1 - \lambda)a$ será un punto de A, es decir

$$\lambda(x + \lambda^{-1}(y - z)) + (1 - \lambda)a = \lambda x + (1 - \lambda)a + (y - z) = y$$

es un punto de A. Al ser y arbitrario en $B(z, \lambda z)$ resulta que $B(z, \lambda z) \subset A$.

□

[13] Teorema de Carathéodory Es una propiedad de los conjuntos convexos en espacios vectoriales de dimensión finita, que enunciamos sin demostración y que complementa a las proposiciones anteriores:

TEOREMA 43 *Si* X *es un espacio vectorial real* n*-dimensional, y* A *es un subconjunto de* X*, entonces cada punto de* $co(A)$ *puede ser expresado como combinación convexa de no más de* $n + 1$ *puntos de* A*.*

Proposición 35 *Sean A y B conjuntos compactos convexos de $(X, \|.\|)$. Entonces el conjunto $co(A \cup B)$ es compacto.*

Prueba:
Podemos usar el hecho de que el conjunto $A \times B$ es compacto en el espacio normado producto $X \times X$, y también que $A \times B \times [0,1]$ es compacto en el espacio normado $X \times X \times \mathbb{R}$. Por la proposición (32) el conjunto $co(A \cup B)$ es la imagen de $A \times B \times [0,1]$ por la aplicación $f : A \times B \times [0,1] \to X$ dada por

$$f(a, b, \lambda) := \lambda a + (1 - \lambda) b,$$

que, evidentemente, es continua y por tanto transforma conjuntos compactos en conjuntos compactos.

$\square$

Observación 1 Si (A_n) en una sucesión decreciente de cerrados acotados y convexos de un espacio de Banach, en general no puede garantizarse que

$$\bigcap A_n$$

sea no vacío, como puede verse tomando en c_0 ,

$$A_n := \{x = (x_i) \in c_0 : \|x\| = 1, x_i = 1, i = 1, \ldots, n\}.$$

Proposición 36 *Si A es un conjunto compacto en el espacio de Banach $(X, \|.\|)$, entonces $cl(co(A))$ es compacto.*

Usaremos para probar este resultado de dos lemas, y de una observación previa sobre compacidad en espacios de Banach:

Lema 6 *Si M, N son conjuntos compactos no vacíos de X, $M + N$ es compacto.*

Prueba : Basta considerar $(u_k + v_k)$ una sucesión de vectores en $M + N$, con $u_k \in M$ y con $v_k \in N$ para $k = 1, 2, \ldots$. Por ser M compacto $u_{k_p} \to m \in M$ para cierta subsucesión (u_{k_p}) de (u_k). Pero (v_{k_p}) es una sucesión de elementos del conjunto compacto N, luego $v_{k_{p_j}} \to v \in N$ para cierta subsucesión $(v_{k_{p_j}})$ de (v_{k_p}). Entonces

$$u_{k_{p_j}} + v_{k_{p_j}} \to u + v \in M + N$$

y $(u_{k_{p_j}} + v_{k_{p_j}})$ es subsucesión de $(u_k + v_k)$, luego hemos probado que $M + N$ es compacto.
$\square$

Lema 7 *Si A, B son subconjuntos de X, con A compacto y B cerrado, entonces $A + B$ es cerrado.*

Prueba : Tomemos $(a_k + b_k)$ una sucesión de vectores en $A + B$, con $a_k \in A, b_k \in B$ y con $a_k + b_k \to x \in X$. Como A es compacto entonces $a_{k_p} \to a \in A$ para una subsucesión (a_{k_p}) de (a_k). Entonces, como B es cerrado,

$$b_{k_p} = a_{k_p} + b_{k_p} - a_{k_p} \to x - a \in B$$

de donde $x \in a + B \subset A + B$.

$\square$

Observación 2 Un espacio métrico cualquiera, es compacto si y sólo si es *precompacto* y completo. Esto quiere decir que todas las sucesiones de Cauchy son convergentes (completitud) y que para cada $\varepsilon > 0$ existe un conjunto finito $F = \{x_1, ..., x_n\} \subset X$ tal que las bolas cerradas con centro en algún punto de F y radio ε recubren a X (precompacidad) .

Si C es un cerrado de un espacio de Banach $(X, \|.\|)$, automáticamente C es completo si lo consideramos como espacio métrico con la métrica que hereda de la norma, y por tanto será compacto si y sólo si, para cualquier $\varepsilon > 0$ existe un conjunto finito $F \subset C$ tal que las bolas cerradas de radio ε centradas en puntos de F recubren a C, lo que a su vez es equivalente a decir que $C \subset F + \varepsilon B$.

Prueba de la Proposición.

Dado $\varepsilon > 0$, por ser A compacto podemos encontrar un conjunto finito $F = \{x_1, ..., x_n\} \subset A$ tal que para cualquier punto de $x \in A$ existe algún $x_i \in F$,con $d(x, x_i) \leq \varepsilon/2$. (O lo que es lo mismo,

$$A \subset F + (\varepsilon/2)B_X.$$

Los segmentos $S_i := [0, x_i]$ ($i = 1, ..., n$) son evidentemente compactos y convexos, y por tanto también lo es $S := S_1 + ... + S_n$. Desde luego, $F \subset S$.

Como S es compacto podemos escoger un conjunto finito G con

$$S \subset G + (\varepsilon/2)B_X.$$

Entonces

$$A \subset F + (\varepsilon/2)B_X \subset S + (\varepsilon/2)B_X \subset G + (\varepsilon/2)B_X + (\varepsilon/2)B_X = G + \varepsilon B_X$$

(Notar que $(\varepsilon/2)B_X + (\varepsilon/2)B_X = \varepsilon B_X$ porque B_X es un conjnto convexo). Por otra parte $S + (\varepsilon/2)B_X$ es cerrado y convexo y contiene al conjunto A, luego $cl(co(A)) \subset S + (\varepsilon/2)B_X \subset G + \varepsilon B$, y al ser ε arbitrario concluimos que $cl(co(A))$ es compacto.

$\square$

Ejemplo 43 Toda bola cerrada de $(\mathbb{R}^n, \|.\|_\infty)$ es un conjunto compacto, porque es la (clausura de la) envoltura convexa de sus vértices, que foman un conjunto finito y por tanto compacto.

OBSERVACIÓN 3 Si A es un subconjunto no vacío de X en el espacio normado $(X, \|.\|)$,

$$A \subset co(A) \subset cl(co(A)) \Rightarrow cl(A) \subset cl(co(A)).$$

Como la clausura de un conjunto convexo es un conjunto convexo, se sigue que

$$co(cl(A) \subset cl(co(A)).$$

Por otra parte, si tratamos de probar la inclusión contraria vemos que

$$A \subset cl(A) \subset co(cl(A)) \Rightarrow co(A) \subset co(cl(A)),$$

pero no podemos afirmar que la envoltura convexa de un conjunto cerrado sea un conjunto cerrado.

Ejemplo 44 En el plano euclídeo ordinario sea

$$B := \{(0,0)\} \cup \{(1,y) : y \in \mathbb{R}\}.$$

Este conjunto es cerrado pero

$$co(B) = \{(0,0)\} \cup (0,1] \times \mathbb{R}$$

no es cerrado. Además

$$cl(co(B)) = cl\big(\{(0,0)\} \cup (0,1] \times \mathbb{R}\big) = [0,1] \times \mathbb{R}.$$

mientras que

$$co(cl(B)) = co(B) = \{(0,0)\} \cup (0,1] \times \mathbb{R} \neq cl(co(B)).$$

5.3. Teorema de Hahn-Banach de extensión de formas lineales

Sea E un espacio vectorial sobre $\mathbb{R}$. Recordemos que una *forma lineal* sobre E es una aplicación lineal definida en E y con valores en $\mathbb{R}$. Este teorema, básico en Análisis Funcional, hace referencia a la posibilidad de extender a *todo* el espacio E, una forma lineal definida sobre un subespacio G de E.

DEFINICIÓN 27 *Sea E un espacio vectorial. Una aplicación $p : X \to \mathbb{R}$ tal que*
1) Es sublineal, es decir para cualesquiera $x, y \in X$, $p(x+y) \leq p(x) + p(y)$.
2) Es positivamente homogénea, es decir, para todo $x \in X$ y para todo $\lambda \in R$, $\lambda \geq 0$ se verifica que $p(\lambda x) = \lambda p(x)$,
se llama forma sublineal sobre E.

El mejor ejemplo de forma sublineal sobre un espacio normado $(X, \|.\|)$ es la propia norma, es decir $p(x) := \|x\|$.

Vamos a dar un resultado sencillo que permite extender una forma lineal.

PROPOSICIÓN 37 *Sean E un espacio vectorial real y V uno de sus subespacios.*

Sea $f : V \to \mathbb{R}$ una forma lineal, mayorada en V por el funcional sublineal $p : E \to \mathbb{R}$, es decir verificando para todo $v \in V$, que $f(v) \leq p(v)$. Sea $w_0 \in E \backslash V$. Entonces existe una forma lineal $F : V + \{tw_0 : t \in \mathbb{R}\}$ tal que:
a) F coincide con f en V.
b) F está mayorada por p en el subespacio $V + \{tw_0 : t \in \mathbb{R}\}$.

Prueba:

Sea $W := V + \{tw_0 : t \in \mathbb{R}\}$ es evidente que W es un subespacio vectorial de E que contiene al V. Observemos que cada vector $x \in W$ se esecribe de manera única como $x = v + tw_0$, con $t \in \mathbb{R}$ y $v \in V$. En efecto, si

$$x = v_1 + t_1 w_0 = v_2 + t_2 w_0$$

con $v_1, v_2 \in V$, $t_1, t_2 \in \mathbb{R}$, entonces $v_1 - v_2 = (t_2 - t_1)w_0$. Como $w_0 \notin V$, si fuera $t_2 - t_1 \neq 0$ llegaríamos inmediatamente a una contradicción, luego ha de ser $t_1 = t_2$ lo cual a su vez implica $v_1 = v_2$. Tiene sentido entonces definir $F : W \to \mathbb{R}$ por

$$F(v + tw_0) := f(v) + at$$

donde a es una constante que elegiremos luego. Es directo ver que F es lineal. Si $v \in V$, puede escribirse como $v + 0w_0$, luego $F(v) = f(v)$, es decir que F extiende a f. Comprobaremos finalmente que F está mayorada por p en W, supuesto se elija adecuadamente el número real constante a. Para ello observemos que si $v_1, v_2 \in V$, teniendo en cuenta que p es sublineal y mayora a f en V,

$$f(v_1) + f(v_2) = f(v_1 + v_2) \leq p(v_1 + v_2) = p(v_1 - w_0 + v_2 + w_0) \leq p(v_1 - w_0) + p(v_2 - w_0).$$

Por tanto

$$f(v_1) - p(v_1 - w_0) \leq p(v_2 + w_0) - f(v_2). \tag{19}$$

Existen entonces los siguientes números reales:

$$\beta_1 := \sup\{f(v) - p(v - w_0) : v \in V\}, \quad \beta_2 := \text{ínf}\{p(v + w_0) - f(v) : v \in V\}$$

Además es inmediato que $\beta_1 \leq \beta_2$. Elijamos cualquier constante a con $\beta_1 \leq a \leq \beta_2$. Si $t > 0$ y $v \in V$, teniendo en cuenta cómo se ha elegido a y que p es positivamente homogénea:

$$F(v + tw_0) := f(v) + at \leq f(v) + \beta_2 t \leq f(v) + [p(\frac{1}{t}v + w_0) - f\big(\frac{1}{t}v)]t = p(v + tw_0).$$

En cambio, si $t < 0$, $at \leq \beta_1 t$, y tenemos

$$\begin{aligned}
F(v + tw_0) &:= f(v) + at \\
&\leq f(v) + \beta_1 t \\
&\leq f(v) + [f\big(\tfrac{-1}{t}v\big) - p\big(\tfrac{-1}{t}v - w_0)]t \\
&= -p(\tfrac{-1}{t}v - w_0)t \\
&= p(\tfrac{-1(-t)}{t}v - (-t)w_0) \\
&= p(v + tw_0).
\end{aligned}$$

Por último, si $t = 0$ es obvio que $p(v + 0w_0) = p(v) \geq f(v) = f(v) + 0a = F(v + 0w_0)$.

$\square$

5.3.1. Conjuntos parcialmente ordenados

Recordemos que si M es un conjunto no vacío una relación binaria $\preceq$ sobre M que sea reflexiva, antisimétrica y transitiva se llama relación de *orden parcial*. Si se ha definido una relación de orden parcial sobre M decimos que M está *parcialmente ordenado* por $\preceq$. Si N es un subconjunto no vacío de M, y M está parcialmente ordenado por $\preceq$, naturalmente N también lo está (en realidad por la restricción de $\preceq$ a N).

Supongamos pues que $(M, \preceq)$ es un conjunto (no vacío) parcialmente ordenado, y que N es un subconjunto no vacío de M (no se excluye el caso $N = M$. Decimos que $m_1 \in M$ es una cota superior de N si para todo $n \in N$, se verifica que $n \preceq m_1$. Análogamente, decimos que $m_2 \in M$ es una cota inferior de N si para todo $n \in N$, se verifica que $m_2 \preceq n$. Decimos que $m \in M$ es un elemento maximal de M si de existir $y \in M$ com $m \preceq y$ se sigue $m = y$.

Ejemplo 45 Sea $\mathcal{M} := \mathcal{P}(\mathbb{N})$ el conjunto de los subconjuntos de los números naturales. Sea $\mathcal{P}_3$ el subconjunto de M formado por aquellos de sus elementos que tengan cardinal menor o igual que 3. La relación de inclusión de conjuntos hace de $\mathcal{M}$, (y por tanto de $\mathcal{P}_3$), un conjunto parcialmente ordenado. Cualquier conjunto de tres elementos (mútuamente distintos) $\{p, q, r\} \in P_3$ es un elemento maximal de P_3. (De ser $\{p, q, r\} \subset A \in \mathcal{P}_3$ se sigue que $A = \{p, q, r\}$.) Observar que $\mathcal{P}_3$ no tiene cotas superiores en $\mathcal{M}$, pero $\emptyset \in \mathcal{P}_3 \subset \mathcal{M}$ es una cota inferior de $\mathcal{P}_3$ (y también de $\mathcal{M}$).

TEOREMA 44 *Teorema de Hahn-Banach (Forma analítica).*

Sea $p : E \to \mathbb{R}$ una aplicación que verifica
1) Es positivamente homogénea, es decir: Para todo $x \in E$, y para todo $\lambda > 0$,

$$p(\lambda x) = \lambda p(x).$$

2) Es sublineal, es decir: Para cualesquiera $x, y \in E$,

$$p(x + y) \leq p(x) + p(y).$$

Sean también $G \subset E$ un subespacio vectorial, y $g : G \to \mathbb{R}$ una aplicación lineal tal que está mayorada por p en G, es decir

3) Para todo $x \in G, g(x) \leq p(x)$.

Entonces g se puede extender a una forma lineal $f : E \to \mathbb{R}$, es decir, existe $f \in \mathcal{L}(E, \mathbb{R})$ tal que

a) Para todo $x \in G, g(x) = f(x)$,

y además f está mayorada por p en E, o sea,

b) Para todo $x \in E$, $f(x) \leq p(x)$.

La demostración de este teorema usa el lema de Zörn; recordamos su enunciado:

Sea P un conjunto dotado de una relación de orden (parcial) que simbolizaremos con $\leq$. El subconjunto $Q \subset P$ se dice que esta *totalmente ordenado* si para cualesquiera a, b de Q, se verifica $a \leq b$ o bien $b \leq a$. Un elemento $c \in P$ se dice que es *cota superior* de Q

si para todo q de Q, se tiene que $q \leq c$. Un elemento m de P es *maximal* en P (respecto de $\leq$) si de ser $m \leq x$ para $x \in P$, necesariamente se sigue $m = x$. Finalmente se dice que P es *inductivo* si todo subconjunto de totalmente ordenado de P tiene alguna cota superior en P. Con este vocabulario, el **Lema de Zörn** establece que

Todo conjunto ordenado, inductivo y *no vacío admite un elemento maximal.*

Prueba del Teorema:

Si $G = E$ nada hay que demostrar. Supongamos que $G \neq$ E. y sea $x_0 \in E \backslash G$. Entonces

$$G_1 := G + \{tx_0 : t \in \mathbb{R}\}$$

es un subespacio vectorial de E que contiene (estrictamente) a G, y podemos definir $g_1 : G_1 \to \mathbb{R}$, por

$$g_1(x + tx_0) := g(x) + t\alpha$$

donde α es una constante que se fijará posteriormente.

Si $x_1 + t_1x_0 = x_2 + t_2x_0$ para $x_1, x_2 \in G$, entonces $x_1 - x_2 = (t_2 - t_1)x_0$, lo cual sólo es posible si $(t_2 - t_1) = 0$, y entonces $x_1 = x_2$. Asi pues, la forma lineal g_1 está bien definida. Además si $x \in G, g_1(x) = g_1(x + 0x_0) = g(x)$, luego g_1 extiende a g.

Como p es sublineal, si $x, y \in G$,

$$g(x) + g(y) = g(x + y) \leq p(x + y) \leq p(x - x_0) + p(x_0 + y),$$

de donde, para $x, y \in G$

$$g(x) - p(x - x_0) \leq p(x_0 + y) - g(y)$$

luego

$$g(x) - p(x - x_0) \leq \text{ínf}\{p(x_0 + y) - g(y) : y \in G\}$$

y de aquí que

$$\sup\{g(x) - p(x - x_0) : x \in G\} \leq \text{ínf}\{p(x_0 + y) - g(y) : y \in G\}$$

por lo que podemos elegir α en $\mathbb{R}$ con

$$\sup\{g(x) - p(x - x_0) : x \in G\} \leq \alpha \leq \text{ínf}\{p(x_0 + y) - g(y) : y \in G\}.$$

Entonces, si $\lambda > 0$, para todo $x \in G$,

$$g_1(x + \lambda x_0) = g(x) + \lambda\alpha \leq g(x) + \lambda[p(x_0 + \lambda^{-1}x) - g(\lambda^{-1}x)] = p(x + \lambda x_0).$$

Del mismo modo, si $\lambda < 0$, para todo $x \in G$

$$g_1(x + \lambda x_0) = g(x) + \lambda\alpha \leq g(x) + \lambda[g(-\lambda^{-1}x) - p(-\lambda^{-1}x - x_0)] = p(x + \lambda x_0).$$

Vemos pues que, por la elección hecha de α, para todo $t \in \mathbb{R}$,

$$g_1(x + tx_0) \leq p(x + tx_0),$$

luego g_1 es una extensión de g mayorada por p en su dominio.

Si $G_1 = E$ la demostración ha terminado. Si $G_1 \neq E$ podemos repetir el proceso anterior partiendo de G_1 y de g_1, y obtendríamos un subespacio G_2 de E con $G_2 \supset G_1, y$ una forma lienal g_2 sobre G_2 que prolonga a g_1 y que está mayorada por p en su dominio.

Continuando de este modo, si en un número finito de pasos el subespacio vectorial $G_n = E$ la demostración ha concluido. (Esto ocurre, por ejemplo, cuando E es de dimensión finita).

Caso contrario se considera el conjunto $\mathcal{P}$ de las formas lineales definidas en algún subespacio de E, que extiendan a g, y que estén mayoradas por p en todo su dominio. $\mathcal{P}$ no es vacío porque $g \in \mathcal{P}$.

En este conjunto definimos la siguiente relación de orden $\preccurlyeq$:
Si $h_1 : D_1 \to \mathbb{R}, h_2 : D_2 \to \mathbb{R}$, son elementos de $\mathcal{P}$, diremos que $h_1 \preccurlyeq h_2$ si y sólo si

$$D_1 \subset D_2 \ \text{ y } \ h_2 \ \text{ extiende a } \ h_1.$$

Vamos a ver que $\mathcal{P}$ es inductivo:

Sea $\mathcal{Q} \subset \mathcal{P}$ un subconjunto totalmente ordenado. Podremos escribir

$$\mathcal{Q} := \{h_i : i \in I\}$$

donde h_i es un forma lineal sobre cierto subespacio D_i de E, en el cual es mayorada por la función p.

Entonces $D := \bigcup_{i \in I} D_i$, es un subespacio vectorial de E ya que si $x, y \in D, y\lambda \in \mathbb{R}$, entonces $x \in D_i$, $y \in D_j$ para i, j en I. Como $\mathcal{Q}$ está totalmente ordenado, se cumple que $h_i \preccurlyeq h_j$ o que $h_j \preccurlyeq h_i$. En caso de ser $h_i \preccurlyeq h_j, D_i \subset D_j$, luego $x + y, \lambda x \in D_j \subset D$; y caso de ser $h_j \preccurlyeq h_i$ tendríamos igualmente que $x + y, \lambda x$ pertenecen a $D_i \subset D$.

La aplicación $h : D \to \mathbb{R}$ por $h(x) := h_i(x)$ si $x \in D_i$ (para algún $i \in I$) está bien definida, ya que si además $x \in D_j$ para $j \in I$, nuevamente por ser $\mathcal{Q}$ totalmente ordenado tenemos que $h_i(x) = h_j(x)$ (pues o bien h_i prolonga a h_j o bien h_j prolonga a h_i). Además, es claro que h es lineal por serlo las aplicaciones $h_i en D_i$ (para cada $i \in I$). Por último tambiń es evidente que $h(x) \leq p(x)$ para cada $x \in D$, (por serlo cada $h_i(x)$)), luego $h \in \mathcal{P}$.

Por último podemos ver que h es una cota superior de $\mathcal{Q}$, ya que, efectivamente, si $h_i \in \mathcal{Q}$, h extiende a h_i pues $D \supset D_i$ y si $x \in D_i$, $h(x) = h_i(x)$.

Podemos pues aplicar el Lema de Zörn, y de él resulta que $\mathcal{P}$ tendrá un elemento maximal, digamos f. Sea D_f el dominio de f. Si vemos que $D_f = E$, habremos completado la prueba.

Si $D_f \neq E$, aplicando el mismo proceso detallado al principio, partiendo del subespacio D_f y de la forma lineal f, y obtendríamos una prolongación de f a un subespacio (estrictamente) más amplio y mayorada por p en su dominio, lo cual contradice la maximalidad de $f en \mathcal{P}$ respecto de la relación $\preccurlyeq$. Así pues, $D_f = E$, y el teorema está demostrado.

$\square$

Es importante notar que la extensión f de la forma lineal g *no* es *única,* ni aun en el caso de ser E de dimensión finita.

Daremos a continuación algunas consecuencias sencillas de este teorema, para el caso particular de ser E un espacio vectorial real en el que se haya definido una norma.

Sea pues $(X, \|.\|)$ un espacio normado real, y X^* su dual topológico, (es decir el espacio vectorial de las formas lineales-continuas sobre X, dotado con la norma dual. Cuando $f \in X^*$ y $x \in X$ se escribe a veces $\langle x, f \rangle$ en lugar de $f(x)$.

Corolario 15 *Sea G un subespacio vectorial de X, y sea $g \in G^*$. Entonces existe $f \in X^*$ tal que f extiende a g, y con*

$$\|g\|_{G^*} = \|f\|_{X^*}.$$

Prueba : Si $G = \{0\}$ el resultado es evidente. Si $G \neq \{0\}$, basta aplicar el Teorema con

$$p(x) := \|g\|_{G*}\|x\|$$

que es sublineal y positivamente homogénea. Obtendremos entonces una forma lineal $f : X \to \mathbb{R}$ que extiende a g y que está mayorada por p, es decir con

$$f(x) \leq p(x) = \|g\|_{G*}\|x\|$$

para todo x de X. Por tanto,

$$f(-x) \leq \|g\|_{G*}\| - x\|,$$

y de aquí resulta inmediatamente que, para todo $x \in X$

$$|f(x)| \leq \|g\|_{G*}\|x\|,$$

lo que directamente prueba que f es continua y que $\|f\|_{X*} \leq \|g\|_{G*}$.
Siendo $G \neq \{0\}$ existirá algún vector $v \in G$ con $\|v\| = 1$, y entonces

$$\|f\|_{X*} \geq |f(v)| = |g(v)|.$$

luego (dado que el razonamiento se cumple para todo vector de G con norma 1),

$$\|g\|_{G^*} = \sup\{|g(v)| : \|v\| = 1\} \leq \|f\|_{X*},$$

desigualdad que completa la prueba.

$\square$

Como el resultado de este corolario es en sí mismo de gran importancia, lo extenderemos al caso de espacios vectoriales complejos. Recordemos que si E es un espacio vectorial complejo, siempre podemos considerarlo como espacio vectorial real:

Si $f : E \to \mathbb{C}$ es una forma lineal, $f(\lambda x) = \lambda f(x)$ para todos los escalares complejos, y en particular reales. Sin embargo, por ser $f : E \to \mathbb{C}$ y cumplir $f(\alpha x) = \alpha f(x)$ para escalares $\alpha \in \mathbb{R}$, no se sigue que f verifique lo mismo para escalares complejos. Luego hay que distinguir entre $\mathbb{C}$-linealidad y $\mathbb{R}$-linealidad.

Si $(X, \|.\|)$ es un espacio normado complejo, y X_r es el espacio vectorial real subyacente, si $f \in X^*$ entonces su parte real

$$(\Re f)(x) := \Re(f(x))$$

es un elemento de X_r^*.

Más sorprendentemente, si $f_r \in X_r^*$, entonces $f : X \to \mathbb{R}$ definida por

$$f(x) := f_r(x) + i f_r(-ix)$$

es un elemento de X^*, es decir es $\mathbb{C}$-lineal, continua y tiene f_r como parte real.

$$f(ix) = f_r(ix) + i f_r(x) = i(f_r(x) - i f_r(ix)) = i(f_r(x) + i f_r(-ix)) = i f(x).$$

Además, $\|f_r\|_{X_r^*} = \|f\|_{X^*}$: Claramente

$$|f_r(x)| \leq |f(x)| \leq \|f\|_{X^*}\|x\|,$$

luego

$$\|f_r\|_{X_r^*} \leq \|f\|_{X^*}.$$

Por otra parte, dado $x \in X$, $f(x)$ es un número complejo, y podemos encontrar un complejo α con $|\alpha| = 1$ y tal que $\alpha f(x) = |f(x)|$.

Entonces

$$|f(x)| = f(\alpha x) = \Re(f(\alpha x)) = f_r(\alpha x) \leq \|f_r\|_{X_r^*}\|\alpha x\| = \|f_r\|_{X_r^*}\|x\|,$$

luego

$$\|f\|_{X^*} \leq \|f_r\|_{X_r^*}.$$

Esta misma argumentación sirve para probar que $\|\Re(f)\|_{X_r^*} = \|f\|_{X^*}$ para todo elemento de X^*.

Sean pues G, X como en el enunciado del Corolario 1, pero ahora $(X, \|.\|)$ un espacio normado complejo. Sea $g : G \to \mathbb{R}$ una forma $\mathbb{B}$-lineal sobre G. Aplicamos el Corolario 1 al espacio normado real X_r, al subespacio G_r y a la forma $\mathbb{R}$-lineal $\Re(g)$ sobre G_r.

Via Corolario 1 obtenemos una forma lineal f_r sobre X_r tal que extiende a $(\Re g)$ y tal que $\|f_r\|_{X_r^*} = \|\Re g\|_{G_r^*}$.

Según hemos visto antes

$$f(x) = f_r(x) + i f_r(-ix)$$

será una forma $\mathbb{C}$-lineal sobre X, con

$$\|f\|_{X^*} = \|f_r\|_{X_r^*} = \|\Re g\|_{G_r^*} = \|g\|_{G^*}.$$

Además, si $x \in G$,

$$f(x) = f_r(x) + i f_r(-ix) = (\Re g)(x) + i(\Re g)(-ix) =$$

$$= \Re(g(x)) + i(\Re(-ig(x))) = g(x)$$

(pues si z es un número complejo, $z = \Re(z) + i\Re(-iz)$). Luego f extiende a g.

$\square$

Corolario 16 *En las condiciones del corolario anterior, para cada $x_0 \in X$ existe $f_0 \in X^*$ tal que*

$$\|f_0\| = \|x_0\|, f_0(x_0) = \|x_0\|^2.$$

Prueba:

Prueba : Si $x_0 = 0$ el resultado es evidente. Caso contrario, tomemos el subespacio vectorial G de X

$$G := \mathbb{R}x_0 = \{\lambda x_0 : \lambda \in \mathbb{R}\},$$

y la forma lineal g sobre G definida por

$$g(tx_0) := t\|x_0\|^2.$$

De la definición se sigue directamente que

$$\|g\|_{G^*} := \sup\{|g(v)| : v \in G, \|v\| = 1\}$$

y como los únicos vectores de G con norma 1 son de la forma tx_0 con

$$|t|\|x_0\| = 1$$

se sigue que en este caso $|g(v)| = |t\|x_0\|^2| = \frac{1}{\|x_0\|}\|x_0\|^2 = \|x_0\|$, luego

$$\|g\|_{G^*} = \|x_0\|.$$

Aplicando el corolario anterior tenemos que existe $f_0 \in X^*$ tal que

$$\|f_0\|_{x^*} = \|x_0\|,$$

y además

$$f_0(x_0) = g(x_0) = \|x_0\|^2,$$

lo que completa la prueba.

□

El elemento f_0 definido en este corolario no es único en general, aunque hay casos particulares importantes en que sí puede garantizarse su unicidad, como, por ejemplo, si $(X, \|.\|)$ es un espacio de Hilbert.

El siguiente corolario debe compararse con la definión de norma (dual) de un elemento de X^*:

$$\|f\| := \sup\{|f(x)| : x \in X, \|x\| \leq 1\}.$$

Sabemos que se trata de una definición y no de un resultado, y sabemos también que en general no puede garantizarse que el supremo sea un máximo. Sin embargo

Corolario 17 *Para todo $x \in X$ se tiene:*

$$\|x\| = \sup\{|f(x)| : f \in X^*, \|f\| \leq 1\} = \text{máx}\{|f(x)| : f \in X^*, \|f\| \leq 1\}.$$

Prueba : Si $x = 0_X$ el resultado *es* evidente. Supongamos $x \neq 0_X$. Claramente se tiene

$$|f(x)| \leq \|f\|\|x\|,$$

luego

$$\sup\{|f(x)| : f \in X^*, \|f\| \leq 1\} \leq \|x\|.$$

Por el corolario anterior se tiene que existe $f \in X^*$ tal que $\|f\| = \|x\|$ y con

$$f(x) = \|x\|^2.$$

Entonces $h := \|x\|^{-1} f \in X^*$ y tiene norma 1, por lo que

$$\|x\| = |h(x)| \leq \sup\{|f(x)| : f \in X^*, \|f\| \leq 1\} \leq \|x\|,$$

y la demostración está completa.

□

Otra consecuencia importante del Teorema de Hahn Banach es que en cualquier espacio dual existe *suficiente cantidad de elementos como para separar puntos de* X, es decir:

COROLARIO 18 *Dados dos vectores u, v del espacio normado $(X, \|.\|)$, existe $f \in X^*$ tal que $f(u) \neq f(v)$.*

Prueba : Basta aplicar el corolario 24 al elemento $x_0 := u - v \neq 0$. Existirá $f \in X^*$ tal que

$$f(x_0) = \|x_0\| \neq 0,$$

o sea

$$f(u) - f(v) \neq 0.$$

□

5.4. Formas geométricas del Teorema de Hahn-Banach

DEFINICIÓN 28 *Un hiperplano vectorial, en un espacio vectorial E, es un subespacio vectorial propio y maximal. Es decir, un subespacio $H \subset E$ es un hiperplano vectorial en E si $H \neq E$ y de ser F un subespacio de E con $H \subset F \subset E$ se sigue $F = H$ ó F=E.*

PROPOSICIÓN 38 *Las siguientes propiedades de un subespacio M de E son equivalentes:*

a)M es un hiperplano vectorial.
b) El espacio vectorial cociente X/M es de dimensión 1.
c) M es el núcleo de alguna forma lineal f no nula sobre E.

Prueba : ((a)⇒(b))

Si M es un hiperplano vectorial y fijamos $e \in E \backslash H$ entonces por definición de hiperplano, $\mathbb{R}e + M$ es un subespacio de E que contiene al M y es *distinto de* M, luego $E = \mathbb{R}e + M$. En otras palabras, cada elemento x de E puede ser escrito en la forma $x = \lambda e + m$, donde $m \in M$. Luego $x - \lambda e \in M$, es decir, en el espacio vectorial cociente,

$$x + M = \lambda e + M = \lambda(e + M).$$

Por tanto, y dado que $e + M \neq M$, obtenemos que X/M es de dimensión uno.

((b)⇒ (c))

Supongamos que $\dim(E/M) = 1$, y sea $\{e + M\}$ una base de E/M. Todo elemento $x + M$ de X/M podrá expresarse en forma única como

$$x + M = \lambda(e + M),$$

y entonces la aplicación $f : E \to \mathbb{K}$ dada por $f(x) = \lambda$ está bien definida y es una forma lineal no nula sobre E, cuyo núcleo es M.

((c)⇒ (a))

Por último, sea $M = \ker(f)$ donde f es una forma lineal no nula sobre E. Evidentemente, por ser f no nula, $M \neq E$.

Supongamos que F es un subespacio de E que contiene al M, y con $F \neq M$. Existe entonces algún vector $xenF \backslash M$. Sea $y \in E$:

$$f(y - \frac{f(y)}{f(x)}x) = 0$$

luego

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

es decir

$$y \in \mathbb{R}x + M \subset F.$$

Como el razonamiento es válido para todo vector de E, concluimos que $F = E$, es decir que M es un hiperplano vectorial.

□

Definición 29 *Se llama hiperplano afín en un espacio vectorial, al trasladado de cualquier hiperplano vectorial, es decir, a todo conjunto de la forma $w + H$ con $w \in E$ y H hiperplano vectorial en E.*

Proposición 39 *Un conjunto $M \subset E$ es un hiperplano afín si y sólo si existen una forma lineal f no nula sobre E, y un escalar α tales que*

$$M = \{x \in E : f(x) = \alpha\}.$$

Prueba : Supongamos que $M = w + H$ con H hiperplano vectorial en E. Entonces $H = \ker(f)$, siendo f una forma lineal no nula sobre E. Si $w \in H$, entonces $M = H$, y basta tomar $\alpha = 0$. Si $w \notin E$, entonces tomamos $\alpha = f(w)$, pues

$$\{x \in E : f(x) = f(w)\} = \{x \in E : x - w \in \ker(f)\} = w + \ker(f).$$

Recíprocamente, si para alguna forma lineal no nula f sobre E y para algún escalar α

$$M = \{x \in E : f(x) = \alpha\},$$

entonces, como fes no nula existe $y \in E$ con $f(y) = 1$, por lo que

$$M = \{x \in E : f(x) = f(\alpha y)\} = \{x \in E : x - \alpha y \in \ker(f)\} = \alpha y + \ker(f),$$

pero $\ker(f)$ es un hiperplano vectorial.

$\square$

Esta última proposición justifica la costumbre que tiene mucha gente de no distinguir entre hiperplanos vectoriales e hiperplanos afínes, usando para ambas clases de conjuntos el término 'hiperplano'.

PROPOSICIÓN 40 *Sea* $(X, \|\cdot\|)$ *un espacio normado, y sea* f *una forma lineal sobre* X*. El hiperplano afín* $\{x \in X : f(x) = \alpha\}$ *es cerrado si y sólo si* f *es continua.*

Prueba : Sea $H := \{x \in X : f(x) = \alpha\}$. Si f es continua, evidentemente H es cerrado.

Recíprocamente, supongamos que H es cerrado. Entonces su complementario es abierto y no vacío, (ya que $f \neq 0$). Sea $v \in X \backslash H$, y supongamos, por ejemplo, que $f(v) < \alpha$. Por ser $X \backslash H$ abierto existirá $r > 0$ tal que $B[v, r] \subset X \backslash H$. Afirmamos que para todo $x \in B[v, r], f(x) < \alpha$.

En efecto, si para algún $y \in B[v, r]$ se tuviera que $f(y) \geq \alpha$, como el segmento

$$\{tv + (1 - t)y : t \in [0, 1]\}$$

está contenido en $B[v, r]$, ha de verificarse que $f(ty + (1 - t)v) \neq \alpha$ para todo $t \in [0, 1]$, pero si tomamos

$$t := \frac{f(y) - \alpha}{f(y) - f(v)}$$

sucede que $t \in [0, 1]$, y

$$f\left[\frac{f(y) - \alpha}{f(y) - f(v)} v + \big(1 - \frac{f(y) - \alpha}{f(y) - f(v)}\big) y\right] = \alpha$$

lo cual es absurdo.

Asi pues, para todo punto $y \in B[v, r]$ ha de tenerse que $f(y) < \alpha$, o, lo que es lo mismo, para todo $z \in B[0, 1]$

$$f(v + rz) < \alpha$$

es decir

$$f(z) < \frac{\alpha - f(v)}{r}.$$

lo que prueba que f es acotada en $B[0,1]$, y por tanto continua, con

$$\|f\| \leq \frac{\alpha - f(v)}{r}.$$

De modo parecido razonaríamos que si $f(v) > \alpha$ existe una bola cerrada $B[v,r]$ para la cual se tiene que $f(y) > \alpha$ si $y \in B[v,r]$. O, equivalentemente, si $z \in B[0,1]$

$$f(v + rz) > \alpha$$

es decir

$$f(-z) < (f(v) - \alpha)/r,$$

lo que también prueba que f es acotada en la bola unidad $B[0,1]$ y por tanto continua, teniéndose en este caso

$$\|f\| \leq \frac{(f(v) - \alpha)}{r}.$$

□

Sea E un espacio vectorial, y sean A y B dos subconjuntos de E. Sea f una forma lineal no nula sobre E. El hiperplano afín $H := \{x \in E : f(x) = \alpha\}$ *separa* los conjuntos A y B **en sentido amplio**, si

$$\forall x \in A, f(x) \leq \alpha \;\text{ y }\; \forall x \in B, f(x) \geq \alpha.$$

Se dice que H separa a los conjuntos A y B **en sentido estricto**, si existe $\epsilon > 0$ tal que

$$\forall a \in A, \forall b \in B, f(a) \leq \alpha - \varepsilon < \alpha + \varepsilon \leq f(b)$$

Ejemplo 46 En el plano euclídeo a los conjuntos abiertos convexos

$$A := \{(x,y) : xy > 1, x > 0, y > 0\}$$

$$B := \{(x,y) : y < 0\}$$

los separa en sentido amplio el hiperplano $H := \{(x,y) : y = 0\}$, pero no pueden ser separados en sentido estricto por hiperplano afín alguno.

Veremos a continuación cómo, utilizando el Teorema de Hahn Banach en su forma analítica, pueden obtenerse algunos resultados que garantizan la posibilidad de separar conjuntos convexos mediante hiperplanos.

Teorema 45 *Teorema de Hahn-Banach (Primera forma geométrica).*

Sea $(X, \|\cdot\|)$ un espacio normado, y sean A, B subconjuntos convexos, no vacíos y disjuntos de X. Supongamos además que A es abierto.

Entonces existe un hiperplano afín cerrado H que separa $A y B$ en sentido amplio.

La prueba se basa en los dos siguientes lemas, que, por sí mismos, tienen importancia. El primero de ellos describe una función real que puede definirse en cualquier espacio normado partiendo de un abierto convexo que contenga al origen. El segundo lema es, en realidad, un teorema de separación entre puntos (conjuntos con un solo elemento), y ciertos conjuntos convexos.

Lema 8 *(Funcional de Minkowski de un abierto convexo). Sea C un abierto convexo de $(X, \|\cdot\|)$, con $0 \in C$. Para todo $x \in X$ se define*

$$p(x) := \text{ínf}\{\alpha > 0 : \alpha^{-1}x \in C\}.$$

Entonces p es una función real sobre X, subaditiva y positivamente homogénea. Además existe $M \in \mathbb{R}$ tal que, para todo $x \in X$,

$$0 \leq p(x) \leq M\|x\|,$$

y, finalmente

$$C = \{x \in X : p(x) < 1\}.$$

La función p que este lema define se llama *Funcional de Minkowski de C* ó calibrador de C.

Prueba : Como C es un abierto que contiene al origen, existirá $r > 0$ tal que $B[0, r] \subset C$, luego para todo $x \in X, x \neq 0$, se tiene que

$$(r/\|x\|)x \in B[0, r] \subset C,$$

lo que prueba que $p(x)$ es un número real no negativo, y además que si tomamos $M := 1/r$,

$$p(x) \leq M\|x\|.$$

Por otra parte, si $x \in C$, dado que Ces abierto, para $\varepsilon > 0$ suficientemente pequeño, $(1 + \varepsilon)x \in C$, luego $p(x) \leq \frac{1}{1+\epsilon} < 1$.

Inversamente, si $p(x) < 1$ y tomamos β con $p(x) < \beta < 1$, existirá $\alpha > 0$ tal que $\alpha^{-1}x \in C$, con $p(x) \leq \alpha < \beta$, pero entonces por ser C un convexo que contiene al origen

$$x = \alpha(\alpha^{-1}x) + (1 - \alpha)0 \in C.$$

Luego hemos probado que $C := \{x \in X : p(x) < 1\}$. Si $\lambda > 0$, y $x \in X$

$$p(\lambda x) := \text{ínf}\{\alpha : \alpha^{-1}\lambda x \in C\} = \text{ínf}\{\lambda(\alpha\lambda^{-1}) : (\alpha\lambda^{-1})^{-1}x \in C\} = \lambda p(x).$$

Por último veremos que p es subaditiva . Sean x, y elementos de X, y $\varepsilon > 0$ arbitrario. Tendremos que

$$p\left(\frac{x}{p(x) + \varepsilon}\right) < 1, \; p\left(\frac{y}{p(y) + \varepsilon}\right) < 1$$

luego

$$\frac{1}{p(x) + \varepsilon}x \in C, \quad \frac{1}{p(y) + \varepsilon}y \in C.$$

Como C es convexo y $0_x \in C$, para todo $t \in [0,1]$

$$t\frac{1}{p(x)+\varepsilon}x + (1-t)\frac{1}{p(y)+\varepsilon}y \in C.$$

y tomando en particular

$$t := \frac{p(x)+\varepsilon}{p(x)+p(y)+2\varepsilon}$$

se obtiene

$$\frac{1}{p(x)+p(y)+2\varepsilon}x + \frac{1}{p(x)+p(y)+\epsilon}y \in C.$$

por lo que

$$p\left(\frac{1}{p(x)+p(y)+2\varepsilon}x + \frac{1}{p(x)+p(y)+\epsilon}y\right) < 1,$$

es decir

$$p(x+y) < p(x)+p(y)+2\varepsilon,$$

y siendo ε arbitrario tenemos finalmente

$$p(x+y) \leq p(x)+p(y),$$

completándose así la prueba de este lema.

$\square$

Lema 9 *(Separación entre puntos y abiertos convexos).*

Sea $(X, \|\cdot\|)$ un espacio normado real. Sea C un conjunto convexo abierto de $(X, \|\cdot\|)$, y sea x_0 un punto de X con $x_0 \notin C$. Entonces existe $f \in X^*$, tal que para todo $x \in C, f(x) < f(x_0)$. En particular el hiperplano

$$\{x \in X : f(x) = f(x_0)\}$$

separa $\{x_0\}$ de C en sentido amplio.

Prueba : Supongamos que $0 \in C$. Sea p el funcional de Minkowski de C. Sean G el subespacio generado por $\{x_0\}$, y $g : G \to \mathbb{R}$ la forma lineal

$$g(tx_0) := t.$$

Es claro que, si $t > 0$, como $x_0 \notin C$

$$g(tx_0) = t < tp(x_0) = p(tx_0),$$

mientras que si $t \leq 0$,

$$g(tx_0) = t \leq -tp(-x_0) = p(tx_0),$$

luego p mayora a g en G. Gracias al Teorema de Hahn Banach en forma analítica existe una forma lineal $f : X \to \mathbb{R}$, mayorada por $penX$, y tal que f extiende a g.

Por las propiedades de p vistas en el lema anterior existirá $M > 0$ tal que, para todo $x \in X$
$$f(x) \leq p(x) \leq M\|x\|,$$
luego
$$|f(x)| = \text{máx}\{f(x), f(-x)\} \leq M\|x\|,$$
y f será continua.

Además, para cada $x \in C$,
$$f(x_0) = g(x_0) = 1 > p(x) \geq f(x).$$
Si C no contiene al origen, para algún $v \in X$ tendremos que $0_X \in v + C$, mientras que $v + x_0 \notin v + C$.

Aplicando el caso particular ya demostrado, existirá una forma lineal continua $f \in X^*$ tal que, para todo $x \in C$,
$$f(v + x_0) > f(v + x)$$
y de la linealidad de f se sigue que también
$$f(x_0) > f(x),$$
y la prueba del lema está completa.

$\square$

Prueba del Teorema:

Sea
$$C := A - B = \bigcup_{y \in B} (A - y).$$
Este conjunto es convexo, pues si x, y son elementos de C y $t \in [0, 1]$, podremos escribir, $x = a - b$, $y = a' - b'$ con $a, a' \in A$, $b, b' \in B$, de donde, por ser A y B convexos,
$$tx + (1-t)y = t(a-b) + (1-t)(a'-b') = ta + (1-t)a' - tb + (1-t)b' \in A - B.$$
Por otra parte $A - B$ es abierto por ser unión de abiertos, y no contiene al origen porque A y B son disjuntos.

Según el lema anterior, existe $f \in X^*$ tal que para todo $z \in C$,
$$f(z) < f(0_X) = 0,$$
luego si $a \in A$ y $b \in B$,
$$f(a - b) < 0 \Leftrightarrow f(a) < f(b).$$
Luego existe $\alpha \in \mathbb{R}$ con
$$\sup\{f(x) : x \in A\} \leq \alpha \leq \text{ínf}\{f(y) : y \in B\},$$
y el hiperplano afín
$$H := \{x \in X : f(x) = \alpha\}$$
separa A y B en sentido amplio.

$\square$

Teorema 46 *Segunda forma geométrica del Teorema de Hahn-Banach.*

Sean A, B conjuntos no vacíos y disjuntos del espacio normado real $(X, \|\cdot\|)$. Si A es cerrado convexo y B es compacto y convexo, entonces existe un hiperplano afín cerrado H que separa A y B en sentido estricto.

Prueba : Fijemos $\varepsilon > 0$ y consideremos los conjuntos

$$A_\varepsilon := A + B(0, \varepsilon), B_\varepsilon := B + B(0, \varepsilon).$$

Tanto A_ϵ como B_ϵ son conjuntos suma de conjuntos convexos y no vacíos, y por tanto son ellos mismos convexos y no vacíos. Comprobaremos que son abiertos sólo con A_ϵ, ya que para $B_\epsilon se$ pueden repetir al pie de la letra los mismos argumentos.

Sea pues $v \in A_\varepsilon$. Podrá escribirse $v = a + b$, con $a \in A, b \in B(0, \varepsilon), yen$ tal caso $\|b\| < \varepsilon$, por lo que existirá $\rho > 0$ con $\|b\| + \rho < \varepsilon$. Se tendrá que $B(b, \rho) \subset B(0, \varepsilon)$ pues para todo $z \in B(b, \rho)$

$$\|z\| \leq \|z - b\| + \|b\| < \rho + \|b\| < \varepsilon.$$

Esto demuestra que v es punto interior de A_ε pues

$$v = a + b \in a + B(b, \rho) \subset A + B(b, \rho) \subset A_\varepsilon$$

y el conjunto $a + B(b, \rho)$ es abierto.

Afirmamos que, para $\varepsilon > 0$ suficientemente pequeño, $A_\epsilon \cap B_\epsilon = \emptyset$. Caso contrario existe un natural n_0 tal que para todo $n \geq n_0$

$$A_{1/n} \cap B_{1/n} \neq \emptyset,$$

es decir, existen sucesiones (v_n) tales que, para $n \geq n_0$ se puede escribir

$$v_n = a_n + b_n = w_n + b'_n$$

con $a_n \in A$, $w_n \in B$, $b_n, b'_n \in B(0, 1/n)$. Como B es compacto, pasando a subsucesiones si fuera necesario, podemos suponer que $w_n \to w_0 \in B, y$ entonces como

$$\|w_n - a_n\| = \|b_n - b'_n\| \leq \frac{2}{n},$$

necesariamente debe tenerse que también $a_n \to w_0$, y siendo A cerrado, $w_0 \in A$. Entonces $w_0 \in A \cap B$, lo que va contra la hipótesis de ser A y B disjuntos. Por tanto hemos probado la afirmación de que los conjuntos A_ϵ y B_ϵ son disjuntos para ϵ suficientemente pequeño.

Como consecuencia de la primera forma geométrica del Teorema de Hahn Banach tenemos que existe un hiperplano afín cerrado $H := \{x \in X : f(x) = \alpha\}$, con $f \in X^*, f \neq 0$, tal que separa a A_ϵ y B_ϵ en sentido amplio. O sea que, para cualesquiera $x \in A, y \in B, z \in B(0, 1)$

$$f(x + \varepsilon z) \leq \alpha \leq f(y - \epsilon z) \leq f(y) - \varepsilon\|f\|$$

de donde resulta que, siendo z arbitrario en $B(0_X, 1)$

$$f(x) + \varepsilon\|f\| \leq \alpha \leq f(y) - \varepsilon\|f\|,$$

de donde se concluye que A y B están separados en sentido estricto por el hiperplano H, y la demostración está completa.

$\square$

Se puede construir un ejemplo de un espacio normado conteniendo dos conjuntos no vacíos, cerrados y disjuntos que no pueden ser separados en sentido amplio con un hiperplano cerrado. Sin embargo, en espacios normados de dimensión finita siempre se pueden separar en sentido amplio dos conjuntos convexos no vacíos y disjuntos, sin más hipótesis.

COROLARIO 19 *t Si F es un subespacio vectorial en el espacio normado $(X, \|.\|)$ tal que* $\mathrm{cl}(F) \neq X$*, entonces existe $f \in X^*$, con $f \neq 0$, tal que $f(x) = 0$ para todo $x \in F$.*

Prueba : Sea $v \in X \backslash cl(F).El$ conjunto $B := \{v\}$ es no vacío compacto y convexo, y el conjunto $A := cl(F)$ es cerrado no vacío y convexo. Luego existe un elemento f del dual (topológico) de X y un escalar α, tal que el hiperplano $\{x \in X : f(x) = \alpha\}$ separa A y B en sentido estricto. Luego, para todo $x \in cl(F)$

$$f(x) < \alpha < f(v),$$

de donde resulta que, en particular, $f(\lambda x) < \alpha$ para todo $\lambda \in \mathbb{R}$, lo que sólo es posible si $f(x) = 0$.

5.5. Teoremas de la aplicación abierta y de la gráfica cerrada

Empezaremos por recordar algunos conceptos, que se suponen conocidos de la teoría de espacios métricos.

5.5.1. Conjuntos diseminados y cateorías

DEFINICIÓN 30 *Sea (X, d) un espacio métrico. Un subconjunto A de X se dice*

- *Diseminado (nowere dense) en X si $int(cl(A) = \emptyset$.*
- *De primera categoría en X si es unión numerable de conjuntos diseminados.*
- *De segunda categoría en X si no es de primera categoría.*

Ejemplo 47 El conjunto $\emptyset$ es siempre diseminado, y por tanto de primera categoría.

Todo conjunto formado por un sólo punto $\{x\} \subset X$ es cerrado (por ser intersección de bolas cerradas, $\{x\} = \bigcap_{n=1}^{\infty} B[x, \frac{1}{n}]$). Por tanto puede suceder:

a) $\{x\}$ es además abierto, en cuyo caso $\{x\}$ es abierto y cerrado y el punto x se dice que es aislado. Esto ocurrirá si, por ejemplo $\{x\} = B(x, r)$ para algún $r > 0$.

b) $\{x\}$ es no abierto, en cuyo caso $int(\{x\}) = \emptyset$ y $\{x\}$ es diseminado. Esto ocurrirá si, por ejemplo, existe $r_0 > 0$ tal que $B(x, r) \neq \{x\}$ para $0 < r < r_0$. En el caso de la recta real y de todos los espacios métricos con métrica procedente de una norma.

Ejemplo 48 El la recta real, $\mathbb{Q}$ (y todo conjunto numerable) es de primera categoría, pues según se ha visto en el ejemplo anterior, cada conjunto formado por un solo elemento es diseminado.

En el plano euclídeo cualquier recta es un conjunto diseminado, y por tanto de primera categoría.

Consecuencias inmediatas de estas definiciones son:

PROPOSICIÓN 41 *Sea (X, d) un espacio métrico.*

1. *Un conjunto de segunda categoría nunca es vacío.*
2. *Un conjunto $A \subset X$ es diseminado si y sólo si su clausura tiene complementario denso en X.*
3. *Un conjunto $A \subset X$ es diseminado si y sólo si el interior de su complementario es denso en X.*
4. *La unión finita o numerable de conjuntos de primera categoría es de primera categoría.*

Pueba:

1. El conjunto $\emptyset$ es diseminado, y por tanto de primera categoría.
2. $\emptyset = \mathrm{int(cl(A))} \Leftrightarrow \mathrm{X} = [\mathrm{int(cl(A))}]^{\mathrm{c}} = \mathrm{cl}[(\mathrm{cl(A)})^{\mathrm{c}}]$.
3. $\emptyset = \mathrm{int(cl(A))} \Leftrightarrow \mathrm{X} = [\mathrm{int(cl(A))}]^{\mathrm{c}} = \mathrm{cl}[(\mathrm{cl(A)})^{\mathrm{c}}] = \mathrm{cl}[\mathrm{int}(\mathrm{A}^{\mathrm{c}})]$.

5.5.2. El Teorema de Baire

Recordamos resultados conocidos de la teoría de espacios métricos.

TEOREMA 47 *En un espacio métrico completo (X, d) el conjunto X es de segunda categoría.*

En consecuencia, en un espacio métrico completo (X, d), si $X = \bigcup_{k=1}^{\infty} F_n$ con F_n cerrado, se tiene que existe $n_0 \in \mathbb{N}$ tal que $\operatorname{int}(F_{n_0})$ es no vacío.

Caso contrario cada F_n sería un cerrado con interior vacío y por tanto diseminado. Entonces X sería de primera categoría, lo que contradice al teorema de Baire.

5.6. El Principio de la Acotación uniforme

TEOREMA 48 *Principio de la acotación uniforme Sean $X; Y$ espacios normados sobre el mismo cuerpo. Sea $\mathcal{G}$ un subconjunto no vacío de $\mathcal{B}(X, Y)$. Si para cada $x \in X$ se cumple*

$$\sup\{\|A(x)\| : A \in \mathcal{G}\} < \infty \tag{20}$$

y $(X, \|.\|)$ es un espacio de Banach, entonces

$$\sup\{\|A\| : A \in \mathcal{G}\} = M < \infty,$$

es decir, $G \subset B[0_{\mathcal{B}(X,Y)}, M]$.

Prueba : Para cada entero positivo n definimos

$$F_n := \{x \in X : \forall A \in \mathcal{G}, \|A(x)\| \neq n\} = \bigcap_{A \in \mathcal{G}} \{x \in X : \|A(x)\| \leq n\}.$$

Es claro que $0_X \in F_n$ para cada n, y que cada F_n es un conjunto cerrado, ya que cada oparador A es continuo. Además

$$\bigcup_{n=1}^{\infty} F_n = X.$$

En efecto, fijado cualquier $x \in X$ la hipótesis de acotación puntual nos dice que

$$s(x) := \sup\{\|A(x)\| : A \in \mathcal{G}\} < \infty$$

luego si $n \geq s(x)$, sea cual sea el operador $A \in \mathcal{G}$,

$$\|A(x)\| \leq s(x) \leq n.$$

Por tanto, $x \in F_n$. Es decir, $X \subset \bigcup_{n=1}^{\infty} F_n \subset X$.

Por el teorema de Baire existe n_0 tal que $\operatorname{int}(F_{n_0}) \neq \emptyset$. Por tanto existen $x_0 \in F_{n_0}$ y $r > 0$ tales que $B[x_0, r] \subset F_{n_0}$.

Sea ahora cualquier operador $A \in \mathcal{G}$, y cualquier vector $x \in X$ con $\|x\| \leq 1$.

$$\begin{aligned}\|A(x)\| &= \tfrac{1}{r}\|A(rx)\| \\ &= \tfrac{1}{r}\|A(x_0+rx) - A(x_0)\| \\ &\leq \tfrac{1}{r}\|A(x_0+rx) - A(x_0)\| \\ &\leq \tfrac{1}{r}(\|A(x_0+rx)\| + \|A(x_0)\|).\end{aligned}$$

Como $x_0+rx \in B[x_0,r] \subset F_{n_0}$, $\|A(x_0+rx)\| \leq n_0$. Igualmente, como $x_0 \in F_{n_0}$, $\|A(x_0)\| \leq n_0$. Por tanto,

$$\|A(x)\| \leq \frac{1}{r}(n_0+n_0) = \frac{2n_0}{r}.$$

La cota superior obtenida $M := \frac{2n_0}{r}$ no depende del punto x, luego

$$\|A\| := \sup\{\|A(x)\| : x \in X\} \leq M,$$

y además la cota M tampoco depende del operador A elegido, de donde

$$\sup\{\|A\| : A \in \mathcal{G}\} \leq M < \infty.$$

$\square$

5.6.1. Prueba que no usa argumentos de categorías

Supongamos que $\sup\{\|A\| : A \in \mathcal{G}\} = \infty$. Construiremos una sucesión de vectores (x_n) de X y una sucesión de operadores (A_n) de $\mathcal{G}$ tal que $\|x_n\| = 4^{-n}$. Entonces la serie $\sum_{i=1}^{\infty} x_n$ será absolutamente convergente, y como X es un espacio de Banach, estará definido $x := \sum_{i=1}^{\infty} x_n \in X$.

Elijamos $x_0 \in S_X$ arbitrario. Sea $M_0 = \sup\{\|A(x_0)\| : A \in \mathcal{G}\}$, que por (20) es finito. Por nuestra suposición, podemos elegir A_1 en $\mathcal{G}$ con $\|A_1\| > 3(M_0+1)4^1$. Por definición de norma de un operador, existe un vector $v_1 \in S_X$, tal que $\|A_1(v_1)\| > \frac{2}{3}\|A_1\|$. Pongamos $x_1 := 4^{-1}v_1$.

Supuesto que hayan sido ya construidos de este modo $x_1, \ldots, x_{n-1}$, $A_1, \ldots, A_{n-1}$, $M_0, \ldots, M_{n-2}$, $x_1, \ldots, x_{n-1}$, sea

$$M_{n-1} := \sup\{\|A(x_1 + \ldots + x_{n-1})\| : A \in \mathcal{G}\}.$$

Desde luego, por (20) $M_{n-1} < \infty$. Por nuestra suposición existe $A_n \in \mathcal{G}$ tal que

$$\|A_n\| > 3(M_{n-1}+n)4^n.$$

Por definición de norma de un operador, existe un vector v_n de norma uno, tal que $\|A_n(v_n)\| > \frac{2}{3}\|A_n\|$. Pongamos $x_n := 4^{-n}v_n$. Obviamente $\|x_n\| = 4^{-n}$ y $\|A_n(4^n x_n)\| = \|A_n(v_n)\| > \frac{2}{3}\|A_n\|$, o, equivalentemente $\|A_n(x_n)\| > \frac{2}{3}\|A_n\|4^{-n} = \frac{2}{3}\|A_n\|\|x_n\|$. Por tanto hemos construido por inducción una sucesión de vectores (x_n) con $\|x_n\| = 4^{-n}$, una sucesión de números reales no negativos (M_n) y una sucesión de operadores en $\mathcal{G}$ verificando que

(i) $\|A_n(x_n)\| > \frac{2}{3}\|A_n\|\|x_n\|$, y
(ii) $\|A_n\| > 3(M_{n-1}+n)4^n$, de donde
(iii) $\|A_n(x_n)\| \overset{(i)}{>} \frac{2}{3}\|A_n\|\|x_n\| \overset{(ii)}{>} \frac{2}{3}3(M_{n-1}+n)4^n\|x_n\| = 2(M_{n-1}+n)$.
(iv) $\|\sum_{k>n} x_k\| \leq \sum_{k>n}\|x_k\| \leq \sum_{k>n} 4^{-k} = \frac{4^{-n-1}}{1-1/4} = \frac{4^{-n}}{3} = \frac{\|x_n\|}{3}$.
Luego
(v) $\left\|A_n\left(\sum_{k>n} x_k\right)\right\| \leq \|A_n\|\left\|\left(\sum_{k>n} x_k\right)\right\| \overset{(iv)}{\leq} \|A_n\|\dfrac{1}{3}\|x_n\| \overset{(i)}{\leq} \dfrac{1}{2}\|A_n(x_n)\|$. Entonces, por la desigualdad triangular y la definición de M_{n-1},

$$\begin{aligned}
&\|A_n(x)\| = \|A_n\left(\textstyle\sum_{n=1}^{\infty} x_n\right)\| = \left\|A_n\left(\textstyle\sum_{k=1}^{n-1} x_k\right) + A_n(x_n) + A_n\left(\textstyle\sum_{k>n} x_k\right)\right\| \geq \\
&\geq [\|A_n(x_n)\| - \left\|A_n\left(\textstyle\sum_{k>n} x_k\right)\right\|] - \left\|A_n\left(\textstyle\sum_{k=1}^{n-1} x_k\right)\right\| \\
&\overset{(v)}{\geq} [\|A_n(x_n)\| - \tfrac{1}{2}\|A_n(x_n)\|] - \left\|A_n\left(\textstyle\sum_{k=1}^{n-1} x_k\right)\right\| \\
&\overset{(iii)}{\geq} [M_{n-1}+n] - \left\|A_n\left(\textstyle\sum_{k=1}^{n-1} x_k\right)\right\| \\
&\overset{\text{def } M_{n-1}}{\geq} [M_{n-1}+n] - M_{n-1} = n.
\end{aligned}$$

En definitiva

$$\|A_n(x)\| = \left\|A_n\left(\sum_{n=1}^{\infty} x_n\right)\right\| \geq n,$$

y esto contradice la hipótesis (20).

$\square$

Corolario 20 *Sea $(X, \|\cdot\|)$ un espacio normado sobre $\mathbb{K}$, y sea C un subconjunto no vacío de X. Si para todo $f \in X^*$ se tiene que el conjunto de números reales*
$\{|f(x)| : x \in C\}$
es acotado, entonces C es acotado.

Prueba : Los elementos $x \in C$ pueden considerarse de modo natural, como aplicaciones lineales y continuas entre X^* y $\mathbb{K}$, sin más que poner

$$x(f) := f(x)$$

donde $x \in X$ y $f \in X^*$. Por otra parte X y $\mathbb{K}$ son espacios de Banach. Entonces identificamos C con el subconjunto $\mathcal{G} \subset \mathcal{B}(X^*, \mathbb{K}) = (X^*)^*$, formado por los propios elementos de C, considerados como funcionales lineales sobre el espacio dual X^*. Además, si $x \in X$, lo consideramos de este modo

$$\begin{aligned}
\|x\|_{X^{**}} &:= \sup\{|x(f)| : f \in X^*, \|f\|_{X^*} \leq 1\} \\
&= \sup\{|f(x)| : f \in X^*, \|f\|_{X^*} \leq 1\}.
\end{aligned}$$

Pero por una de las consecuencias del Teorema de Hahn- Banach, (Corolario (**??**)), $\sup\{|f(x)| : f \in X^*, \|f\|_{X^*} \leq 1\} = \|x\|$. Si para todo $f \in X^*$ se tiene que el conjunto de números reales $\{|f(x)| : x \in C\}$, estamos diciendo que

$$\sup\{|x(f)| : x \in \mathcal{G}\} < \infty.$$

Luego podemos aplicar el teorema anterior y obtenemos que existe $M < \infty$ de modo que

$$\sup\{\|x\|_{X^{**}} : x \in \mathcal{G}\} \leq M < \infty,$$

Es decir

$$\sup\{\|x\| : x \in C\} \leq M < \infty,$$

o sea que M es acotado.

$\square$

5.7. El espacio dual de $\mathcal{C}([a,b])$

5.7.1. Funciones de variación acotada

Sea $f : [a,b] \to \mathbb{C}$. Sea π una partición (finita) de [a,b] es decir $\pi = \{t_0, t_1, \ldots, t_n\}$ con $a = t_0 < t_1 < \ldots < t_n = b$. Llamaremos *variación de f correspondiente a π* al número real

$$V_\pi(f) := \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|.$$

Se llama *variación total* de f en $[a,b]$ al elemento de $[0,\infty]$ dado por

$$V_a^b(f) := \sup\{V_\pi(f) : \pi \in \mathcal{P}\},$$

donde $\mathcal{P}$ es es conjunto de todas las particiones de $[a,b]$. Se dice que esta función f es *de variación acotada* en $[a,b]$ si $V_a^b(f) < \infty$.

Ejemplo 49 Toda función real monótona creciente $f : [a,b] \to \mathbb{R}$ es de variación acotada en $[a,b]$ con $V_a^b(f) = f(b) - f(a)$.

Ejemplo 50 Si $f : [a,b] \to \mathbb{R}$ tiene derivada f' definida y acotada en $[a,b]$, es decir $|f'(x)| \leq M < \infty$ entonces f es de variación acotada en $[a,b]$ pues basta aplicar el teorema del valor medio:

$$V_\pi(f) := \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \leq \sum_{i=1}^{n} M|x_i - x_{i-1}| = M(b-a).$$

Ejemplo 51 La función $f : [0,1] \to \mathbb{R}$ dada por

$$f(x) = \begin{cases} x \cos\frac{\pi}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

es continua en $[0,1]$ pero no es de variación acotada en este intervalo ya que si damos la partición

$$\pi_n := \{0, \frac{1}{2n}, \frac{1}{2n-1}, ..., \frac{1}{2}, 1\}$$

se tiene que

$$V_{\pi_n}(f) = |1\cos\pi - \frac{1}{2}\cos 2\pi| + +|\frac{1}{2}\cos 2\pi - \frac{1}{3}\cos 3\pi| + \ldots +$$

$$\ldots + |\frac{1}{2n-1}\cos((2n-1)\pi) - \frac{1}{2n}\cos(2n\pi)| + |\frac{1}{2n}\cos(2n\pi) - 0| =$$
$$= |-1-\frac{1}{2}| + |\frac{1}{2}+\frac{1}{3}| + \ldots + |-\frac{1}{2n-1}-\frac{1}{2n}| + |\frac{1}{2n}| =$$
$$= 1 + 2\left(\frac{1}{2}+\frac{1}{3}+\ldots\frac{1}{2n}\right)$$

y, como la serie armónica es divergente,

$$\sup\{V_{\pi_n}(f)\} = +\infty.$$

Si escribimos $f_n := \chi_{[\frac{1}{n},1]}f$, es claro que

$$\begin{aligned}
\|f_n - f\|_\infty &= \sup\{|f_n(t)-f(t)| : t\in[0,1]\}\\
&= \sup\{|f_n(t)-f(t)| : t\in[0,\tfrac{1}{n}[\}\\
&= \sup\{|f(t)| : t\in[0,\tfrac{1}{n}[\}\\
&= \sup\{|t\cos\left(\tfrac{\pi}{t}\right)| : t\in]0,\tfrac{1}{n}[\}\\
&\leq \sup\{|t| : t\in]0,\tfrac{1}{n}[\}\\
&\leq \tfrac{1}{n}
\end{aligned}$$

Luego $\|f_n - f\|_\infty \to 0$. Es fácil ver que cada una de las funciones f_n es de variación acotada. Por tanto vemos que, en general, el límite uniforme de una sucesión de funciones de variación acotada puede no ser una función de variación acotada.

5.7.2. Propiedades

Se cumple:

1) *Toda función de variación acotada en $[a,b]$ es acotada en dicho intervalo.*

2) *La función compleja $f = (f_1 + if_2) : [a,b] \to \mathbb{C}$ es de variación acotada en $[a,b]$, si y sólo si lo son sus componentes.*

3) *Si f y g son funciones reales (o complejas) de variación acotada en $[a,b]$, entonces $f+g$ y fg son también de variación acotada en $[a,b]$.*

Además,

$$V_a^b(f+g) \leq V_a^b(f) + V_a^b(g).$$

$$V_a^b(fg) \leq M_f V_a^b(g) + M_g V_a^b(f)$$

donde M_f y M_g son cotas superiores de $|f|$ y de $|g|$ en $[a,b]$ respectivamente.

Dos consecuencias son

a) *El conjunto $\mathcal{BV}([a,b])$ de las funciones reales o complejas de variación acotada sobre $[a,b]$ tiene estructura de espacio vectorial real.*

b) *La diferencia de dos funciones reales monótonas en $[a,b]$ es una función de variación acotada en $[a,b]$.*

4) *Sea f una función de $\mathcal{BV}([a,b])$.*

Si $a < c < b$ entonces $f \in \mathcal{BV}([a,c])$ y $f \in \mathcal{BV}([c,b])$, siendo además

$$V_a^b(f) = V_a^c(f) + V_c^b(f).$$

5) *La función real $x \mapsto V_a^x$ no negativa es monótona no decreciente en $[a,b]$.*

6) Teorema de Jordan: *Toda función real de variación acotada en un intervalo compacto $[a,b]$ puede descomponerse en diferencia de dos funciones monótonas no decrecientes.*

5.7.3. El espacio de Banach $(\mathcal{BV}([a,b]), \|\cdot\|)$

La aplicación

$$f \mapsto \|f\| := |f(a)| + V_a^b(f)$$

es una norma sobre $\mathcal{BV}([a,b])$, como se comprueba fácilmente.

Aunque $\mathcal{BV}([a,b]) \subset \mathcal{B}([a,b])$ (donde $\mathcal{B}([a,b])$ es el espacio vectorial de las funciones acotadas sobre $[a,b]$), en general esta norma no coincide con la norma del supremo que es la usual en $\mathcal{B}([a,b])$. No obstante observemos que si $t \in [a,b]$, y $g \in \mathcal{BV}([a,b])$, entonces $V_a^t(g) \leq V_a^b(g)$, y como $\pi := \{a,t\}$ es una partición del intervalo cerrado $[a,t]$ entonces

$$|g(t)| \leq |g(t) - g(a)| + |g(a)| = |g(a)| + V_\pi(g) \leq |g(a)| + V_a^t(g) \leq |g(a)| + V_a^b(g) = \|g\|$$

lo que en particular implica que $\|g\|_\infty \leq \|g\|$.

Si (f_n) es una sucesión de Cauchy en $(\mathcal{BV}([a,b]), \|\cdot\|)$, entonces como para cada $t \in [a,b]$ y para cualesquiera enteros positivos p, q

$$|f_p(t) - f_q(t)| \leq \|f_p - f_q\|_\infty \leq \|f_p - f_q\|$$

tenemos que la sucesión de números complejos $(f_n(t))$ será de Cauchy, y que la sucesión de funciones (f_n) será de Cauchy en el espacio normado $(\mathcal{B}([a,b]), \|.\|_\infty)$. Por tanto, $(f_n(t))$ será convergente a un número complejo, digamos $f(t)$, (por ser $\mathbb{C}$ completo), y además la sucesión de funciones (f_n) converge a uniformemente (en norma $\|\cdot\|_\infty$) a cierta función, digamos φ de $\mathcal{B}([a,b])$ (por ser $(\mathcal{B}([a,b]), \|.\|_\infty)$ un espacio de Banach). Ahora bien, como la convergencia uniforme implica la convergencia puntual, para todo $t \in [a,b]$, $\varphi(t) = \lim_n f_n(t) = f(t)$, es decir que $\varphi = f \in \mathcal{B}([a,b])$. En definitiva podemos afirmar que si (f_n) es una sucesión de Cauchy en $(\mathcal{BV}([a,b]), \|\cdot\|)$, entonces existe una función acotada $f : [a,b] \to \mathbb{C}$ que es límite uniforme (y puntual, claro) de la sucesión (f_n).

Lema 10 *Sea $P = \{t_0, t_1, \ldots, t_m\}$ una partición del intervalo $[a,b]$. Sean f, φ funciones acotadas en $[a,b]$ con $\|f - \varphi\|_\infty < \varepsilon$. Si $V_P(f) > M$ entonces $V_P(\varphi) > M - 2m\varepsilon$.*

Prueba :

$$\begin{aligned}
V_P(\varphi) &:= \sum_{i=1}^{m} |\varphi(t_i) - \varphi(t_{i-1})| \\
&= \textstyle\sum_{i=1}^{m} |(f(t_i) - f(t_{i-1}) - \big((f(t_i) - \varphi(t_i)) - (f(t_{i-1}) - \varphi(t_{i-1})) \big)| \\
&\geq \textstyle\sum_{i=1}^{m} \big| |(f(t_i) - f(t_{i-1})| - \big| (f(t_i) - \varphi(t_i)) - (f(t_{i-1}) - \varphi(t_{i-1})) \big| \big| \\
&\geq \textstyle\sum_{i=1}^{m} |(f(t_i) - f(t_{i-1})| - \sum_{i=1}^{m} \big| (f(t_i) - \varphi(t_i)) - (f(t_{i-1}) - \varphi(t_{i-1})) \big| \\
&= V_P(f) - \textstyle\sum_{i=1}^{m} \big| (f(t_i) - \varphi(t_i)) - (f(t_{i-1}) - \varphi(t_{i-1})) \big|
\end{aligned}$$

Por ser $|f(t_i) - \varphi(t_i)| \leq \|f - \varphi\|_\infty < \varepsilon$ y también $|f(t_{i-1}) - \varphi(t_{i-1})| < \|f - \varphi\|_\infty < \varepsilon$ se tiene que

$$\big| (f(t_i) - \varphi(t_i)) - (f(t_{i-1}) - \varphi(t_{i-1})) \big| \leq |f(t_i) - \varphi(t_i)| + |f(t_{i-1}) - \varphi(t_{i-1})| < 2\varepsilon\,.$$

Luego

$$V_P(\varphi) \geq V_P(f) - \sum_{i=1}^{m} \big| (f(t_i) - \varphi(t_i)) - (f(t_{i-1}) - \varphi(t_{i-1})) \big| > M - \sum_{i=1}^{m} 2\varepsilon = M - 2m\varepsilon$$

$\square$

TEOREMA 49 *El espacio normado $(\mathcal{BV}([a,b], \|.\|)$ es completo.*

Prueba : Sea (f_n) una Sucesión de Cauchy de elementos de $(\mathcal{BV}([a,b], \|.\|)$. Por los comentarios hechos maás arriba sabemos que existe una función f de $(\mathcal{B}([a,b], \|.\|)$ tal que $\|f_n - f\|_\infty \to 0$. Toda sucesión de Cauchy es acotada, luego existe un número real $M > 0$ con $\|f_n\| \leq M$ para $n = 1, 2, \ldots$. Supongamos, para llegar a una contradicción que f no es de variación acotada. Existe entonces una partición $P = \{t_0, t_1, \ldots, t_m\}$ del intervalo $[a,b]$ tal que $V_P(f) > 2M$. Elegimos $n_0 \in \mathbb{N}$ tal que $\|f - f_{n_0}\|_\infty < \frac{M}{4m}$. Por el lema anterior se tiene que

$$V_P(f_{n_0}) > 2M - 2m\frac{M}{4m} > M$$

y de aquí que

$$\|f_{n_0}\| \geq V_a^b(f_{n_0}) \geq V_P(f_{n_0}) > M$$

lo que contradice al hecho de ser $\|f_{n_0}\| \leq M$. Así pues tenemos que la sucesión (f_n) converge uniformemente a una función de variación acotada $f \in \mathcal{BV}([a,b])$.

Veremos finalmente que $\|f_n - f\| \to 0$, es decir que la sucesión (f_n) converge en la norma de $\mathcal{BV}([a,b])$ a la función f.

Caso contrario existe $\varepsilon > 0$ y existe una subsucesión (f_{n_k}) de (f_n) de manera que $\|f_{n_k} - f\| \geq \varepsilon$ para $k = 1, 2, \ldots$. Como $f_n(a) \to f(a)$ podemos suponer que $|f_{n_k}(a) - f(a)| < \frac{\varepsilon}{3}$.

Elijamos pues f_{n_1}. Dado que esta función es de variación acotada y que

$$|f_{n_1}(a) - f(a)| + V_a^b(f_{n_1} - f) =: \|f_{n_1} - f\| \geq \varepsilon$$

mientras que $|f_{n_1}(a) - f(a)| < \frac{\varepsilon}{3}$ tendremos que $V_a^b(f_{n_1} - f) > \frac{2}{3}\varepsilon$. Existe entonces una partición $\pi_1 := \{t_0, t_1, \ldots, t_{m_1}\}$ del intervalo $[a, b]$ de modo que

$$V_{\pi_1}(f_{n_1} - f) > \frac{2}{3}\varepsilon\,.$$

Pero $\|f_n - f\|_\infty \to 0$. Existe un entero positivo n_2 de modo que

$$\|f_{n_2} - f\|_\infty < \frac{1}{6m_1}\varepsilon$$

Siendo $\|f_{n_2} - f\| \geq \varepsilon$, del mismo modo que para f_{n_1}, elegimos una partición $\pi_2 := \{t_0, t_1, \ldots, t_{m_2}\}$ del intervalo $[a, b]$ de manera que

$$V_{\pi_2}(f_{n_2} - f) > \frac{2}{3}\varepsilon\,.$$

Observemos que

$$\begin{aligned}
\|f_{n_2} - f_{n_1}\| &\geq V_{\pi_1}(f_{n_2} - f_{n_1}) \\
&= \textstyle\sum_{i=1}^{m_1} |(f_{n_2} - f_{n_1})(t_i) - (f_{n_2} - f_{n_1})(t_{i-1})| \\
&= \textstyle\sum_{i=1}^{m_1} |(f_{n_2} - f + f - f_{n_1})(t_i) - (f_{n_2} - f + f - f_{n_1})(t_{i-1})| \\
&\geq \textstyle\sum_{i=1}^{m_1} \big(|(f - f_{n_1})(t_i) - (f - f_{n_1})(t_{i-1})| - |(f - f_{n_2})(t_i) - (f - f_{n_2})(t_{i-1})|\big) \\
&\geq V_{\pi_1}(f_{n_1} - f) - \textstyle\sum_{i=1}^{m_1} |(f - f_{n_2})(t_i) - (f - f_{n_2})(t_{i-1})| \\
&\geq V_{\pi_1}(f_{n_1} - f) - \textstyle\sum_{i=1}^{m_1} \big(|(f - f_{n_2})(t_i)| + |(f - f_{n_2})(t_{i-1})|\big) \\
&\geq V_{\pi_1}(f_{n_1} - f) - \textstyle\sum_{i=1}^{m_1} \big(2\|f - f_{n_2}\|_\infty\big) \\
&= V_{\pi_1}(f_{n_1} - f) - 2m_1\|f - f_{n_2}\|_\infty \\
&> V_{\pi_1}(f_{n_1} - f) - \tfrac{1}{3}\varepsilon \\
&> \tfrac{1}{3}\varepsilon
\end{aligned}$$

Elegiremos a continuación n_3 suficientemente grande para que $\|f_{n_3} - f\|_\infty < \frac{\varepsilon}{6(m_1+m_2)}$. Elegiremos también $\pi_3 := \{t_0, t_1, \ldots, t_{m_3}\}$ partición de $[a, b]$ de modo que $V_{\pi_2}(f_{n_2} - f) >$

$\frac{2}{3}\varepsilon$. También se tendrá que

$$\begin{aligned}
\|f_{n_3} - f_{n_1}\| &\geq V_{\pi_1}(f_{n_3} - f_{n_1}) \\
&= \textstyle\sum_{i=1}^{m_1} |(f_{n_3} - f_{n_1})(t_i) - (f_{n_3} - f_{n_1})(t_{i-1})| \\
&= \textstyle\sum_{i=1}^{m_1} |(f_{n_3} - f + f - f_{n_1})(t_i) - (f_{n_3} - f + f - f_{n_1})(t_{i-1})| \\
&\geq \textstyle\sum_{i=1}^{m_1} \big(|(f - f_{n_1})(t_i) - (f - f_{n_1})(t_{i-1})| - |(f - f_{n_3})(t_i) - (f - f_{n_3})(t_{i-1})|\big) \\
&\geq V_{\pi_1}(f_{n_1} - f) - \textstyle\sum_{i=1}^{m_1} |(f - f_{n_3})(t_i) - (f - f_{n_3})(t_{i-1})| \\
&\geq V_{\pi_1}(f_{n_1} - f) - \textstyle\sum_{i=1}^{m_1} \big(|(f - f_{n_3})(t_i)| + |(f - f_{n_3})(t_{i-1})|\big) \\
&\geq V_{\pi_1}(f_{n_1} - f) - \textstyle\sum_{i=1}^{m_1} \big(2\|f - f_{n_3}\|_\infty\big) \\
&= V_{\pi_1}(f_{n_1} - f) - 2m_1\|f - f_{n_3}\|_\infty \\
&> V_{\pi_1}(f_{n_1} - f) - 2m_1 \tfrac{\varepsilon}{6(m_1+m_2)} \\
&\geq V_{\pi_1}(f_{n_1} - f) - \tfrac{1}{3}\varepsilon \\
&> \tfrac{1}{3}\varepsilon
\end{aligned}$$

y que

$$\begin{aligned}
\|f_{n_3} - f_{n_2}\| &\geq V_{\pi_2}(f_{n_3} - f_{n_2}) \\
&= \textstyle\sum_{i=1}^{m_2} |(f_{n_3} - f_{n_2})(t_i) - (f_{n_3} - f_{n_2})(t_{i-1})| \\
&= \textstyle\sum_{i=1}^{m_2} |(f_{n_3} - f + f - f_{n_2})(t_i) - (f_{n_3} - f + f - f_{n_2})(t_{i-1})| \\
&\geq \textstyle\sum_{i=1}^{m_2} \big(|(f - f_{n_2})(t_i) - (f - f_{n_2})(t_{i-1})| - |(f - f_{n_3})(t_i) - (f - f_{n_3})(t_{i-1})|\big) \\
&\geq V_{\pi_2}(f_{n_2} - f) - \textstyle\sum_{i=1}^{m_2} |(f - f_{n_3})(t_i) - (f - f_{n_3})(t_{i-1})| \\
&\geq V_{\pi_2}(f_{n_2} - f) - \textstyle\sum_{i=1}^{m_2} \big(|(f - f_{n_3})(t_i)| + |(f - f_{n_3})(t_{i-1})|\big) \\
&\geq V_{\pi_2}(f_{n_2} - f) - \textstyle\sum_{i=1}^{m_2} \big(2\|f - f_{n_3}\|_\infty\big) \\
&= V_{\pi_2}(f_{n_2} - f) - 2m_2\|f - f_{n_3}\|_\infty \\
&> V_{\pi_2}(f_{n_2} - f) - 2m_2 \ \tfrac{\varepsilon}{6(m_1+m_2)} \\
&> \tfrac{2}{3}\varepsilon - \tfrac{1}{3}\varepsilon \\
&= \tfrac{1}{3}\varepsilon
\end{aligned}$$

Prosiguiendo de este modo obtendremos una subsucesión (f_{n_k}) de (f_n) con

$$\|f_{n_i} - f_{n_j}\| \geq \frac{1}{3}\varepsilon$$

para $i > j$ Pero tal subsucesión no admite subsucesiones de Cauchy (respecto de la norma $\|.\|$). Esto contradice al hecho de que (f_n) es una sucesión de Cauchy (respecto de dicha norma). Hemos llegado a una contradicción que completa la prueba del teorema.

$\square$

5.7.4. La integral de Riemann-Stieltjes: Resumen

Sean $f \in \mathcal{C}([a,b])$ y $w \in \mathcal{BV}([a,b])$. Si $P = \{t_0, t_1, \ldots, t_n\}$ es una partición del intervalo $[a,b]$ podemos calcular el número complejo

$$s(P,f,w) := \sum_{j=1}^{n} f(t_j)[w(t_j) - w(t_{j-1})].$$

(Si la función w fuera la identidad la expresión anterior sería una Suma de Riemann de la función f respecto de la partición P).

Existe un número complejo z tal que para todo $\varepsilon > 0$ existe $\delta > 0$ de modo que para toda partición P del intervalo $[a,b]$ con $\|P\| := \text{máx}\{t_i - t_{i-1} : i = 1, \ldots, n\} < \delta$ se tiene que

$$|z - s(P,f,w)| < \varepsilon$$

Dicho número complejo se llama integral de Riemann-Stieltjes de f respecto de w en $[a,b]$, y se denota por

$$\int_a^b f(t)dw(t). \tag{21}$$

Se puede demostrar que si w tiene una derivada w' integrable Riemann en $[a,b]$, entonces

$$\int_a^b f(t)dw(t) = \int_a^b f(t)w'(t)dt\,.$$

Por supuesto que la integral de Riemann-Stieltjes mantiene la linealidad respecto del integrando que se cumple, en la integral de Riemann, es decir que si $f_1, f_2 \in \mathcal{C}([a,b])$, y α y β son escalares entonces existe la integral de Riemann-Stieltjes de $\alpha f_1 + \beta f_2$ respecto de w en $[a,b]$, y se cumple que

$$\int_a^b (\alpha f_1 + \beta f_2)(t)dw(t) = \alpha \int_a^b f_1(t)dw(t) + \beta \int_a^b f_2(t)dw(t)\ .$$

También hay linealidad en la función de variación acotada w (que a veces se llama integrador) de (21), es decir que si $w_1, w_2 \in \mathcal{BV}([a,b])$, y γ y δ son escalares entonces existe la integral de Riemann-Stieltjes de f respecto de $\gamma w_1 + \delta w_2$ en $[a,b]$, y se cumple que

$$\int_a^b f(t)d[\gamma w_1 + \delta w_2](t) = \gamma \int_a^b f(t)dw_1(t) + \delta \int_a^b f(t)dw_2(t)\ .$$

También se cumple la acotación siguiente:

$$\left|\int_a^b f(t)dw(t)\right| \leq \text{máx}\{|f(t)| : t \in [a,b]\}\, V_a^b(w)\,. \tag{22}$$

Esta acotación también generaliza una desigualdad análoga que cumple la integral de Riemann. De hecho, si w es la función identidad, $V_a^b(w) = b - a$, y

$$\left|\int_a^b f(t)dt\right| \leq \int_a^b |f(t)|dt \leq \text{máx}\{|f(t)| : t \in [a,b]\}\,(b-a).$$

Terminamos aquí el resumen de las propiedades más importantes de la integral de Riemann-Stieltjes de funciones continuas respecto a funciones de variación acotada.[14]

[14]Más detalles sobre la integral de Riemann-Stieltjes pueden leerse, por ejemplo, en el libro de W. Rudin *Principios de Análisis Matemático*.

5.7.5. El espacio dual de $\mathcal{C}([a,b])$

Trataremos de estudiar una representación del espacio dual del espacio de Banach complejo $(\mathcal{C}([a,b]), \|.\|_\infty)$ de las funciones continuas en el intervalo cerrado $[a,b]$ con valores complejos, con la norma del supremo.

Evidentemente, si g es una función de variación acotada sobre $[a,b]$ tiene sentido para cada $\varphi \in \mathcal{C}([a,b])$ la integral de Stieltes

$$\int_a^b \varphi(t) dg(t).$$

Fijada g, la aplicación $f : \mathcal{C}([a,b]) \to \mathbb{C}$ dada por

$$f(\varphi) := \int_a^b \varphi(t) dg(t)$$

es lineal y

$$|f(\varphi)| \leq \|\varphi\|_\infty V_a^b(g)$$

lo que nos demuestra que f es continua, es decir que f es un elemento del dual de $(\mathcal{C}([a,b]), \|.\|_\infty)$. Vamos a ver que, en cierto modo, todos los elementos de dicho espacio dual son del mismo tipo que f. Con más precisión:

Teorema 50 *F. Riesz* [15] *Sea f un elemento del espacio dual de $(\mathcal{C}([a,b]), \|.\|)$. Entonces existe una función g de variación aotada sobre $[a,b]$ de manera que para cada $\varphi \in \mathcal{C}([a,b])$,*

$$f(\varphi) = \int_a^b \varphi(t) dg(t)$$

Además $\|f\| = V_a^b(g)$.

Recordemos que, dado cualquier número complejo z se define el llamado signo de z por

$$sg(z) := \begin{cases} 0 & z = 0_{\mathbb{C}} \\ \frac{z}{|z|} & z \neq 0 \end{cases}$$

Entonces se tiene que si $z \neq 0_{\mathbb{C}}$,

$$sg(\bar{z}) := \frac{\bar{z}}{|\bar{z}|} = \frac{\bar{z}}{|z|}$$

de donde

$$z\ sg(\bar{z}) = z\frac{\bar{z}}{|z|} = \frac{|z|^2}{|z|} = |z|$$

[15] En el libro *Real and Functional Analysis*, de Mukherjea y Pothoven se dice que .$^{\text{es}}$te teorema fué descubierto por F. Riesz en 1909. Más tarde, en 1937, fue extendido por Banach al caso de un espacio métrico compacto (en lugar de $[a,b]$) y posteriormente por Kakutani en 1941 al caso de un espacio topolgíco de Hausdorff compacto (en lugar también de $[a,b]$). Kakutani consideró medidas con signo en lugar de funciones de variación acotada".

igualdad que también se cumple para $z = 0_{\mathbb{C}}$.

Consideraremos $\mathcal{C}([a,b])$ como subespacio cerrado de $\mathcal{B}([a,b])$, entendiendo que en ambos se asume definida la norma del supremo.

Sea f un elemento del espacio dual de $(\mathcal{C}([a,b]), \|.\|_\infty)$. Por el teorema de Hahn-Banach existe un elemento F del espacio dual de $(\mathcal{C}([a,b]), \|.\|_\infty)$ que extiende a f y con la misma norma, es decir $\|f\| = \|F\|$.

A partir de esta prolongación F definiremos una función compleja acotada sobre $[a,b]$, del siguiente modo: $g(a) = 0$, y para $s \in (a,b]$,

$$g(s) = F(\chi_{[a,s]})$$

Como $\chi_{[a,s]}$ (la función característica del intervalo cerrado $[a,s]$) es (real) acotada (y en general no continua) existe $F(\chi[a,s])$.

Nos proponemos ver que

La función g es de variación acotada sobre $[a,b]$ Sea $P = \{t_0, t_1, \ldots, t_n\}$ una partición arbitraria del intervalo $[a,b]$. Esto significa que $a = t_0 < t_1 < \ldots < t_n = b$. Para $i = 1, \ldots, n$ escribamos

$$\alpha_i := \overline{sg\big(g(t_i) - g(t_{i-1}\big)}$$

y definamos la función acotada $y : [a,b] \to \mathbb{C}$ por

$$y(t) := \alpha_1 \chi_{[t_0,t_1]} + \alpha_2 \chi_{]t_1,t_2]} + \ldots + \alpha_n \chi_{]t_{n-1},t_n]}.$$

Es claro que $\|y\|_\infty \leq 1$. Pongamos por mayor comodidad $y_0 \equiv 0_{\mathbb{R}}$

$$y_i = \chi_{[a,t_i]}$$

para $i = 1, \ldots, n$. (Observemos que estas funciones y_i dependen de P). Vamos a ver que para todo $t \in [a,b]$,

$$y(t) = \sum_{i=1}^{n} \alpha_i \big(y_i(t) - y_{i-1}(t)\big) \tag{23}$$

Esto es (tedioso pero) claro pues para $t \in [a,b]$

$$\sum_{i=1}^{n} \alpha_i \big(y_i(t) - y_{i-1}(t)\big) = \alpha_1 \big(y_1(t) - y_0(t)\big) + \alpha_2 \big(y_2(t) - y_1(t)\big) + \ldots + \alpha_n \big(y_n(t) - y_{n-1}(t)\big)$$

$$= y_1(t)\big(\alpha_1 - \alpha_2\big) + y_2(t)\big(\alpha_2 - \alpha_3\big) + \ldots + y_{n-1}(t)\big(\alpha_{n-1} - \alpha_n\big) + \alpha_n y_n(t).$$

Si $t \in [t_0, t_1]$ el primer miembro de esta igualdad queda

$$\alpha_1 \big(1 - 0)\big) + \alpha_2 \big(1 - 1\big) + \ldots + \alpha_n \big(1 - 1\big) = \alpha_1$$

mientras que el último queda

$$= 1\big(\alpha_1 - \alpha_2\big) + 1\big(\alpha_2 - \alpha_3\big) + \ldots + 1\big(\alpha_{n-1} - \alpha_n\big) + \alpha_n 1 = \alpha_1 \ .$$

De igual modo razonamos si $t \in]t_{i-1}, t_i]$ para $i = 2, \ldots, n$.

Observemos que

$$\begin{aligned} F(y) &= \textstyle\sum_{i=1}^n \alpha_i \big(F(y_i) - F(y_{i-1})\big) \\ &= \textstyle\sum_{i=1}^n \alpha_i \big(F(\chi_{[a,t_i]}) - F(\chi_{[a,t_{i-1}]})\big) \\ &= \textstyle\sum_{i=1}^n \alpha_i \big(g(t_i) - g(t_{i-1})\big) \\ &= \textstyle\sum_{i=1}^n \overline{sg\big(g(t_i) - g(t_{i-1}\big)}\big(g(t_i) - g(t_{i-1})\big) \\ &= \textstyle\sum_{i=1}^n \big|g(t_i) - g(t_{i-1})\big| \end{aligned}$$

Por tanto,

$$V_P(g) := \sum_{i=1}^n \big|g(t_i) - g(t_{i-1})\big| = F(y) = |F(y)| \leq \|F\|\|y\|_\infty = \|f\|\|y\|_\infty \leq \|f\|$$

y como el razonamiento no depende de la partición P elegida,

$$\sup\{V_P(g) : P \in \mathcal{P}([a,b])\} \leq \|f\| < \infty$$

es decir g es una función de variación acotada en $[a,b]$ y

$$V_a^b(g) \leq \|f\|. \tag{24}$$

Observemos que g en realidad depende de F. A su vez F depende de f, y el teorema de Hahn-Banach no nos garantiza la unicidad de F.

La función g *representa* al funcional f Supongamos $f \neq 0_{\mathcal{L}}$, pues para $f = 0_{\mathcal{L}}$ basta tomar $g = 0_{\mathcal{BV}([a,b])}$. Demos $\varepsilon > 0$, y fijemos una función continua $\varphi \in \mathcal{C}([a,b])$. Al ser g de variación acotada en $[a,b]$ está definida la integral de Stieltjes

$$\int_a^b \varphi(t)dg(t)$$

Por tanto existirá $\delta > 0$ tal que para toda partición $P = \{t_0, t_1, \ldots, t_n\}$ de $[a,b]$ con $\|P\| := \text{máx}\{t_i - t_{i-1} : i = 1, \ldots, n\} < \delta$ se cumple

$$\Big| \int_a^b \varphi(t)dg(t) - S(\varphi, P, g)\Big| < \frac{\varepsilon}{2}$$

donde $S(\varphi, P, g)$ son las sumas de Riemann-Stieltjes

$$S(\varphi, P, g) := \sum_{i=1}^n \varphi(t_i)\big(g(t_i) - g(t_{i-1})\big) \ .$$

Dadas la función continua φ y cualquier partición P del intervalo $[a,b]$ podemos definir la función acotada $\psi_P : [a,b] \to \mathbb{C}$

$$\psi_P = \varphi(t_1)\chi_{[t_0,t_1]} + \varphi(t_2)\chi_{]t_1,t_2]} + \ldots + \varphi(t_n)\chi_{]t_{n-1},t_n]}$$

Esta función (que depende de φ) se aproxima a φ respecto de la norma del supremo cuando P es suficientemente fina, pues como φ es uniformemente continua en $[a,b]$, dado $\varepsilon > 0$ existe $\eta > 0$ de modo que si $s_1, s_2 \in [a,b]$ con $|s_1 - s_2| < \eta$, entonces $|\varphi(s_1) - \varphi(s_2)| < \frac{\varepsilon}{\|f\|}$, donde k es cualquier constante positiva. En consecuencia, si $\|P\| < \eta$, cuando $t \in [t_0, t_1]$, obviamente $|t - t_1| < \eta$ luego

$$|\psi_P(t) - \varphi(t)| = |\varphi(t_1) - \varphi(t)| < \frac{\varepsilon}{2\|f\|}$$

Lo mismo ocurre si $t \in]t_{i-1}, t_i]$ para $i = 2, 3, \ldots, n$: En tal caso $|t - t_i| < \eta$, luego

$$|\psi_P(t) - \varphi(t)| = |\varphi(t_i) - \varphi(t)| < \frac{\varepsilon}{2\|f\|}.$$

Por tanto $\|\psi_P - \varphi\|_\infty = \sup\{|\psi_P(t) - \varphi(t)| : t \in [a,b]\} \leq \frac{\varepsilon}{2\|f\|}$. Por otra parte el funcional F es lineal y continuo en $\mathcal{B}([a,b])$, luego si $\psi \in \mathcal{B}([a,b])$

$$|F(\psi) - F(\varphi)| \leq \|F\| \, \|\psi - \varphi\|_\infty = \|f\| \, \|\psi - \varphi\|_\infty.$$

Luego si $\|P\| < \eta$,

$$|F(\psi_P) - F(\varphi)| \leq \|f\| \, \|\psi_P - \varphi\|_\infty < \|f\| \frac{\varepsilon}{2\|f\|} = \frac{\varepsilon}{2}.$$

Observemos que

$$\begin{aligned}
F(\psi_P) &= \varphi(t_1)F(\chi_{[t_0,t_1]}) + \varphi(t_2)F(\chi_{]t_1,t_2]}) + \ldots + \varphi(t_n)F(\chi_{]t_{n-1},t_n]}) \\
&= \varphi(t_1)F(\chi_{[t_0,t_1]}) + \varphi(t_2)F(\chi_{[t_0,t_2]} - \chi_{[t_0,t_1]}) + \ldots + \varphi(t_n)F(\chi_{[t_0,t_n]} - \chi_{[t_0,t_{n-1}]}) \\
&= \varphi(t_1)\big(g(t_1) - g(t_0)\big) + \varphi(t_2)\big(g(t_2) - g(t_1)\big) + \ldots + \varphi(t_n)\big(g(t_n) - g(t_{n-1})\big) \\
&= S(P, \varphi, g).
\end{aligned}$$

Por lo tanto, si P es una partición con $\|P\| < \text{mín}\{\eta, \delta\}$.

$$\begin{aligned}
\Big| \int_a^b \varphi(t)dg(t) - f(\varphi)\Big| &\leq \Big| \int_a^b \varphi(t)dg(t) - S(P,\varphi,g)\Big| + \Big|S(P,\varphi,g) - f(\varphi)\Big| \\
&= \Big| \int_a^b \varphi(t)dg(t) - S(P,\varphi,g)\Big| + \Big|F(\psi_P) - F(\varphi)\Big| \\
&< \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon\,.
\end{aligned}$$

Siendo $\varepsilon > 0$ arbitrario hemos probado que $\int_a^b \varphi(t)dg(t) = f(\varphi)$.

Obsevemos que esta igualdad se sigue inmediatamente (es una propiedad de la integral de Riemann-Stieltjes, ver (22)) que $|f(\varphi)| \leq \|\varphi\|_\infty V_a^b(g)$. De aquí se sigue (por definición de norma de un operador), que $\|f\| \leq V_a^b(g)$. Como la desigualdad contraria se ha visto ya en (24), la prueba está completa.

$\square$

5.8. Normalización en $(\mathcal{BV}([a,b]), \|.\|)$

El teorema anterior establece una correspondencia entre los elementos del espacio dual de $(\mathcal{C}([a,b]), \|.\|_\infty)$ y las funciones de variación acotada sobre el mismo intervalo. Si analizamos la prueba del teorema vemos que la función g obtenida para representar.[a] f depende de la prolongación F de f obtenida por aplicacón del teorema de Hahn-Banach, que en general no tiene por qué ser única. Una vez *concretada* F se obtienen una función g escalares α_i que dependen de g (y por tanto de F y de f). Si se examina la definición de estos escalares puede constatarse que cualquier función de la forma $g(.)+c$ con c constante define los mismos α_i. Vamos a tratar de afinar este proceso de construcción de la función g.

En $\mathcal{BV}([a,b], \|.\|)$ definimos la relación binaria $w_1 \sim w_2$ si y sólo si para toda función $y \in \mathcal{C}([a,b])$,

$$\int_a^b y(t)dw_1(t) = \int_a^b y(t)dw_2(t)\,.$$

En otras palabras, son identificadas sos funciones de variación acotada sobre $[a,b]$ si dan lugar al mismo funcional sobre $\mathcal{C}([a,b])$.

Lema 11 *$w \in \mathcal{BV}([a,b], \|.\|)$ es equivalente a la función nula si y sólo si para todo $c \in (a,b)$,*

$$w(a) = w(c+0) = w(c-0) = w(b)\,,$$

donde, como es habitual, $w(c+0) := \lim_{t\to c^+} w(t)$ y $w(c-0) := \lim_{t\to c^-} w(t)$.

Prueba : Supongamos que $w \sim 0_{\mathcal{BV}}$. Entonces, para toda función y continua en $[a,b]$,

$$\int_a^b y(t)dw(t) = 0$$

y si tomamos en particular como y la función constante 1,

$$\int_a^b dw(t) = 0\,.$$

Dado que w es de variación acotada,

$$w(b) - w(a) = \int_a^b dw(t) = 0\,.$$

Vamos ahora a demostrar que $w(c+0) = w(a)$. Para $h > 0$, consideremos la función continua

$$y_h(t) := \begin{cases} 1 & a \le t \le c \\ 1 - \frac{t-c}{h} & c < t < c+h \\ 0 & c+h \le t \le b \end{cases}$$

Como $w \sim 0_{\mathcal{BV}}$, para toda función y continua en $[a, b]$,

$$\int_a^b y(t)dw(t) = 0$$

y si tomamos en particular como y la función y_h,

$$0 = \int_a^b y_h dw(t) = \int_a^c dw(t) + \int_c^{c+h} y_h dw(t) = w(c) - w(a) + \int_c^{c+h} y_h(t)dw(t)$$

6. Notas

6.1. A

Se ha dado por supuesto cierto conocimiento de la Teoría de la Medida e integración, para poder leer estas páginas y también de los hechos básicos de la teoría de espacios de Hilbert.

6.2. B

Este material se ha escrito en períodos de tiempo muy distintos y distantes en el tiempo, siguiendo algunos de los libros recomendados en la bibliografía, entre otros.

6.3. C

En la red existen muchos profesores que han dejado cursos de Análisis Funcional al alcance de quien quiera imprimirlos y leerlos. En particular pueden buscarse los de B. Maurey (escrito en francés), P. Garret y T.B. Ward (en inglés) entre otros muchos.

Referencias

[1] J. B. Conway, A Course of Functional Analysis.Second Edition. Ed. Springer, 1990

[2] G.J.O. Jameson, Topology and Normed Spaces. Ed. Chapman and Hall, 1982.

[3] L. V. Kantorovich; G. P. Akilov , Functional Analysis, 2nd Edition, Ed. Elsevier, 1982.

[4] E. Kreyszig, Introductory Functional Analysis with aplications, Ed. Wiley and Sons, 1978.

[5] L. Lusternik; V. Sobolev, Precis d'Analyse Fonctionelle. Ed. Mir, 1989.

[6] Saxe, C. Beginning Functional Analysis, E. Springer, 2002.

[7] F.G. Tricomi, Integral equations, Dover Publ. Inc. NY. 1985.