Análisis Funcional

Ignacio Villanueva Díez

Índice general

Introducción

1. El Análisis Funcional

El *Análisis Funcional* como tal fue surgiendo a principios del siglo XX como el marco abstracto adecuado para solucionar una serie de problemas del Análisis muy importantes en esos momentos.

Desde entonces ha experimentado un gran desarrollo y en este momento es una herramienta muy sofisticada útil para abordar una amplia variedad de problemas. En **[11]** se puede ver un descripción del proceso histórico que desembocó en la aparición del Análisis Funcional. Incluimos aquí tan sólo una breve descripción de este proceso que nos permita ubicar nuestra materia en el seno de la matemática.

Desde el desarrollo del Cálculo Diferencial, al considerar las soluciones de una ecuación diferencial, se vio que en ocasiones era necesario considerar propiedades del *espacio (o conjunto) de soluciones* de la ecuación, pero no estaba claro cuál era la estructura que poseía dicho espacio de soluciones. Los trabajos de D. Bernouilli, Lagrange y sobre todo Fourier acerca de la resolución de ecuaciones diferenciales se empiezan a enfrentar a problemas que anticipan lo que será el desarrollo posterior del Análisis Funcional. Una de las características comunes a varios de estos procesos era el paso de un problema finito, por ejemplo la solución de un sistema finito de ecuaciones lineales, a la versión infinita del problema, lo que les fuerza a enfrentarse a problemas de convergencia que en esa época no eran entendidos.

Merece también la pena mencionar los trabajos de Sturm (1836) y Liouville (1837) sobre la solución de ecuaciones diferenciales de la forma

$$y'' - q(x)y + \lambda y = 0$$

con ciertas condiciones de contorno, que influyeron notablemente en el desarrollo de la *Teoría Espectral*.

Con el desarrollo del *Cálculo de Variaciones* aparece ya explícitamente la noción de *campo funcional* y de *distancia entre funciones* y aparecen de forma implícita algunos de los problemas provocados por la no compacidad de los conjuntos cerrados y acotados en espacios infinito dimensionales.

Uno de los problemas que más influyó en la aparición del Análisis Funcional fue el estudio de las *Ecuaciones Integrales*. Los trabajos de varios matemáticos en este área llevaron al desarrollo por Fredholm de una técnica para solucionar algunos de estos problemas que dio lugar a lo que hoy conocemos como *Alternativa de Fredholm*.

A partir de ese momento la situación estaba suficientemente madura para el comienzo de la cristalización del Análisis Funcional. A principios del siglo XX, Hilbert se interesó por los resultados de Fredholm y se dedicó al estudio de las ecuaciones integrales, publicando una serie de artículos en los que van apareciendo de forma más o menos explícita las nociones de autovalor, ortogonalidad de los distintos autoespacios, y la misma noción de bola unidad de un espacio de Hilbert (obviamente no en esos términos).

En el último de dichos artículos, Hilbert abandona el marco de las ecuaciones integrales y da el salto de abstracción que caracteriza al Análisis Funcional: intenta desarrollar una teoría abstracta de formas bilineales y cuadráticas de infinitas variables para luego aplicar dicha teoría el estudio de las ecuaciones integrales. En el desarrollo de esta teoría aparece ya de forma bastante explícita el espacio ℓ_2, la convergencia débil de sucesiones en ℓ_2 y su *principio de elección*, que en términos modernos podemos pensar como el Teorema de Eberlein para el caso particular de ℓ_2.

A continuación, en los trabajos de una serie de autores (Frechet, Schmidt, F. Riesz, Fischer entre otros) se fueron desarrollando muchas nociones que hoy son centrales en el Análisis Funcional como la *distancia* en un conjunto abstracto y desde ahí diversas nociones topológicas como *completitud, compacidad, separabilidad*, los espacios L_p y ℓ_p, el estudio de sus duales y el dual de $C[a,b]$. En un artículo seminal de F. Riesz en 1918 se desarrolla la Teoría de Operadores Compactos sobre espacios $C[a,b]$, pero el propio Riesz dice que la teoría se puede extender a otros espacios funcionales. En ese artículo aparece la definición de la *norma* de $C[a,b]$.

Hemos citado algunos autores y artículos, pero hay muchos más de esa época que hicieron que la situación estuviese madura para la aparición en 1922 de la Tesis de Banach, en la que se desarrolla una teoría general de espacios normados y operadores lineales entre ellos. Desde ese momento y hasta 1932 el Análisis Funcional vivió una década de vertiginoso desarrollo, en la que se sistematizó la aplicación de métodos topológicos al estudio de los espacios de Banach (B-espacios en la notación del propio Banach) y se probaron entre otros muchos resultados el Principio de Acotación Uniforme (que ya había aparecido en una forma más débil en la tesis de Banach), el Teorema de Hahn-Banach y el Teorema de la Aplicación Abierta, tres de los principales resultados de la asignatura que proponemos.

En esa misma década, von Neumann utiliza la noción de *espacio de Hilbert abstracto* (ya no exclusivamente ℓ_2 o L_2) para satisfacer una necesidad de los físicos de la época: la construcción del lenguaje formal adecuado para la Mecánica Cuántica, que sigue siendo esencialmente el utilizado hoy en día.

En 1932 el Análisis Funcional alcanza la mayoría de edad con la aparición de los libros *Theorie des Operations Lineaires*, de Banach, *Grundlagen der Quantenmechanik*, de von Neumann, y *Linear Transformations in Hilbert Space*, de Stone. Prácticamente todo el contenido de nuestra asignatura (y mucho más) se halla (con otro lenguaje, por supuesto) en esos libros.

Desde entonces el desarrollo del Análisis Funcional ha sido muy amplio, y no nos atrevemos ni siquiera a incluir aquí un pequeño resumen de este, aunque sí nos gustaría mencionar brevemente que la necesidad de estudiar espacios funcionales no normados dio lugar a la Teoría de Espacios Vectoriales Topológicos y en particular al estudio de los Espacios Localmente Convexos de donde surgió la *Teoría de las Distribuciones* que ha tenido tantas aplicaciones en el estudio de las Ecuaciones Diferenciales.

También a lo largo del siglo pasado numerosas áreas de la matemática han ido adoptando parte del lenguaje y las herramientas del Análisis Funcional, debido a su gran potencia.

1.1. Las fuentes. Mencionamos en esta sección los principales libros que hemos utilizado en la elaboración de esta memoria, ordenados de mayor a menor importancia para nuestro proyecto. Dichos libros son [**31**], [**24**] (de este libro hay una segunda versión, [**21**], actualizada y con más autores, más adecuada como referencia bibliográfica para los alumnos), [**13**], [**33**], [**8**], [**15**], [**16**] (estos tres últimos probablemente no sean adecuados para los alumnos por su excesiva complejidad), [**39**] (sólo para Teoría Espectral en espacios de Hilbert), [**12**] (sólo para el Teorema de Stone-Weierstrass) y los siempre atractivos libros de Rudin [**40, 41, 42**]. En menor medida hemos usado también [**1**] (una buena referencia para los alumnos por exhaustivo), [**37**], [**29**], [**5**], [**22**], [**26**], [**30**] y [**44**].

Los libros [**31**] y [**24**] son excelentes fuentes de ejercicios interesantes para este curso.

Contenidos

Espacios normados. Espacios de Banach.

Norma. Espacios normados. Espacios de Banach. Subespacios. Cocientes. Normas equivalentes. Lema de Riesz. Compacidad. Series en espacios de Banach. Desigualdades de Hölder y de Minkowski. Espacios ℓ_p. Breve presentación de la Teoría de Integración de Lebesgue. Espacios $L_p[0,1]$.

Aplicaciones lineales continuas entre espacios normados.

Aplicaciones lineales continuas. Espacios isométricos e isomorfos. Normas en $\mathcal{L}(X;Y)$. Espacio dual. Hiperplanos y formas lineales.

Teoremas de Hahn-Banach.

Teorema de Hahn-Banach en forma analítica. Forma normante. Espacios vectoriales topológicos. Espacios localmente convexos. Seminormas. Funcionales sublineales. Conjuntos absorbentes, equilibrados, convexos. Teoremas de Hahn-Banach en forma geométrica. Reflexividad. Dual de un espacio cociente. Aplicaciones.

El Teorema de Baire y sus consecuencias: el Principio de Acotación Uniforme y el Teorema de la Gráfica Cerrada

Teorema de Ascoli-Arzela. Principio de Acotación Uniforme. Teorema de Banach-Steinhaus. Gráfica cerrada. Teorema de la Gráfica Cerrada. Aplicaciones abiertas. Teorema de la Aplicación Abierta. Proyecciones en un espacio de Banach. Subespacios complementados. Teorema de Stone-Weierstrass. Aplicaciones.

Espacios duales y operadores traspuestos.

Espacio dual. Duales de c_0 y ℓ_p $(1 \leq p < \infty)$. Funciones de variación acotada. Dual de $C[0,1]$. Duales de $L_p[0,1]$ $(1 \leq p < \infty)$. Traspuesto de un operador.

Topologías débil y débil*.

Topología débil. Convergencia débil de sucesiones. Caso finito dimensional. Formas continuas con la topología débil. Teorema de Mazur. Propiedad de Schur de ℓ_1. Topología débil*. Teorema de Alaoglu. Teorema de Goldstine. Reflexividad. Aplicaciones.

Operadores compactos.

Operadores compactos. Operadores de rango finito. Propiedad de ideal. Teorema de Schauder. Aproximación de operadores compactos por operadores de rango finito. Operadores completamente continuos.

Teoría espectral de operadores compactos.

Operadores inversibles. Perturbaciones de la identidad inversibles. Espectro de un operador. Autovalores. Compacidad del espectro. Teorema de Gelfand. Teoría espectral de operadores compactos. Alternativa de Fredholm. Aplicaciones.

Espacios de Hilbert.

Producto escalar. Desigualdad de Schwarz. Espacios de Hilbert. Identidad de Polarización. Ortogonalidad. Teorema de Pitágoras. Ley del Paralelogramo. Distancia mínima a un convexo y cerrado. Proyección ortogonal. Complementación. Conjuntos ortogonales. Desigualdad de Bessel. Teorema de Riesz-Fischer. Bases Hilbertianas. Desarrollo en serie de Fourier. Fórmula de Parseval. Dimensión Hilbertiana. Teorema de representación de Riesz. Aplicaciones.

Teoría espectral en espacios de Hilbert: Operadores compactos.

Operador adjunto. Operadores normales y autoadjuntos. Espectro de un operador compacto autoadjunto. Espectro de un operador compacto normal. Teorema espectral para operadores compactos normales. Teorema espectral para operadores compactos autoadjuntos. Teorema espectral para operadores compactos.

Capítulo 1

Espacios normados. Espacios de Banach

La base abstracta del Análisis Funcional consiste en el estudio de *Espacios Funcionales*. Estos espacios tienen una doble estructura que nos permitirá su estudio. De un lado una estructura algebraica de espacio vectorial (habitualmente de dimensión infinita) y de otro lado una estructura topológica "bien relacionada" con la estructura algebraica. La forma más natural de definir una topología en un espacio vectorial, siempre que sea posible, es mediante una distancia asociada a una norma, y de hecho los primeros espacios funcionales y probablemente los más utilizados son *espacios normados*, cuya definición y estudio introducimos en este capítulo.

A continuación introducimos la noción de espacio de Banach, puesto que la *completitud* será una de las herramientas más útiles a nuestra disposición.

En muchas ocasiones necesitaremos estudiar subespacios y cocientes de espacios de Banach y normados, y por ellos vemos aquí que ambos respetan esas estructuras.

Mediante el Lema de Riesz probamos la no *compacidad* de la bola unidad de los espacios de dimensión infinita. Este es uno de los leitmotiv del Análisis Funcional, puesto que buena parte de las técnicas que estudiaremos van precisamente destinadas a remedar dicha falta de compacidad. Estudiamos también la compacidad en el caso finitodimensional, lo que nos da pie a hablar de normas equivalentes, que volverán a aparecer más adelante cuando hablemos de isomorfismos entre espacios de Banach.

Los espacios de Banach habitualmente llamados clásicos son $C(K)$, c_0, los ℓ_p y los L_p. $C(K)$ es fácil de definir. La construcción de los ℓ_p es sencilla una vez probada la desigualdad de Hölder, útil por sí misma en muchas ocasiones.

Puesto que nuestros alumnos previsiblemente no conocerán Teoría de la Medida, no podemos abordar la construcción de los espacios $L_p(\mu)$ en general. Sin embargo, no creemos adecuado enseñar Análisis Funcional sin hablar al menos de los espacios $L_p[0,1]$. Dado que no podemos estar seguros de que nuestros alumnos conozcan la Teoría de la Integración de Lebesgue, y dado que una presentación rigurosa de esta nos llevaría demasiado tiempo, hemos decidido incluir en este capítulo una presentación de dicha teoría escueta y sin demostraciones, pero suficiente para poder construir los espacios $L_p[0,1]$ y demostrar su completitud.

Para la preparación de este capítulo hemos seguido sobre todo [**24**], [**31**] y [**13**]. La exposición de la Teoría de la Integración de Lebesgue se basa principalmente en [**40**] y [**42**].

Espacios normados y espacios de Banach

Empezamos recordando la definición de *norma* de un espacio vectorial.

DEFINICIÓN 1.1. *Sea X un espacio vectorial sobre $\mathbb{R}$ o $\mathbb{C}$. Una función $||\cdot|| : X \longrightarrow [0,+\infty)$ es una norma en X si:*

1. $||x|| = 0$ *si y sólo si* $x = 0$.
2. $||\lambda x|| = |\lambda|\,||x||$, *para todo* $x \in X$ *y* $\lambda \in \mathbb{R}$ *(ó* $\mathbb{C}$*).*
3. $||x+y|| \leq ||x|| + ||y||$, *para todos* $x, y \in X$.

Un espacio vectorial con una norma $(X, ||\cdot||)$ *es un espacio vectorial normado.*

Veremos a lo largo de la asignatura que la completitud juega un importante papel en muchos de los razonamientos característicos del Análisis Funcional, por lo que habitualmente trabajaremos con espacios normados y *completos*. Esta es precisamente la definición de un *Espacio de Banach* llamados así porque fue Stefan Banach quien comenzó su estudio sistemático en [**6**] y sobre todo en [**7**].

Dado un espacio vectorial normado $(X, \|\cdot\|)$, automáticamente posee una estructura natural de espacio métrico, donde la métrica $d : X \times X \longrightarrow [0,+\infty)$ viene dada por

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

para todos $x, y \in X$. Es fácil ver que, así definida, d es efectivamente una distancia. Para comprobar la propiedad triangular, observamos que

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

Dos observaciones son importantes: en primer lugar, observar que la norma es uniformemente continua, dado que

$$|\|x\| - \|y\|| \leq \|x - y\|$$

(Para demostrar esto, nótese que

$$\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|$$

y por tanto

$$\|x\| - \|y\| \leq \|x - y\|.$$

Análogamente se tiene que

$$\|y\| - \|x\| \leq \|x - y\|.)$$

En segundo lugar es muy importante notar que la norma dota a X de una topología "bien relacionada" con su estructura vectorial, en el sentido de que tanto la suma como el producto por escalares son continuos. Veámoslo.

Sea $\epsilon > 0$. Entonces dados $x, y \in X$, para todos $z_1 \in B(x, \frac{\epsilon}{2})$, $z_2 \in B(y, \frac{\epsilon}{2})$,

$$\|z_1+z_2-(x+y)\| = \|x+z_1-x+y+z_2-y-x-y\| \leq \|z_1-x\|+\|z_2-y\| = \epsilon.$$

Análogamente, dados $x \in X$, $k \in \mathbb{K}$ ($\mathbb{R}$ ó $\mathbb{C}$), entonces tomando por ejemplo

$$\delta_1 = \text{mín}\{\frac{\epsilon}{3\|x\|}, \frac{\sqrt{\epsilon}}{\sqrt{3}}\}, \quad \delta_2 = \text{mín}\{\frac{\epsilon}{3|k|}, \frac{\sqrt{\epsilon}}{\sqrt{3}}\}$$

se tiene que, para todos $\alpha \in B(k, \delta_1)$ $z \in B(x, \delta_2)$,

$$\|\alpha z - kx\| = \|(k + \alpha - k)(x + z - x) - kx\| \leq \epsilon.$$

De aquí se sigue que la topología definida por la norma viene caracterizada por los entornos del origen.

EJEMPLO 1.2.

1. *Ya se han estudiado en cursos anteriores normas en espacios finito dimensionales. En particular se sabe que*

$$\ell_i^n := (\mathbb{K}^n, \| \cdot \|_i)$$

es un espacio de Banach para $i = 1, 2, \infty$ *con las definiciones habituales. Mas adelante veremos que se sigue de la desigualdad de Minkowsky que* ℓ_p^n *es un espacio de Banach para todo* $1 \leq p \leq \infty$.

2. *Dado un conjunto* T, *podemos definir* $B(T)$ *como el espacio de las funciones escalares acotadas definidas sobre* T. *Se ve fácilmente que la norma del supremo* $\| \cdot \|_\infty$ *definida como*

$$\|f\|_\infty = \sup_{x \in T} |f(x)|$$

es efectivamente una norma en $B(T)$. *Veamos que con esta norma* $B(T)$ *es un espacio de Banach.*

Sea $(f_n)_{n\in\mathbb{N}} \subset B(T)$ *una sucesión de Cauchy. Entonces para cada* $x \in T$ *la sucesión* $(f_n(x))_{n\in\mathbb{N}}$ *también es de Cauchy. Por tanto existe el* $\lim_n f_n(x)$. *Definimos entonces la función* $f : T \longrightarrow \mathbb{K}$ *como*

$$f(x) = \lim_n f_n(x)$$

Por ser la sucesión (f_n) *de Cauchy en la norma del supremo, se tiene que la sucesión está uniformemente acotada, y por tanto* f *está acotada. Sólo nos falta ver que* f_n *tiende a* f *en la norma del supremo. Sea* $\epsilon > 0$. *Entonces existe* $n_0 \in \mathbb{N}$ *tal que para todos* $n, m \geq n_0$ $\|f_n - f_m\| < \epsilon$. *Sea ahora* $x \in T$. *Sabemos que existe* $m \geq n_0$ *tal que* $|f(x) - f_m(x)| < \epsilon$. *Entonces*

$$|f(x) - f_{n_0}(x)| \leq |f(x) - f_m(x)| + |f_m(x) - f_{n_0}(x)| \leq 2\epsilon.$$

Nótese que para cada $x \in T$ m *puede variar, pero siempre existe algún* $m \geq n_0$ *para el que se tiene el resultado, y por tanto*

$$\lim_n \|f - f_n\|_\infty = 0.$$

La demostración es interesante no tanto por el resultado en sí mismo, sino porque esta es a menudo la técnica seguida para verificar que un espacio normado es completo: se parte de una sucesión de Cauchy, se encuentra un candidato a límite (a menudo, como aquí, utilizando la completitud de los escalares), y luego se verifica que (1) el candidato a límite está en el espacio, y (2) la sucesión converge efectivamente a ese límite en la norma del espacio.

3. *Sea* K *un espacio compacto de Hausdorff. Definimos* $C(K)$ *como el espacio de las funciones* $f : K \longrightarrow \mathbb{K}$ *continuas, dotado de la norma del supremo,* $\|\cdot\|_\infty$. *Se ve igual que en el ejemplo anterior que* $\|\cdot\|_\infty$ *es una norma. Veamos que con esta norma* $C(K)$ *es un espacio de Banach: dado que* $C(K)$ *es un subespacio vectorial de* $B(K)$, *las funciones acotadas sobre* K, *puesto que ya hemos visto que* $B(K)$ *es completo sólo es necesario demostrar que* $C(K)$ *es un subespacio cerrado de* $B(K)$. *Sea entonces* $(f_n) \subset C(K)$ *una sucesión tal que* $f_n \to f \in B(K)$. *Sólo hay que ver que* f *es continua. Esto es un ejercicio de topología, demostrar que el límite uniforme de funciones continuas es una función continua. Lo detallamos: Sea* $\epsilon > 0$. *Existe* $n_0 \in \mathbb{N}$ *tal que para todo* $n \geq n_0$

$$\|f_n - f\| < \epsilon.$$

Sea ahora $t_0 \in K$*. Existe un entorno* V *de* t_0 *tal que para todo* $t \in V$

$$|f_{n_0}(t_0) - f_{n_0}(t)| < \epsilon.$$

Entonces, para todo $t \in V$

$$|f(t) - f(t_0)| \leq |f(t) - f_{n_0}(t)| + |f_{n_0}(t) - f_{n_0}(t_0)| +$$
$$+ |f_{n_0}(t_0) - f(t_0)| \leq 3\epsilon$$

y por tanto f *es continua.*

4. *Veremos más adelante que siempre que podamos definir un producto escalar* $\langle \cdot, \cdot \rangle$ *en* X*, éste define una norma en* X *mediante*

$$\|x\|^2 = \langle x, x \rangle.$$

Los espacios de Banach cuya norma procede de un producto escalar se llaman espacios de Hilbert. Su importancia justifica que se estudien en un capítulo aparte.

Sabemos del estudio de los espacios vectoriales la importancia que tiene, dado un espacio vectorial, el estudio de sus subespacios y sus espacios cocientes. Por tanto, resulta también importante ver cómo ambos heredan la estructura de espacio de Banach del espacio original.

Proposición 1.3. *Sea* $(X, \|\cdot\|)$ *un espacio vectorial normado y sea* Y *un subespacio vectorial de* X*. Entonces tanto* Y *como su clausura* $\overline{Y}$ *son espacios normados con la norma heredada. Además, si* X *es un espacio de Banach* $\overline{Y}$ *también lo es.*

Demostración. Nótese primero que dado que la suma y el producto por escalares son continuos para la norma, $\overline{Y}$ es efectivamente un subespacio vectorial de X: Veamos por ejemplo que $\overline{Y}$ está cerrado para la suma: Sean $x, y \in \overline{Y}$, y sea $\epsilon > 0$. Existen $x', y' \in Y$ tales que $\|x' - x\| \leq \epsilon$, $\|y' - y\| \leq \epsilon$. Entonces $x' + y' \in Y$ y

$$\|x' + y' - (x + y)\| \leq 2\epsilon.$$

Es trivial que $\|\cdot\|$ induce una norma tanto en Y como en $\overline{Y}$. Además, si X es un espacio de Banach entonces $\overline{Y}$ es un subespacio cerrado de un espacio completo, y por tanto él mismo es completo. $\square$

Proposición 1.4. *Sea* $(X, \|\cdot\|)$ *un espacio vectorial normado y sea* $Y \subset X$ *un subespacio vectorial cerrado. Dado* $[x] \in X/Y$*, definimos*

$$|||[x]||| = \inf_{y \in Y} \|x + y\|.$$

Entonces $|||\cdot|||$ *es una norma en* X/Y *llamada la norma cociente. Además si* X *es un espacio de Banach entonces* X/Y *también es un espacio de Banach.*

Demostración. Sea $x \in X$. Claramente $|||[x]||| \geq 0$. Si $|||[x]||| = 0$ entonces existe una sucesión $(y_n)_n \subset Y$ tal que

$$\|x - y_n\| < \frac{1}{n}.$$

Entonces $y_n \to x$ y como Y es cerrado se tiene que $x \in Y$, luego $[x] = 0$.

Para ver la desigualdad triangular, sean $x_1, x_2 \in X$ y sean $y_1, y_2 \in Y$ tales que

$$\|x_1 + y_1\| < |||[x_1]||| + \epsilon$$

y

$$\|x_2 + y_2\| < |||[x_2]||| + \epsilon.$$

Entonces

$$|||[x_1]+[x_2]||| = |||[x_1+x_2]||| = \inf_{y\in Y} \|x_1+x_2+y\| \leq \|x_1+x_2+y_1+y_2\| \leq$$

$$\leq \|x_1 + y_1\| + \|x_2 + y_2\| < |||[x_1]||| + |||[x_2]||| + 2\epsilon$$

y como esto es cierto para todo $\epsilon > 0$ se sigue la propiedad triangular.

Finalmente, si $x \in X$ y $0 \neq k \in \mathbb{K}$, se tiene

$$|||k[x]||| = |||[kx]||| = \inf_{y\in Y} \|kx + y\| = \inf_{y\in Y} |k| \left\|x + \frac{y}{k}\right\| =$$

$$= |k| \inf_{y\in Y} \|x + y\| = |k||||[x]|||,$$

y por tanto $||| \cdot |||$ es una norma en X/Y.

Veamos finalmente que si X es un espacio de Banach entonces X/Y también lo es. Sea $([x_n])_{n\in\mathbb{N}} \subset X/Y$ una sucesión de Cauchy. Elijamos una subsucesión $(n_k)_k$ tal que

$$|||[x_{n_k}] - [x_{n_{k+1}}]||| < 2^{-k}.$$

Como

$$\|[x_{n_1}] - [x_{n_2}]\| = \|[x_{n_1} - x_{n_2}]\| < \frac{1}{2},$$

se tiene que existe $y_{n_2} \in Y$ tal que

$$\|x_{n_1} - (x_{n_2} + y_{n_2})\| < \frac{1}{2}.$$

A continuación, puesto que

$$\|[x_{n_2}] - [x_{n_3}]\| = \|[x_{n_2} - x_{n_3}]\| < \frac{1}{4},$$

existe $y_{n_3} \in Y$ tal que

$$\|(x_{n_2} + y_{n_2}) - (x_{n_3} + y_{n_3})\| < \frac{1}{4}.$$

Continuando el proceso se tiene la existencia de una sucesión $(y_{n_k})_k \in Y$ tal que

$$\|(x_{n_k} + y_{n_k}) - (x_{n_{k+1}} + y_{n_{k+1}})\| < 2^{-k}$$

y por tanto podemos elegir los representantes x_{n_k} de manera que

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

De aquí se sigue que $(x_{n_k})_k$ es de Cauchy en X y por la completitud de X existe $x \in X$ tal que x_{n_k} tiende a x. Como

$$\|[x_{n_k}] - [x]\| = \|[x_{n_k} - x]\| \leq \|x_{n_k} - x\|$$

se sigue que $[x_{n_k}] \to [x]$ y por ser $([x_n])_n$ de Cauchy se tiene finalmente que $[x_n] \to [x]$ □

Observación 1.5. *A la hora de pensar en la norma cociente, es a menudo útil darse cuenta de que*

$$\|[x]\| = \inf_{y \in Y} \|x + y\| = \inf_{y \in Y} \|x - y\| = dist(x, Y).$$

Queremos estudiar la compacidad de los subconjuntos de espacios normados. Empecemos recordando la situación en dimensión finita. Previamente necesitamos la noción de normas equivalentes, que volverá a aparecer más adelante.

Definición 1.6. *Sea X un espacio vectorial. Dadas dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$ sobre X, decimos que son equivalentes si inducen la misma topología.*

Proposición 1.7. *Dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$ sobre X son equivalentes si y sólo si existen dos constantes $c, C \in (0, +\infty)$ tales que, para todo $x \in X$*

$$c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1$$

Demostración. Supongamos que existen c, C como en el enunciado. Sea $x_0 \in X$ y $\epsilon > 0$. Entonces

$$B_1(x_0, \frac{\epsilon}{C}) \subset B_2(x_0, \epsilon)$$

y

$$B_2(x_0, c\epsilon) \subset B_2(x_0, \epsilon)$$

lo que muestra que las dos topologías son equivalentes.

Recíprocamente, supongamos que las dos topologías son equivalentes. Por tanto el conjunto $B_1(0,1)$ es un entorno abierto del 0 en la topología dada por la norma 2, y por ello existe $r > 0$ tal que

$$B_2(0,r) \subset B_1(0,1).$$

De aquí se sigue que, para todo $x \in X$,

$$\|x\|_1 \leq r^{-1}\|x\|_2.$$

La otra desigualdad se obtiene análogamente. □

Sea X un espacio vectorial en el que hay definidas dos normas equivalentes $\|\cdot\|_1$ y $\|\cdot\|_2$. Entonces la aplicación identidad

$$Id : (X, \|\cdot\|_1) \longrightarrow (X, \|\cdot\|_2)$$

es lineal, continua, biyectiva, y su inversa también es lineal y continua, como se ve al principio del Capítulo 3. Dos espacios normados entre los que podemos establecer una aplicación con esas características se dicen *isomorfos*.

Proposición 1.8. *Sea X un espacio vectorial de dimensión finita. Entonces dos normas cualesquiera en X son equivalentes.*

Demostración. Suponemos $X = \mathbb{K}^n$ y consideramos en el la norma 1 $\|\cdot\|_1$ y otra norma cualquiera $\|\cdot\|$. Si $(e_i)_{i=1}^n$ es la base canónica de $\mathbb{K}^n$, se tiene

$$\|x\| = \left\|\sum x_i e_i\right\| \leq \sum |x_i|\|e_i\| \leq \left(\max_i \|e_i\|\right)\|x\|_1.$$

Esto nos da una de las desigualdades buscadas y nos dice que

$$Id : (X, \|\cdot\|_1) \longrightarrow (X, \|\cdot\|)$$

es continua, de donde $\|\cdot\|$ es una función continua sobre $(X, \|\cdot\|_1)$. Puesto que $S_{(X,\|\cdot\|_1)}$ es un compacto (probamos este hecho más adelante), la función $\|\cdot\|$, que es estrictamente positiva en $S_{(X,\|\cdot\|_1)}$ alcanza su mínimo ahí al que llamamos δ y se tiene $\delta > 0$. Por tanto

$$\|\cdot\|_1 \leq \frac{1}{\delta}\|\cdot\|$$

que era la otra desigualdad buscada.

Sólo resta por tanto comprobar que efectivamente $S_{(X,\|\cdot\|_1)}$ es un compacto. Consideremos la sucesión $(x_k)_{k\in\mathbb{N}} \subset S_{(X,\|\cdot\|_1)}$ donde para todo $k \in \mathbb{N}$

$$x_k = (x_k^1, \ldots, x_k^n).$$

Se tiene que $\|x_k\| = \sum_{i=1}^n |x_k^i| = 1$ y por tanto para todo $1 \leq i \leq n$ la sucesión $(x_k^i)_{k\in\mathbb{N}}$ está acotada. Podemos entonces, tomando subsucesiones repetidas veces, extraer una sucesión de índices $(k_l)_{l\in\mathbb{N}}$ de manera que para todo $1 \leq i \leq n$ la sucesión $(x_{k_l}^i)_{l\in\mathbb{N}}$ converge a x^i. Entonces

$$\lim_{l\to\infty} \sum_{i=1}^n |x_{k_l}^i - x^i| = 0$$

y por tanto, llamando $x = (x^1, \ldots, x^n)$, se tiene que

$$\lim_{l\to\infty} \|x_{k_l} - x\|_1 = 0.$$

Además, puesto que $x_{k_l} \in S_{(X,\|\cdot\|_1)}$ para todo l, se tiene que $x \in S_{(X,\|\cdot\|_1)}$, de donde se concluye que efectivamente $S_{(X,\|\cdot\|_1)}$ es compacto. □

COROLARIO 1.9. *Si X es un espacio normado e $Y \subset X$ es un subespacio de dimensión finita, entonces Y está cerrado en X*

DEMOSTRACIÓN. Y es un espacio de dimensión finita (digamos n) con la norma heredada de X. Por el teorema anterior esta norma es equivalente a, por ejemplo, la norma del supremo, de manera que Y es (isomorfo a) ℓ_∞^n. Sabemos que este espacio es completo (esto está sugerido como ejercicio). La completitud no se conserva por homeomorfismos, pero sí se mantiene por homeomorfismos *uniformemente continuos*. Se verá al principio del siguiente capítulo (y de hecho es un ejercicio muy simple) que las aplicaciones lineales continuas son uniformemente continuas. Por tanto $Y \subset X$ es completo y se tiene que Y debe ser cerrado en X. □

El Teorema de Heine-Borel nos dice que los subconjuntos cerrados y acotados de $\mathbb{K}^n$ (con cualquiera de sus normas equivalentes) son compactos. Para probar la falsedad de este resultado en el caso de espacios de dimensión infinita necesitamos un resultado interesante en sí mismo.

TEOREMA 1.10 (Lema de Riesz, F. Riesz 1918). *Sea X un espacio normado y sea $Y \subset X$ un subespacio vectorial cerrado propio. Entonces para todo $0 \leq r < 1$ existe $x_r \in S_X$ tal que*

$$r < dist(x_r, Y) \leq 1.$$

DEMOSTRACIÓN. Sea $[z] \in X/Y$ tal que $1 > \|[z]\| > r$. Sea ahora un $z \in [z]$ (es decir, un representante de $[z]$) tal que $\|z\| \leq 1$ y sea $x_r = \frac{z}{\|z\|}$. Entonces se tiene que

$$dist(x_r, Y) = \frac{dist(z, Y)}{\|z\|} = \frac{\|[z]\|}{\|z\|} \geq \|[z]\| > r.$$

□

El siguiente teorema es ahora fácil de probar

TEOREMA 1.11. *Sea* $(X, \|\cdot\|)$ *un espacio normado. Entonces son equivalentes:*

1. *Todo conjunto cerrado y acotado de* X *es compacto*
2. B_X *es compacto*
3. X *tiene dimensión finita*

DEMOSTRACIÓN. Claramente (1) implica (2). Para ver que (2) implica (3), supongamos que X es de dimensión infinita. Usando el Lema de Riesz e inducción construimos una sucesión $(x_n) \subset S_X$ tal que $dist(x_n span\{x_1, \ldots, x_{n-1}\}) > \frac{1}{2}$ para todo $n \in \mathbb{N}$. Entonces, para todo $n \neq m$, $dist(x_n, x_m) > \frac{1}{2}$, por lo que la sucesión (x_n) no puede tener subsucesiones convergentes.

Veamos finalmente que (3) implica (1). Supongamos que $\dim X = m$, sea $E \subset X$ un conjunto cerrado y acotado y sea $(x_n)_n \subset E$. Sea $\{e_1, \ldots, e_m\}$ una base de X y para todo $n \in \mathbb{N}$ sea $x_n = \alpha_n^1 e_1 + \cdots + \alpha_n^m e^m$. Usando que $(x_n)_n \subset E$ y que E está acotado es fácil ver que para todo $1 \leq i \leq m$ la sucesión $(\alpha_n^i)_n \subset \mathbb{K}$ está acotada. Así, existe una subsucesión $(n_l)_l$ tal que para todo $1 \leq i \leq m$ la sucesión $(\alpha_{n_l}^i)_l$ tiende a $\alpha^i \in \mathbb{K}$ cuando l tiende a infinito. Entonces x_{n_l} tiende a $x = \alpha^1 e_1 + \cdots + \alpha^n e_n$ (esto se puede ver estableciendo un isomorfismo entre X y ℓ_1^n, o ℓ_∞^n, que lleve los e_i a los elementos de la base canónica). Por ser E cerrado, se tiene que $x \in E$ y con eso acabamos. □

Continuamos el capítulo dando una caracterización a menudo útil de los espacio de Banach en términos de la convergencia de series. En [**16**, Chapter 1. Notes and Remarks] aparece una gran cantidad de información acerca de resultados profundos aunque elementales en su formulación relacionados con la sumabilidad en espacios normados, incluso en espacios finito-dimensionales.

Dado un espacio normado y una sucesión $(x_n) \subset X$, podemos definir la serie $\sum_{n=1}^{\infty} x_n$ como

$$\sum_{n=1}^{\infty} x_n = \lim_{m \to \infty} \sum_{n=1}^{m} x_n,$$

al igual que hacíamos con las series numéricas. Si ese límite existe y vale x, decimos que la serie es *convergente*. Decimos que la serie $\sum_n x_n$ es *absolutamente convergente* si la serie numérica $\sum_{n=1}^{\infty} \|x_n\|$ converge. Parece obvio que la convergencia absoluta debería implicar la convergencia en norma de X. El problema podría ser la completitud. De hecho, se tiene

TEOREMA 1.12. *Un espacio normado X es completo si y sólo si toda serie absolutamente convergente es convergente.*

DEMOSTRACIÓN. Supongamos que X es un espacio de Banach y sea $\sum_n x_n$ una serie absolutamente convergente. Entonces, dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que para todos $p, q \geq n_0$

$$\sum_{n=p}^{q} \|x_n\| < \epsilon.$$

Por tanto, llamando $s_m = \sum_{n=1}^{m} x_n$, tenemos

$$\|s_p - s_q\| = \left\| \sum_{n=p+1}^{q} x_n \right\| \leq \sum_{n=p+1}^{q} \|x_n\| < \epsilon.$$

Es decir, $(s_m)_{m\in\mathbb{N}}$ es una sucesión de Cauchy y por tanto (s_m) converge a $x \in X$, ya que X es un espacio de Banach.

Recíprocamente, supongamos que toda serie absolutamente convergente es convergente, y sea $(x_n) \subset X$ una sucesión de Cauchy. Veamos que (x_n) tiene una subsucesión convergente. Empezamos construyendo inductivamente una subsucesión $(x_{n_m})_m \subset (x_n)_n$: Sea $n_1 \in \mathbb{N}$ tal que, para todo $n \geq n_1$,

$$\|x_{n_1} - x_n\| \leq 2^{-1}.$$

Supuestos ya elegidos $n_1 < n_2 < \cdots < n_{m-1}$, elegimos n_m de manera que $n_m > n_{m-1}$ y tal que, para todo $n \geq n_m$,

$$\|x_n - x_{n_m}\| \leq 2^{-m}.$$

Definimos además $x_{n_0} = 0$ y definimos ahora la sucesión $(y_i)_{i\in\mathbb{N}}$ como

$$y_i = x_{n_i} - x_{n_{i-1}}.$$

Entonces

$$x_{n_m} = \sum_{i=1}^{m} y_i$$

y

$$\|y_i\| = \|x_{n_i} - x_{n_{i-1}}\| \leq 2^{-(i-1)}.$$

Por tanto la serie $\sum_i y_i$ es absolutamente convergente y, por hipótesis, convergente. Es decir $\sum_i y_i$ converge a $x \in X$. Puesto que $x_{n_m} = \sum_{i=1}^{m} y_i$, tenemos que x_{n_m} converge a x y, por ser (x_n) de Cauchy, de aquí se sigue que x_n converge a x, y por tanto X es completo. $\square$

La construcción de los ℓ_p

Los espacios de Banach clásicos son los ℓ_p, los L_p y los espacios $C(K)$. Vamos a ver que los ℓ_p son espacios de Banach. Los espacios $C(K)$ ya están estudiados en el ejemplo 1.2 y el estudio de los L_p lo haremos si los alumnos ya saben teoría de la medida, o si finalmente les enseñamos la integral de Lebesgue. Para definir los espacios ℓ_p y comprobar que son espacios normados necesitamos primeramente estudiar las desigualdades de Hölder y de Minkowski.

Primeramente necesitamos un lema.

Lema 1.13. *Sean $a \geq 0$, $b \geq 0$, $p, q \in (1, \infty)$ tales que $\frac{1}{p} + \frac{1}{q} = 1$. Entonces*

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

Demostración. Sea $b > 0$ y definamos la función

$$\varphi(a) = \frac{a^p}{p} + \frac{b^q}{q} - ab$$

Es fácil ver que $\varphi'(a) > 0$ si $a > b^{q-1}$ y que $\varphi'(a) < 0$ si $0 < a < b^{q-1}$. Por tanto φ alcanza su mínimo en $a = b^{q-1}$ y $\varphi(b^{q-1}) = 0$. De ahí se sigue que

$$\varphi(a) = \frac{a^p}{p} + \frac{b^q}{q} - ab \geq 0 \text{ para todo } a \geq 0$$

y por lo tanto

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

□

Teorema 1.14 (Desigualdad de Hölder). *Sean $(a_k)_{k=1}^m, (b_k)_{k=1}^m \subset \mathbb{R}$, dos sucesiones finitas de términos positivos, y sean $p, q \in (1, \infty)$ dos números tales que $\frac{1}{p} + \frac{1}{q} = 1$. Entonces*

$$\sum_{k=1}^m a_k b_k \leq \left(\sum_{k=1}^m a_k^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^m b_k^q \right)^{\frac{1}{q}}.$$

Demostración. Podemos suponer que $a_k \neq 0 \neq b_k$ para todo k. Entonces, para todo $1 \leq k \leq m$ definimos

$$A_k = \frac{a_k}{\left(\sum_{j=1}^m a_j^p \right)^{\frac{1}{p}}} \text{ y } B_k = \frac{b_k}{\left(\sum_{j=1}^m b_j^p \right)^{\frac{1}{p}}}$$

y observamos que

$$\sum_{k=1}^{m} A_k^p = \sum_{k=1}^{m} B_k^p = 1$$

Usando el Lema 1.13 tenemos que

$$A_k B_k \leq \frac{A_k^p}{p} + \frac{B_k^q}{q}$$

y sumando esta desigualdad para todo k tenemos que

$$\sum_{k=1}^{m} A_k B_k \leq \frac{1}{p} \sum_{k=1}^{m} A_k^p + \frac{1}{q} \sum_{k=1}^{m} B_k^q = \frac{1}{p} + \frac{1}{q} = 1,$$

y por tanto

$$\sum_{k=1}^{m} \frac{a_k}{\left(\sum_{j=1}^{m} a_j^p\right)^{\frac{1}{p}}} \cdot \frac{b_k}{\left(\sum_{j=1}^{m} b_j^p\right)^{\frac{1}{p}}} \leq 1$$

lo que termina la demostración. □

TEOREMA 1.15 (Desigualdad de Minkowski). *Sean dos sucesiones $(a_k)_{k=1}^n$, $(b_k)_{k=1}^n$ tales que, para todo k, $a_k \geq 0$ y $b_k \geq 0$, y sea $1 < p < \infty$. Entonces*

$$\left(\sum_{k=1}^{n} (a_k + b_k)^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{n} a_k^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} b_k^p\right)^{\frac{1}{p}}$$

DEMOSTRACIÓN. Si $p = 1$ el resultado es trivial. Sea $1 < p < \infty$ y sea $q \in (1, \infty)$ tal que $\frac{1}{p} + \frac{1}{q} = 1$. Entonces, usando la desigualdad de Hölder y el hecho de que $(p-1)q = p$ tenemos

$$\sum_{k=1}^{n} (a_k + b_k)^p = \sum_{k=1}^{n} (a_k + b_k)^{p-1} (a_k + b_k) =$$

$$= \sum_{k=1}^{n} (a_k + b_k)^{p-1} a_k + \sum_{k=1}^{n} (a_k + b_k)^{p-1} b_k \leq$$

$$\leq \left(\sum_{k=1}^{n} (a_k + b_k)^{(p-1)q}\right)^{\frac{1}{q}} \left(\sum a_k^p\right)^{\frac{1}{p}} +$$

$$+ \left(\sum_{k=1}^{n} (a_k + b_k)^{(p-1)q}\right)^{\frac{1}{q}} \left(\sum b_k^p\right)^{\frac{1}{p}} =$$

$$= \left(\sum_{k=1}^{n}(a_k+b_k)^p\right)^{\frac{1}{q}} \left(\sum a_k^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n}(a_k+b_k)^p\right)^{\frac{1}{q}} \left(\sum b_k^p\right)^{\frac{1}{p}}$$

Dividiendo todo por $(\sum_{k=1}^{n}(a_k+b_k)^p)^{\frac{1}{q}}$, tenemos que

$$\left(\sum_{k=1}^{n}(a_k+b_k)^p\right)^{\frac{1}{p}} = \left(\sum_{k=1}^{n}(a_k+b_k)^p\right)^{1-\frac{1}{q}} \leq \left(\sum a_k^p\right)^{\frac{1}{p}} + \left(\sum b_k^p\right)^{\frac{1}{p}}$$

como queríamos comprobar. □

El caso $p = 1$ de las desigualdades de Hölder y Minkowski se deja como ejercicio.

Con esto ya podemos definir los espacios ℓ_p, ℓ_p^n y comprobar que son espacios de Banach.

Definición 1.16. *Definimos ℓ_∞^n como el espacio vectorial n dimensional de las n-uplas de números reales o complejos ($\mathbb{R}^n$ ó $\mathbb{C}^n$) dotado con la norma del supremo $\|\cdot\|_\infty$, donde para todo $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ ó $\mathbb{C}^n$, su norma del supremo se define como:*

$$\|x\|_\infty = \max_i |x_i|$$

Sobre el mismo espacio vectorial $\mathbb{R}^n$ o $\mathbb{C}^n$ podemos definir otras normas. En particular tenemos

Definición 1.17. *Sea $1 \leq p < \infty$. Definimos ℓ_p^n como el espacio vectorial $\mathbb{R}^n$ ó $\mathbb{C}^n$ dotado con la norma p $\|\cdot\|_p$, definida como:*

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$$

Observación 1.18. *Es sencillo ver que $\|\cdot\|_p$ verifica las propiedades (1) y (2) de la definición de norma. La propiedad (3) (la desigualdad triangular) es precisamente la desigualdad de Minkowski.*

Es interesante representar las distintas bolas unidad de los espacios ℓ_p^2 para hacerse una idea de las distintas geometrías de estos espacios.

En los ejercicios veremos que los espacios normados finito dimensionales son completos, por lo que en particular ℓ_p^n es un espacio de Banach para todo $1 \leq p \leq \infty$ y para todo $n \in \mathbb{N}$.

Análogamente podemos definir los espacios ℓ_p $(1 \leq p \leq \infty)$.

Definición 1.19. *Sea $1 \leq p < \infty$. Definimos ℓ_p como el espacio vectorial de las sucesiones $x := (x_i)_{i\in\mathbb{N}}$ de números (reales o complejos) tales que $\sum_{i=1}^{\infty} |x_i|^p < \infty$, dotado con la norma p definida por*

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

Observación 1.20. *Para ver que ℓ_p está bien definido es necesario ver que efectivamente es un espacio vectorial y que $\|\cdot\| : \ell_p \longrightarrow [0,\infty)$ es efectivamente una norma. Veamóslo. La única dificultad reside en comprobar que, si $(x_i),(y_i) \in \ell_p$, entonces $(x_i)+(y_i) \in \ell_p$ y $\|(x_i) + (y_i)\|_p \leq \|(x_i)\|_p + \|(y_i)\|_p$:*

Para todo $n \leq m$, usando la desigualdad de Minkowski, se tiene

$$\left(\sum_{i=1}^{n} |x_i+y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}} \leq$$
$$\leq \left(\sum_{i=1}^{m} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{m} |y_i|^p\right)^{\frac{1}{p}}$$

Así pues, dejando tender m a infinito se tiene que

$$\left(\sum_{i=1}^{n} |x_i+y_i|^p\right)^{\frac{1}{p}} \leq \|(x_i)\|_p + \|(y_i)\|_p$$

para todo $n \in \mathbb{N}$, de forma que, al dejar que n tienda a infinito tenemos lo buscado.

Definición 1.21. *Definimos ℓ_∞ como el espacio vectorial de las sucesiones acotadas $x := (x_i)_{i\in\mathbb{N}}$ de números (reales o complejos), dotado con la norma del supremo $\|\cdot\|_\infty$ definida por*

$$\|(x_i)\|_\infty = \sup_{i\in\mathbb{N}} |x_i|.$$

Se ve fácilmente que ℓ_∞ está bien definido, es decir, que es un espacio vectorial y que $\|\cdot\|_\infty$ es en efecto una norma.

El espacio ℓ_∞ es "muy grande" y tienen gran importancia sus siguientes subespacios cerrados.

Definición 1.22. *Definimos c como el espacio vectorial de las sucesiones convergentes $x := (x_i)_{i\in\mathbb{N}}$ de números (reales o complejos), dotado con la norma del supremo $\|\cdot\|_\infty$, y definimos c_0 como el espacio vectorial de las sucesiones $x := (x_i)_{i\in\mathbb{N}}$ convergentes a cero de números (reales o complejos), dotado de nuevo con la norma del supremo.*

Todos los espacios definidos anteriormente son espacios de Banach. Ese es el contenido de la siguiente proposición.

PROPOSICIÓN 1.23. *Sea $1 \le p \le \infty$. Entonces ℓ_p es un espacio de Banach. Además c y c_0 también son espacios de Banach.*

DEMOSTRACIÓN. Veamos primero el caso $1 \le p < \infty$. Consideremos $(x^n)_{n\in\mathbb{N}} \subset \ell_p$ una sucesión de Cauchy, donde $x^n = (x_i^n)_i$. Dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que para todo $n, m \ge n_0$

$$\|x^n - x^m\| < \epsilon, \text{ es decir } \left(\sum_{i=1}^{\infty} |x_i^n - x_i^m|^p\right)^{\frac{1}{p}} < \epsilon.$$

En particular, para todo $i \in \mathbb{N}$, se tiene que $|x_i^n - x_i^m| < \epsilon$, y por tanto $(x_i^n)_n \subset \mathbb{K}$ es una sucesión de Cauchy. Usando la completitud de $\mathbb{K}$ tenemos que para todo $i \in \mathbb{N}$ existe $x_i \in \mathbb{K}$ tal que $x_i^n \to x_i$ cuando n tiende a ∞. Con esto ya tenemos el candidato a límite, que es $x = (x_i)_{i\in\mathbb{N}}$. Veamos que $x \in \ell_p$ y que x^n tiende a x en ℓ_p. Para lo primero, por ser $(x^n)_n$ una sucesión de Cauchy, sabemos que está acotada, y por tanto existe $C > 0$ tal que, para todo $n \in \mathbb{N}$, $\|x^n\|_p^p \le C$ es decir

$$\sum_{i=1}^{\infty} |x_i^n|^p \le C.$$

Entonces, para todos $j, n \in \mathbb{N}$ se tiene que

$$\sum_{i=1}^{j} |x_i^n|^p \le C.$$

Tomando límites cuando n tiende a infinito, se tiene que

$$\sum_{i=1}^{j} |x_i|^p \le C,$$

y dejando ahora que j tienda a infinito tenemos que

$$\sum_{i=1}^{\infty} |x_i|^p \le C$$

y por tanto $x \in \ell_p$.

Veamos finalmente que x^n tiende a x en $\|\cdot\|_p$. Sea $\epsilon > 0$. Sea $n_0 \in \mathbb{N}$ tal que, para todo $n, m \ge n_0$,

$$\|x^n - x^m\|_p < \epsilon.$$

Entonces, para todo $j \in \mathbb{N}$ tenemos que

$$\left(\sum_{i=1}^{j} |x_i^n - x_i^m|^p\right)^{\frac{1}{p}} < \epsilon.$$

Fijando $j \in \mathbb{N}$, eligiendo $n \geq n_0$ y dejando que m tienda a infinito obtenemos que

$$\left(\sum_{i=1}^{j} |x_i^n - x_i|^p\right)^{\frac{1}{p}} < \epsilon \text{ para todo } j \in \mathbb{N},\ n \geq n_0.$$

Dejando ahora que j tienda a infinito, obtenemos que

$$\left(\sum_{i=1}^{\infty} |x_i^n - x_i|^p\right)^{\frac{1}{p}} < \epsilon \text{ para todo } n \geq n_0$$

y por tanto x^n tiende a x en ℓ_p.

Para ver que ℓ_∞ es Banach, basta notar que $\ell_\infty = B(\mathbb{N})$, el conjunto de las funciones acotadas

$$f : \mathbb{N} \longrightarrow \mathbb{K}.$$

Ya hemos visto en el Ejemplo 1.2 que ese espacio es completo. También, si los alumnos conocen suficiente topología, se puede ver que $\ell_\infty = C(\beta\mathbb{N})$, las funciones continuas sobre la compactificación de Stone-Čech de $\mathbb{N}$ con la métrica usual. También en el Ejemplo 1.2 vimos que este espacio es un espacio de Banach.

Veamos finalmente que c y c_0 son espacios de Banach. Puesto que ambos son subespacios de ℓ_∞ y este es completo, sólo tenemos que ver que ambos son cerrados en ℓ_∞. Lo vemos para c, el caso de c_0 es algo más sencillo y se deja como ejercicio.

Sea $(x^n)_{n\in\mathbb{N}} \subset c$ una sucesión tal que x^n tiende a $x \in \ell_\infty$. Hemos de ver que $x \in c$. Para todo $n \in \mathbb{N}$ sea $l_n = \lim_{i\to\infty} x_i^n$. Veamos que $(l_n)_n$ es una sucesión de Cauchy. Sea $\epsilon > 0$. Existe $n_0 \in \mathbb{N}$ tal que, para todos $n, m \geq n_0$ se tiene

$$\|x^n - x^m\|_\infty < \epsilon.$$

Entonces para todo $i \in \mathbb{N}$

$$|x_i^n - x_i^m| < \epsilon.$$

Fijando $n, m \geq n_0$ y dejando que i tienda a infinito obtenemos que

$$|l_n - l_m| < \epsilon.$$

Por tanto $(l_n)_n$ es una sucesión de Cauchy y existe $l \in \mathbb{K}$ tal que (l_n) tiende a l. Veamos que $\lim_{i\to\infty} x_i = l$: sea $\epsilon > 0$. Existe $n_0 \in \mathbb{N}$ tal que para todo $n \geq n_0$

$$\|x^n - x\| < \epsilon \text{ y } |l_n - l| < \epsilon.$$

Además existe $j \in \mathbb{N}$ tal que para todo $i \geq j$

$$|x_i^{n_0} - l_{n_0}| < \epsilon.$$

Entonces, para todo $i \geq j$

$$|x_i - l| \leq |x_i - x_i^{n_0}| + |x_i^{n_0} - l_{n_0}| + |l_{n_0} - l| \leq 3\epsilon$$

y por tanto x_i tiende a l y $x \in c$. □

La construcción de los $L_p[0,1]$

Puesto que es posible que nuestros alumnos no conozcan la integral de Lebesgue, incluimos en esta memoria una breve descripción sin demostraciones de los resultados básicos de la Teoría de la Integral de Lebesgue. Creemos que debería ser posible explicar en una o dos horas los resultados de esta Teoría necesarios para poder construir los espacios $L_p[0,1]$.

Comenzamos con una definición

Definición 1.24. *Sea Ω un conjunto y sea $\mathcal{F} \subset \mathcal{P}(\Omega)$ una colección de subconjuntos de Ω. Decimos que $\mathcal{F}$ es una σ-*álgebra *si*

1. $\Omega \in \mathcal{F}$.
2. *Si $A \in \mathcal{F}$ entonces también su complementario $A^c \in \mathcal{F}$.*
3. *Si $(A_n)_{n\in\mathbb{N}} \subset \mathcal{F}$ entonces*

$$\cup_{n\in\mathbb{N}} A_n \in \mathcal{F}.$$

Se sigue de la definición que si $\mathcal{F}$ es una σ-álgebra y $A, B \in \mathcal{F}$ entonces $A \setminus B \in \mathcal{F}$, y si $(A_n)_{n=1}^\infty \subset \mathcal{F}$ entonces también $\cap_{n=1}^\infty A_n \in \mathcal{F}$.

Se puede demostrar que dada $T \subset \mathcal{P}(\Omega)$ una familia de subconjuntos de Ω, existe una σ-álgebra mínima (en el sentido de la inclusión) que contiene a T. A esta σ-álgebra la denominamos la σ-álgebra *generada* por T, y la escribimos $\sigma(T)$.

No pretendemos desarrollar una Teoría de la Medida abstracta, sino simplemente definir y justificar parcialmente los resultados principales de la Teoría de la Medida de Lebesgue. Por lo tanto, de ahora en adelante consideraremos $\Omega = [0,1]$ (todo es totalmente análogo para cualquier otro intervalo cerrado y acotado $[a,b] \subset \mathbb{R}$) y la σ-álgebra que consideraremos será $\mathcal{B}$, la σ-álgebra generada por los conjuntos abiertos de $[0,1]$. A $\mathcal{B}$ se le denomina habitualmente la *σ-álgebra de Borel.*

Un par $(\Omega, \mathcal{F})$, formado por un espacio y una σ-álgebra definida sobre él se denomina un *espacio de medida*. Nuestro espacio de medida será siempre $([0,1], \mathcal{B})$.

DEFINICIÓN 1.25. *Dado $([0,1], \mathcal{B})$, un conjunto $A \subset [0,1]$ se dice* medible *si $A \in \mathcal{B}$.*

En Teoría de la Medida es conveniente a menudo definir funciones con valores en la *recta real ampliada* $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ ó en $[0, \infty]$. En el caso complejo, diremos que una función toma valores en $\overline{\mathbb{C}}$ si f es de la forma $f = g + \imath h$, y tanto g como h toman valores en $\overline{\mathbb{R}}$. Finalmente, si queremos denotar una función que toma valores en $\overline{\mathbb{R}}$ o $\overline{\mathbb{C}}$, según el cuerpo sobre el que estemos trabajando, diremos que f toma valores en $\overline{\mathbb{K}}$. Esta notación *no* es estándar, pero creemos que simplifica algo la escritura de varios de los resultados siguientes.

DEFINICIÓN 1.26. *Una función $f : [0,1] \longrightarrow \overline{\mathbb{R}}$ se dice* medible *si para todo conjunto* abierto *$A \subset \mathbb{R}$ y para todo $a \in \mathbb{R}$ se tiene que $f^{-1}(A)$, $f^{-1}([-\infty, a))$ y $f^{-1}((a, +\infty])$ son medibles.*

DEFINICIÓN 1.27. *Una función $f : [0,1] \longrightarrow \overline{\mathbb{C}}$ se dice* medible *si $f = g + \imath h$, con $f, g : [0,1] \longrightarrow \overline{\mathbb{R}}$ funciones medibles.*

Se tiene la siguiente proposición que no demostraremos.

PROPOSICIÓN 1.28. *Sea $(f_n)_{n\in\mathbb{N}}$ una sucesión de funciones $f_n : [0,1] \longrightarrow \overline{\mathbb{R}}$ medibles. Entonces*

1. *$g = \sup_n f_n$ y $h = \limsup_n f_n$ son medibles.*
2. *Si $\lim_n f_n(t) = f(t)$ para todo $t \in [0,1]$ entonces f es medible.*
3. *Si $f, g : [0,1] \longrightarrow \overline{\mathbb{R}}$ son medibles entonces las funciones* $\max\{f, g\}$ *y* $\min\{f, g\}$ *son medibles. En particular* $\max\{0, f\}$ *y* $\min\{0, f\}$ *son medibles.*

DEFINICIÓN 1.29. *Una función $f : [0,1] \longrightarrow \mathbb{K}$ se dice* simple *si existe $n \in \mathbb{N}$, escalares $\{a_1, \ldots, a_n\}$ y conjuntos $A_1, \ldots, A_n \in \mathcal{B}$ tales que*

$$f = \sum_{i=1}^{n} a_i \chi_{A_i}$$

Claramente, toda función simple es medible.

Del siguiente resultado esbozamos la demostración.

TEOREMA 1.30. *Sea $f : [0,1] \longrightarrow \overline{\mathbb{R}}$ una función medible. Entonces existe una sucesión de funciones simples $(s_n)_n$ tal que*

1. $0 \leq s_1 \leq s_2 \leq \cdots$
2. *Para todo $t \in [0,1]$*

$$\lim_n s_n(t) = f(t)$$

DEMOSTRACIÓN. Para todo $n \in \mathbb{N}$ y para todo $1 \leq i \leq n2^n$ sea

$$E_{n,i} = f^{-1}\left(\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)\right)$$

y

$$F_n = f^{-1}([n, \infty])$$

y sea

$$s_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{E_{n,i}} + n\chi_{F_n}.$$

□

Si bien sólo vamos a utilizar la medida de Lebesgue, damos la definición general de medida (positiva contablemente aditiva)

DEFINICIÓN 1.31. *Dado un espacio de medida $(\Omega, \mathcal{F})$, una medida sobre dicho espacio es una función*

$$m : \mathcal{F} \longrightarrow [0, \infty]$$

tal que

1. $m(\emptyset) = 0$
2. *Para toda sucesión de conjuntos medibles $(A_n)_{n\in\mathbb{N}}$ dos a dos disjuntos (esto es, $A_n \cap A_m = \emptyset$ para todo $m \neq n$) se tiene*

$$m\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} m(A_n).$$

No es obvia la existencia de medidas no triviales. A continuación damos como teorema (sin demostración) la existencia de la medida de Lebesgue.

TEOREMA 1.32. *En el espacio de medida $([0,1], \mathcal{B})$ existe una única medida*

$$\lambda : \mathcal{B} \longrightarrow [0,1]$$

tal que para todo $0 \leq a \leq b \leq 1$

$$\lambda((a,b)) = \lambda([a,b]) = b - a.$$

A esta λ la denominamos medida de Lebesgue.

Nótese que para todo medida positiva m se tiene que si A, B son dos conjuntos medibles con $A \subset B$ entonces $m(A) \leq m(B)$. La siguiente proposición (que tampoco probamos) nos dice que podemos aproximar (en el sentido de la medida de Lebesgue) cualquier conjunto medible desde dentro por compactos y desde fuera por abiertos.

PROPOSICIÓN 1.33. *La medida de Lebesgue λ es* regular *en el siguiente sentido: para todo $A \subset [0,1]$ y para todo $\epsilon > 0$ existe un conjunto cerrado (equivalentemente compacto) $F \subset A$ y un conjunto abierto $G \supset A$ tales que*

$$\lambda(G) - \epsilon \leq \lambda(A) \leq \lambda(F) + \epsilon,$$

equivalentemente

$$\lambda(A \setminus F) \leq \epsilon \text{ y } \lambda(G \setminus A) \leq \epsilon$$

Necesitamos ahora definir *integración* respecto de la medida de Lebesgue. Comenzamos definiendo la integral de las funciones simples de la única forma razonable y la vamos extendiendo a las funciones medibles ganando generalidad. No es difícil ver que cada una de las definiciones va siendo coherente con las anteriores.

DEFINICIÓN 1.34. *Si $s = \sum_{i=1}^n a_i \chi_{A_i} : [0,1] \longrightarrow \mathbb{R}$ es una función simple y $E \subset \mathcal{B}$ es un conjunto medible, se define la* integral de s en E respecto de λ *como*

$$\int_E s d\lambda = \sum_{i=1}^n a_i \lambda(A_i \cap E).$$

DEFINICIÓN 1.35. *Si $f : [0,1] \longrightarrow [0,\infty]$ es una función medible y positiva, y $E \subset \mathcal{B}$ es un conjunto medible, se define la integral de f en E respecto de λ como*

$$\int_E f d\lambda = \sup_s \int_E s d\lambda,$$

donde el supremo se toma sobre todas las funciones simples s tales que $s \leq f$.

Nótese que la integral así definida puede valer ∞.

DEFINICIÓN 1.36. *Si $f : [0,1] \longrightarrow \overline{\mathbb{R}}$ es una función medible, se definen f^+ y f^- como*

$$f^+ = \max\{f, 0\} \text{ y } f^- = -\min\{f, 0\}.$$

De la Proposición 1.28 se sigue que tanto f^+ como f^- son medibles (y claramente positivas). Entonces, dado un conjunto medible $E \subset \mathcal{B}$ consideramos las dos integrales

$$\int_E f^+ d\lambda \ \text{ y } \int_E f^- d\lambda.$$

Si al menos una de esas dos integrales es finita, se define la integral de f *en* E *respecto de* λ *como*

$$\int_E f d\lambda = \int_E f^+ d\lambda - \int_E f^- d\lambda.$$

Finalmente,

DEFINICIÓN 1.37. *Si* $f : [0,1] \longrightarrow \overline{\mathbb{C}}$ *es una función compleja medible y* $f = g + \imath h$, *con* $g, h : [0,1] \longrightarrow \overline{\mathbb{R}}$ *medibles, dado un conjunto medible* $E \subset \mathcal{B}$ *se define la integral de* f *en* E *respecto de* λ *como*

$$\int_E f d\lambda = \int_E g d\lambda + \imath \int_E h d\lambda$$

si las integrales involucradas en la definición existen.

OBSERVACIÓN 1.38. *Se demuestra que si* f *es integrable Riemann en un intervalo* $[a, b]$, *entonces la integral de Riemann de* f *en* $[a, b]$ *y la integral de Lebesgue de* f *en* $[a, b]$ *coinciden. Se demuestra también que existen funciones no integrables Riemann que sí son integrables Lebesgue.*

Resumimos en las dos siguientes proposiciones las propiedades básicas de la integral de Lebesgue que necesitaremos más adelante, sin demostración. Necesitamos una definición previa.

DEFINICIÓN 1.39. *Sea* $f : [0,1] \longrightarrow \overline{\mathbb{R}}$ *una función medible. Decimos que* f *es* integrable *si las dos integrales que aparecen en su definición son finitas. Análogamente, si* $f : [0,1] \longrightarrow \overline{\mathbb{C}}$ *es una función medible, decimos que* f *es* integrable *si las dos integrales que aparecen en su definición son finitas.*

PROPOSICIÓN 1.40. *Sea* $f, g : [0,1] \longrightarrow \overline{\mathbb{R}}$ *y* $E \in \mathcal{B}$. *Entonces*

1. *Si* f *es medible y acotada en* E *entonces* f *es integrable.*
2. *Si* $a \leq f(x) \leq b$ *para todo* $x \in E$ *entonces*
$$a\lambda(E) \leq \int_E f d\lambda \leq b\lambda(E).$$
3. *Si* f, g *son integrables en* E *y* $f(x) \leq g(x)$ *para todo* $x \in E$ *entonces*
$$\int_E f d\lambda \leq \int_E g d\lambda.$$

PROPOSICIÓN 1.41. *Sea* $f, g : [0,1] \longrightarrow \overline{\mathbb{K}}$ *y* $E \in \mathcal{B}$. *Entonces*

1. *Si f, g son integrables, entonces para todo $\alpha, \beta \in \mathbb{K}$*
$$\int_E (\alpha f + \beta g) d\lambda = \alpha \int_E f d\lambda + \beta \int_E g d\lambda.$$
(Si bien es razonablemente inmediato que los escalares pueden salir fuera del signo integral, para demostrar la aditividad de la integral se suele usar el Teorema de la Convergencia Monótona, que veremos más adelante, o algún resultado equivalente.)
2. *Si f es medible y $\lambda(E) = 0$ entonces*
$$\int_E f d\lambda = 0.$$
3. *Si f es integrable en E y $A \in \mathcal{B}$ verifica $A \subset E$ entonces f es integrable en A.*
4. $\left| \int_E f d\lambda \right| \leq \int_E |f| d\lambda.$

De los resultados anteriores se deduce que los conjuntos de medida nula son despreciables en la integración:

Proposición 1.42. *Sean $f : [0,1] \longrightarrow \overline{\mathbb{K}}$ y $A, E \in \mathcal{B}$. Si $A \subset E$ y $\lambda(E \setminus A) = 0$ entonces*
$$\int_E f d\lambda = \int_A f d\lambda$$

Este último resultado hace que, si lo que nos interesa de las funciones sea su integral, parezca que tenga sentido considerar la siguiente *relación de equivalencia*. Dadas dos funciones medibles
$$f, g : [0,1] \longrightarrow \overline{\mathbb{K}}$$
decimos que $f \sim g$ si
$$\lambda \left(\{x \in [0,1] \text{ tales que } f(x) \neq g(x)\}\right) = 0.$$
Es claro que $\sim$ es una relación de equivalencia y que si $f \sim g$ entonces para todo $E \in \mathcal{B}$
$$\int_E f d\lambda = \int_E g d\lambda.$$

La siguiente notación útil está relacionada con lo anterior: Si una propiedad P ocurre para todo $x \in [0,1] \setminus A$ siendo A un conjunto de medida nula, decimos que P ocurre en casi todo punto, abreviado c.t.p.

En particular, $f \sim g$ si y sólo si $f(x) = g(x)$ c.t.p.

Los dos resultados que enunciamos a continuación, sin demostración, son fundamentales en la teoría de integración de Lebesgue, y nos permitirán intercambiar límites e integrales en numerosas ocasiones.

TEOREMA 1.43 (de la Convergencia Monótona). *Sea* $E \subset \mathcal{B}$. *Sea* (f_n) *una sucesión de funciones medibles* $f_n : [0,1] \longrightarrow [0,\infty]$ *tales que para todo* $x \in E$

$$0 \leq f_1(x) \leq f_2(x) \leq \cdots$$

Sea $f : [0,1] \longrightarrow [0,\infty]$ *la función definida como*

$$f(x) = \lim_n f_n(x) \text{ para todo } x \in [0,1].$$

Entonces

$$\lim_{n\to\infty} \int_E f_n d\lambda = \int_E f d\lambda.$$

TEOREMA 1.44 (de la Convergencia Dominada). *Sea* $E \subset \mathcal{B}$. *Sea* (f_n) *una sucesión de funciones medibles* $f_n : [0,1] \longrightarrow [0,\infty]$ *y sea* $f : [0,1] \longrightarrow [0,\infty]$ *tales que*

$$f(x) = \lim_n f_n(x) \text{ para todo } x \in [0,1].$$

(Nótese que f *es medible como consecuencia de la Proposición 1.28). Si existe una función* $g : [0,1] \longrightarrow [0,\infty]$ *integrable en* E *y tal que para todo* $n \in \mathbb{N}$ *y para todo* $x \in [0,1]$

$$|f_n(x)| \leq g(x)$$

entonces

$$\lim_{n\to\infty} \int_E f_n d\lambda = \int_E f d\lambda.$$

Una vez que tenemos a nuestra disposición los resultados básicos de la Teoría de integración de Lebesgue, podemos iniciar la construcción de los espacios $L_p[0,1]$.

Empezamos proponiendo como ejercicio la demostración de las desigualdades de Hölder y Minkowski en la versión integral.

EJERCICIO 1.1 (Desigualdad de Hölder). *Sean* $1 < p, q < \infty$ *con* $\frac{1}{p} + \frac{1}{q} = 1$, *sean* $f, g : [0,1] \longrightarrow [0,\infty]$ *medibles y sea* $E \subset \mathcal{B}$. *Entonces*

$$\int_E fg d\lambda \leq \left(\int_E f^p d\lambda\right)^{\frac{1}{p}} \left(\int_E g^q d\lambda\right)^{\frac{1}{q}}.$$

EJERCICIO 1.2 (Desigualdad de Minkowski). *Sean* $1 < p, q < \infty$ *con* $\frac{1}{p} + \frac{1}{q} = 1$, *sean* $f, g : [0,1] \longrightarrow [0,\infty]$ *medibles y sea* $E \subset \mathcal{B}$. *Entonces*

$$\left(\int_E (f+g)^p d\lambda\right)^{\frac{1}{p}} \leq \left(\int_E f^p d\lambda\right)^{\frac{1}{p}} + \left(\int_E g^p d\lambda\right)^{\frac{1}{p}}.$$

Definimos ahora las normas p para las funciones medibles.

DEFINICIÓN 1.45. *Sea* $f : [0,1] \longrightarrow \overline{\mathbb{K}}$ *una función medible y sea* $1 \leq p < \infty$. *Se define la norma* p *de* f *como*

$$\|f\|_p = \left(\int_0^1 |f|^p d\lambda\right)^{\frac{1}{p}}.$$

Veamos ahora el caso $p = \infty$.

DEFINICIÓN 1.46. *Sea* $f : [0,1] \longrightarrow [0,\infty]$ *una función medible. Sea*

$$S = \{\alpha \in \mathbb{R} \text{ tales que } \lambda(f^{-1}((\alpha,\infty])) = 0\}$$

Si $S = \emptyset$ *hacemos* $\beta = \infty$. *Si* $S \neq \emptyset$, *decimos que* $\alpha \in S$ *es una* cota esencial *de* S *y hacemos* $\beta = \inf S$. *Llamamos a* β *el* supremo esencial *de* f.

Ahora, dada una función medible $f : [0,1] \longrightarrow \overline{\mathbb{K}}$, *llamamos* β *al supremo esencial de* $|f|$ *y definimos la norma* ∞ *de* f *como*

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

Puesto que

$$f^{-1}((\beta,\infty]) = \cup_{n=1}^{\infty} f^{-1}((\beta + \frac{1}{n}, \infty])$$

y dado que la unión numerable de conjuntos de medida 0 tiene medida 0, se sigue que el supremo esencial es una cota esencial y por lo tanto

$$f(x) \leq \|f\|_\infty \text{ c.t.p.}$$

Ya podemos definir los espacios $\mathcal{L}_p$.

DEFINICIÓN 1.47. *Sea* $1 \leq p \leq \infty$. *Se define el espacio vectorial* $\mathcal{L}_p[0,1]$ *como*

$$\mathcal{L}_p[0,1] = \{f : [0,1] \longrightarrow \overline{\mathbb{K}} \text{ medibles tales que } \|f\|_p < \infty\}.$$

Quisiéramos ver ahora que el espacio $\mathcal{L}_p[0,1]$ con la norma p es un espacio de Banach. Lamentablemente la norma p *no* es una norma en $\mathcal{L}_p[0,1]$. Lo único que falla es que f puede ser distinta de 0 (en el sentido que esperaríamos de "ser distinto de 0") mientras que $\|f\|_p = 0$, tómese por ejemplo

$$f(x) = \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{si } x \neq 0. \end{cases}$$

La solución es definir los espacios $L_p[0,1]$ como el cociente de $\mathcal{L}_p[0,1]$ por la relación de equivalencia $\sim$ definida anteriormente.

Con esta definición se tiene

EJERCICIO 1.3. *Sea $1 \leq p \leq \infty$. El espacio $L_p[0,1]$ con la norma $\|\cdot\|_p$ es un espacio normado.*

Es algo más complicado ver que $L_p[0,1]$ es un espacio de Banach. Siguiendo la costumbre habitual, al escribir no somos totalmente rigurosos en la distinción formal entre $f \in \mathcal{L}_p[0,1]$ y su clase $[f] \in L_p[0,1]$. Creemos que no hay confusión posible y que ser totalmente formalistas confunde más que aclara.

TEOREMA 1.48. *Sea $1 \leq p \leq \infty$. El espacio $L_p[0,1]$ con la norma $\|\cdot\|_p$ es un espacio de Banach.*

DEMOSTRACIÓN. Consideramos primero el caso $1 \leq p < \infty$. Sea $(f_n)_n \subset L_p[0,1]$ una sucesión de Cauchy. Queremos ver que (f_n) converge a una función $f \in L_p[0,1]$, y por ser (f_n) una sucesión de Cauchy, está claro que basta probar que (f_n) tiene una subsucesión convergente. Así, tomando una subsucesión podemos suponer sin perdida de generalidad que

$$\|f_{n+1} - f_n\|_p \leq \frac{1}{2^n}.$$

Sea $f_0 = 0$ y para todos $n \in \mathbb{N}$, $t \in [0,1]$ sea

$$g_n(t) = \sum_{j=0}^{n} |f_{j+1}(t) - f_j(t)|$$

y

$$g(t) = \sum_{j=0}^{\infty} |f_{j+1}(t) - f_j(t)|$$

Entonces, por la desigualdad triangular se tiene

$$\|g_n\|_p \leq \sum_{j=0}^{n} \|f_{j+1} - f_j\|_p \leq \|f_1 - f_0\|_p + \sum_{j=1}^{n} \frac{1}{2^j} \leq \|f_1\|_p + 1.$$

Como además la sucesión (g_n) es creciente y converge a g del Teorema de la Convergencia Monótona se tiene que

$$\|g\|_p = \lim_n \|g_n\|_p \leq \|f_1\|_p + 1$$

y por tanto $g \in L_p[0,1]$, y en particular g es finita c.t.p.

Es decir, la serie $\sum_{j=0}^{\infty} f_{j+1}(t) - f_j(t)$ converge absolutamente c.t.p. y por tanto converge c.t.p.

Para todo $t \in [0,1]$ sea entonces

$$f(t) = \sum_{j=0}^{\infty} f_{j+1}(t) - f_j(t).$$

Como

$$f_n(t) = \sum_{j=0}^{n-1} f_{j+1}(t) - f_j(t),$$

se tiene que

$$\lim_n f_n(t) = f(t)$$

y

$$|f_n(t)| \leq \sum_{j=0}^{n-1} |f_{j+1}(t) - f_j(t)| \leq g(t) \text{ para todo } t \in [0,1].$$

Ahora el Teorema de la Convergencia Dominada nos dice que

$$\int_0^1 |f|^p d\lambda = \lim_n \int_0^1 |f_n|^p d\lambda \leq \int_0^1 g^p d\lambda < \infty$$

y por tanto $f \in L_p[0,1]$.

En ese caso también $|f| + g \in L_p[0,1]$ y $|f - f_n|^p \leq (|f| + g)^p$ para todo $n \in \mathbb{N}$. Por tanto, de nuevo el Teorema de la Convergencia Dominada nos dice que

$$\|f - f_n\|_p^p = \int_0^1 |f_n - f|^p d\lambda \to 0$$

y por tanto $f_n \to f$ en $\|\cdot\|_p$, lo que termina la demostración en el caso $1 \leq p < \infty$.

Si ahora $p = \infty$, sea $(f_n) \subset L_\infty[0,1]$ una sucesión de Cauchy. Para todos $n, m \in \mathbb{N}$ sea

$$A_n = \{t \in [0,1] \text{ tales que } |f_n(t)| \geq \|f\|_\infty\}$$

y sea

$$B_{n,m} = \{t \in [0,1] \text{ tales que } |f_m(t) - f_n(t)| \geq \|f_m - f_n\|_\infty\}.$$

Si

$$F = \left(\bigcup_{n=1}^{\infty} A_n \cup \bigcup_{n,m=1}^{\infty} B_{n,m} \right)$$

entonces $m(F) = 0$ y está claro que (f_n) tiende uniformemente en F^c a una función f acotada. Por lo tanto $f \in L_\infty[0,1]$ y $\|f_n - f\|_\infty \to 0$. $\square$

En muchas ocasiones la forma más cómoda de estudiar un espacio $L_p[0,1]$ es considerarlo como el completado de $C[0,1]$ con la norma p. Contamos eso a continuación sin detallar las demostraciones, puesto que eso nos llevaría a internarnos en la Teoría de la Integral de Lebesgue más de lo que consideramos adecuado para este curso.

Comenzamos enunciando sin demostración el Teorema de Lusin. Al presentarlo en clase podemos hacer un esquema de la demostración, utilizando la regularidad de λ y el Lema de Urysohn, aunque sin completar los detalles.

TEOREMA 1.49 (Lusin). *Sea $f : [0,1] \longrightarrow \overline{\mathbb{K}}$ una función integrable, sea $E \subset \mathcal{B}$ con $f(x) = 0$ si $x \notin E$ y sea $\epsilon > 0$. Entonces existe $g \in C[0,1]$ tal que*

$$\lambda(\{x; f(x) \neq g(x)\}) < \epsilon$$

y

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

Necesitamos otro resultado, interesante en sí mismo.

PROPOSICIÓN 1.50. *Sea S el conjunto de las funciones $s : [0,1] \longrightarrow \mathbb{K}$ simples y sea $1 \leq p < \infty$. Entonces S es denso en $L_p[0,1]$.*

DEMOSTRACIÓN. Vemos primero el caso real. Sea $f \in L_p[0,1]$, $f \geq 0$, y sea $(s_n) \subset S$ una sucesión tal que $0 \leq s_1 \leq s_2 \leq \cdots$ y

$$\lim_n s_n(x) = f(x) \text{ c.t.p.}$$

Puesto que $|f - s_n|^p \leq |f|^p$, del Teorema de la Convergencia Dominada se sigue que

$$\|f - s_n\|_p \to 0.$$

Una vez probado el caso $f \geq 0$ se puede probar el caso de f con valores en $\overline{\mathbb{R}}$ descomponiendo $f = f^+ - f^-$, y a continuación se prueba el caso de f con valores en $\overline{\mathbb{C}}$ descomponiendo $f = g + \imath h$. $\square$

Finalmente se tiene

TEOREMA 1.51. *Sea $1 \leq p < \infty$. Entonces $(C[0,1], \|\cdot\|_p)$ es denso en $L_p[0,1]$.*

DEMOSTRACIÓN. La demostración se sigue del resultado anterior más el Teorema de Lusin. $\square$

Finalmente podemos comentar que se puede demostrar que $C[0,1]$ es separable y que $L_\infty[0,1]$ no lo es, por lo que $(C[0,1], \|\cdot\|_\infty)$ no puede ser denso en $L_\infty[0,1]$.

Prácticas sugeridas.

Ejercicio 1.4. *Comprobar que todo espacio normado de dimensión finita es completo.*

Ejercicio 1.5. *Sea* $(X, \|\cdot\|)$ *un espacio normado. Entonces* $\|\cdot\| : X \longrightarrow [0, +\infty)$ *es una función continua.*

Ejercicio 1.6. *Para todo* $n \in \mathbb{N}$ *sea* X_n *un espacio de Banach. Probar que, para todo* $1 \leq p < \infty$, *el espacio*

$$\ell_p(X_n) = \{(x_n)_{n\in\mathbb{N}}$$

$$\textit{tales que } x_n \in X_n \textit{ para todo } n \in \mathbb{N} \textit{ y } \sum_{n=1}^{\infty} \|x_n\|^p < \infty\}$$

es un espacio de Banach con la norma

$$\|(x_n)_{n\in\mathbb{N}}\| = \left(\sum_{n=1}^{\infty} \|x_n\|^p\right)^{\frac{1}{p}}.$$

Análogamente probar que $c(X_n)$, $c_0(X_n)$ *y* $\ell_\infty(X_n)$ *con las definiciones obvias también son espacios de Banach.*

Ejercicio 1.7. *En muchas ocasiones (en particular en la construcción de los duales de estos espacios) necesitaremos una "aproximación por un número finito de coordenadas" a los elementos de* ℓ_p, c *y* c_0. *Para esto resulta muy adecuado definir el espacio*

$$c_{00} = \{(x_n)_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}} \textit{ tales que existe } n_0 \in \mathbb{N} \textit{ de manera que}$$

$$\textit{para todo } n \geq n_0 \quad x_n = 0\},$$

es decir el espacio de las sucesiones de escalares con un número finito de términos no nulos. Comprobar que, para todo $1 \leq p \leq \infty$, $(c_{00}, \|\cdot\|_p)$ *es un espacio normado pero* **no** *es un espacio de Banach. Para todo* $1 \leq p < \infty$ *comprobar que* $c_{00} \subset \ell_p$ *y que* c_{00} *es denso en* ℓ_p. *Comprobar también que* $c_{00} \subset c_0$ *y que* c_{00} *es denso en* c_0.

Ejercicio 1.8. *Aplicar el Teorema 1.12 para demostrar el recíproco de la Proposición 1.4: Si* X *es un espacio normado,* $Y \subset X$ *es un subespacio vectorial cerrado y* X/Y *es un espacio de Banach, entonces* X *es un espacio de Banach.*

EJERCICIO 1.9. *Bases de Schauder.* [**31**, p. 133]

EJERCICIO 1.10. *Separabilidad. Definición y propiedades básicas. Comprobar que* ℓ_p *es separable si y sólo si* $1 \leq p < \infty$*, y que* c_0 *es separable. Demostrar que* $C[0,1]$ *es separable.*

EJERCICIO 1.11. *Probar que si* $1 \leq p \leq \infty$ $L_p[0,1]$ *es un espacio de Banach.*

EJERCICIO 1.12. *Dar condiciones necesarias y suficientes para que se tenga la igualdad en las desigualdades de Hölder y Minkowski.*

EJERCICIO 1.13. *Demostrar que si* $1 \leq q < p \leq \infty$ *entonces la inclusión* $\ell_q \subset \ell_p$ *es propia.*

EJERCICIO 1.14. *Encontrar una sucesión que pertenezca a* c_0 *pero no a* ℓ_p *para ningún* $p \in [1, \infty)$.

EJERCICIO 1.15. *Hacer un boceto de las bolas unidad en* ℓ_p^2 *con* $p = 1, \frac{3}{2}, 2, 4, \infty$.

EJERCICIO 1.16. *Demostrar que si* X *tiene dimensión finita e* $Y \subset X$ *es un subespacio vectorial propio entonces existe* $x \in S_X$ *tal que* $d(x, Y) = 1$.

EJERCICIO 1.17. *Sea* $C^{(n)}[0,1]$ *el conjunto de las funciones continuas con derivada n-sima continua, con la norma*

$$\|f\| = \max_{0 \leq k \leq n} \|f\|_\infty.$$

Demostrar que $C^{(n)}[0,1]$ *con esa norma es un espacio de Banach.*

Capítulo 2

Aplicaciones lineales continuas entre espacios normados

Como ya hemos dicho, el objeto natural de trabajo del Análisis Funcional van a ser espacios vectoriales, en general de dimensión infinita, dotados de algún tipo de topología vectorial (en este curso sólo consideramos espacios normados). Resulta pues natural que en el estudio de estos espacios sea de gran importancia el estudio de las *aplicaciones lineales y continuas* entre ellos, dado que estas son las aplicaciones que preservan las dos estructuras (vectorial y topológica) con las que estamos trabajando. De hecho muchas de las aplicaciones del Análisis Funcional consisten precisamente en estudiar ciertas aplicaciones lineales y continuas destacadas entre espacios normados.

Tras ver las caracterizaciones más utilizadas de la continuidad de una aplicación lineal $T : X \longrightarrow Y$, estudiamos la situación cuando X tiene dimensión finita y vemos que si X tiene dimensión infinita siempre existen aplicaciones lineales no continuas. Definimos seguidamente cuándo dos espacios son *isomorfos* o *isométricos*.

Es fácil ver que $\mathcal{L}(X;Y)$, el espacio de las operadores de X a Y tiene estructura de espacio vectorial. Lo interesante es que también se puede definir una norma que le convierte en espacio normado, completo si Y lo es. En particular se sigue que el dual de todo espacio normado es un espacio de Banach.

Para terminar, incluimos en este capítulo un breve estudio de la relación entre formas lineales e hiperplanos, necesario para estudiar el Teorema de Hahn-Banach en el siguiente capítulo.

En la preparación de este capítulo hemos seguido principalmente **[31]**, **[24]** y **[13]**.

Aplicaciones lineales continuas

Empezamos con el siguiente teorema

TEOREMA 2.1. *Sean X, Y dos espacios normados, y sea $T : X \longrightarrow Y$ una aplicación lineal. Entonces, son equivalentes:*

(i) T *es continuo.*
(ii) T *es continuo en el origen.*
(iii) *Existe* $C > 0$ *tal que* $\|T(x)\| \leq C\|x\|$ *para todo* $x \in X$
(iv) T *es Lipschitziana, es decir, existe* $C > 0$ *tal que*

$$\|T(x) - T(y)\| \leq C\|x - y\|$$

para todos $x, y \in X$.
(v) $T(B_X)$ *está acotado en* Y.
(vi) $\ker(T)$ *es un subespacio cerrado y la aplicación lineal* $\overline{T} : X/\ker(T) \longrightarrow Y$ *definida como*

$$\overline{T}([x]) = T(x)$$

es continua.

Demostración. Claramente (i) implica (ii).

Para ver que (ii) implica (iii), si T es continuo en el origen, entonces, dado $\epsilon > 0$ existe $\delta > 0$ tal que, siempre que $\|x\| \leq \delta$ se tiene

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

Entonces, para todo $0 \neq x \in X$ se tiene que

$$\|T(x)\| = \left\| T\left(\frac{\|x\|}{\delta}\frac{x\delta}{\|x\|}\right)\right\| = \frac{\|x\|}{\delta}\left\| T\left(\frac{x\delta}{\|x\|}\right)\right\|.$$

Como

$$\left\|\frac{x\delta}{\|x\|}\right\| = \delta$$

entonces

$$\|T(x)\| \leq \frac{\epsilon\|x\|}{\delta}$$

y $C = \dfrac{\epsilon}{\delta}$ verifica (iii).

Que (iii) implica (iv) se sigue de la linealidad de T, puesto que

$$\|T(x) - T(y)\| = \|T(x - y)\| \leq C\|x - y\|$$

Claramente (iv) implica (i).

Si (iii) es cierto, entonces, para todo $x \in B_X$, $\|T(x)\| \leq C$, luego $T(B_X)$ está acotado y se tiene (v).

Recíprocamente, si tenemos (v), y, por ejemplo $\|T(x)\| \leq C$ para todo $x \in B_X$, entonces para todo $0 \neq x \in X$ se tiene

$$\|T(x)\| = \|x\| \left\| T\left(\frac{x}{\|x\|}\right)\right\| \leq C\|x\|$$

lo que implica (iii)

Veamos que (i) implica (vi). Claramente, si T es continuo entonces $\ker(T)$ es un subespacio cerrado. Por tanto ya vimos que $X/\ker(T)$ es un espacio normado con la norma inducida. Es conocido (y fácil de ver) que la aplicación $\overline{T}$ está bien definida y es lineal. Además

$$\|\overline{T}([x])\| = \|T(x)\| \leq C\|x\|$$

si T es continuo.

Recíprocamente, si se tiene (vi), existe C tal que

$$\|\overline{T}([x])\| \leq C\|[x]\|$$

y por tanto, para todo $x \in X$,

$$\|T(x)\| = \|\overline{T}([x])\| \leq \|[x]\| \leq C\|x\|$$

ya que

$$\|[x]\| = \inf_{z \in \ker(T)} \|x + z\| \leq \|x\|$$

por tanto (iii), y de ahí (i), son ciertos. □

Veamos como corolario (aunque es fácil probarlo directamente) que en espacios de dimensión finita todo operador lineal es continuo.

Corolario 2.2. *Sean X, Y dos espacios normados con* $\dim X < \infty$ *y sea $T : X \longrightarrow Y$ una aplicación lineal. Entonces T es continuo.*

Demostración. Puesto que ya hemos visto anteriormente que en los espacios de dimensión finita todas las normas son equivalentes, suponemos sin perdida de generalidad $X = \ell_1^n$. Sea $(e_i)_{i=1}^n$ la base canónica de ℓ_1^n y sea

$$M = \sup_i \|T(e_i)\|.$$

Entonces, para todo $x = (x_1, \ldots, x_n) \in B_{\ell_1}$,

$$\|T(x)\| = \left\| T\left(\sum_i x_i e_i\right)\right\| = \left\|\sum_i x_i T(e_i)\right\| \leq \sum_i |x_i|\|T(e_i)\| \leq$$

$$\leq M \sum_i |x_i| \leq M$$

y el teorema anterior nos garantiza que T es continuo. $\square$

Veamos que de hecho esa condición caracteriza a los espacios de dimensión finita.

TEOREMA 2.3. *Sea X un espacio normado de dimensión infinita e Y un espacio normado no nulo. Entonces existe una aplicación $T : X \longrightarrow Y$ lineal no continua.*

DEMOSTRACIÓN. Sea $(x_n) \subset B_X$ una familia de vectores linealmente independientes, consideremos una base B de X que contenga a la sucesión (x_n) y un vector $0 \neq y \in Y$. Definamos ahora la aplicación lineal $T : X \longrightarrow Y$ en todos los elementos de la base B mediante

$$T(b) = \begin{cases} ny\,, & \text{si } b = x_n \\ \\ 0\,, & \text{si } b \neq x_n \text{ para todo } n \end{cases}$$

y la extendemos por linealidad a todo X. Entonces claramente $T(B_X)$ no está acotado, por lo que T no es continuo. $\square$

Desde el punto de vista de las categorías, siempre es importante establecer alguna condición de "equivalencia" entre los objetos de la categoría. En la categoría de espacios de Banach, y de espacios normados, hay dos nociones que pueden jugar este papel, según en que aspectos de la teoría deseemos fijarnos.

DEFINICIÓN 2.4. *Dos espacios normados X e Y se dicen* isomorfos *si existe una aplicación lineal biyectiva y continua $\theta : X \longrightarrow Y$ tal que su inversa θ^{-1} (que sabemos por Álgebra Lineal que es lineal) es continua.*

Veremos que del Teorema de la Aplicación Abierta se sigue que si θ es como en la definición entonces θ^{-1} siempre verifica lo pedido.

Observemos que si X e Y son isomorfos, entonces X es Banach si y sólo si Y lo es, ya que los homeomorfismos lineales ya vimos que son *uniformemente continuos.*

DEFINICIÓN 2.5. *Dos espacios normados X e Y se dicen isométricos si existe una aplicación lineal y biyectiva $\theta : X \longrightarrow Y$ tal que, para todo $x \in X$,*

$$\|\theta(x)\| = \|x\|.$$

Los isomorfismos conservan la estructura lineal y topológica de los espacios. Las isometrías conservan además la *geometría* de los espacios.

Se distingue entonces a menudo entre la Teoría Isométrica, más preocupada de problemas de Geometría de Espacios de Banach, y la Teoría Isomórfica, más preocupada de problemas que se suelen denominar *estructurales*.

A continuación damos algunos ejemplos de aplicaciones lineales y continuas (operadores de ahora en adelante) o no continuas entre espacios normados.

EJEMPLO 2.6. *Sea* $A = (a_{i,j})_{i,j\in\mathbb{N}}$ *una matriz infinita y sea* $1 \leq p < \infty$.

Si $p = 1$*, supongamos que* $\gamma(j) = \sum_{i=1}^{\infty} |a_{i,j}| < \infty$ *para todo* $j \in \mathbb{N}$ *y que* $\gamma(j) \to 0$ *cuando* $j \to \infty$.

Si $1 < p < \infty$ *y* $\frac{1}{p} + \frac{1}{q} = 1$*, supongamos que*

$$\beta_p = \left(\sum_{i=1}^{\infty} \left(\sum_{j=1}^{\infty} |a_{i,j}|^q \right)^{\frac{p}{q}} \right)^{\frac{1}{p}} < \infty.$$

Entonces el operador

$$T_A : \ell_p \longrightarrow \ell_p$$

dado por

$$T_A(x) = \left(\sum_{j=1}^{\infty} a_{i,j} x_j \right)_j$$

(está bien definido y) es continuo.

Comencemos viendo el caso $p = 1$. *Notemos que si* $\gamma(j) \to 0$ *en particular se tiene que* $\sup_j \gamma(j) = C < \infty$. *Entonces*

$$\|T_A(x)\|_1 = \left\| \left(\sum_{j=1}^{\infty} a_{i,j} x_j \right)_i \right\|_1 = \sum_{i=1}^{\infty} \left| \sum_{j=1}^{\infty} a_{i,j} x_j \right| \leq \sum_i \sum_j |a_{i,j} x_j| =$$

$$= \sum_j \sum_i |a_{i,j}||x_j| \leq \sum_j |x_j| \sum_i |a_{i,j}| = \sum_j |x_j| \gamma(j) \leq C\|x\|_1,$$

por lo que T_A *es continuo.*

Veamos ahora el caso $1 < p < \infty$. *Nótese el uso de la desigualdad de Hölder.*

$$\|T_A(x)\|_p = \left\| \left(\sum_{j=1}^{\infty} a_{i,j} x_j \right)_i \right\|_p \leq \left\| \left(\sum_{j=1}^{\infty} |a_{i,j} x_j| \right)_i \right\|_p \leq$$

$$\leq \left\| \left(\left(\sum_{j=1}^{\infty} |a_{i,j}|^q \right)^{\frac{1}{q}} \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{\frac{1}{p}} \right)_i \right\|_p =$$

$$= \|x\|_p \left\| \left(\left(\sum_{j=1}^{\infty} |a_{i,j}|^q \right)^{\frac{1}{q}} \right)_i \right\|_p = \|x\|_p \beta_p,$$

y por tanto T_A es continuo.

Veamos un análogo continuo del ejemplo de arriba, lo que se conoce como operadores integrales de Fredholm.

EJEMPLO 2.7. *Sea $k(\cdot,\cdot) : [0,1] \times [0,1] \longrightarrow \mathbb{K}$ una función a la que de ahora en adelante nos referiremos como núcleo de Fredholm. El núcleo k nos permite definir una aplicación entre espacios funcionales (de momento por precisar)*

$$x \mapsto T_k(x)(s) = \int_0^1 \int_0^1 k(s,t)x(t)dt$$

Se pueden estudiar las características de este operador en función del núcleo k y del espacio en que lo definamos.

En este ejemplo, consideremos $1 < p \leq \infty$. Sea $X = L_p[0,1]$, sea $Y = L_q[0,1]$ con $\frac{1}{p} + \frac{1}{q} = 1$ y supongamos que $k \in L_q([0,1] \times [0,1])$. Entonces, para todo $x \in X$ y $s \in [0,1]$ usando la versión integral de la desigualdad de Hölder análogamente a como lo hicimos en el Ejemplo 7.10 tenemos

$$|T_k(x)(s)| \leq \int_0^1 |k(s,t)||x(t)|dt \leq \|x\|_p \left(\int_0^1 |k(s,t)|^q dt \right)^{\frac{1}{q}}$$

y por lo tanto

$$\|T_k(x)\|_q = \left(\int_0^1 |T_k(x)(s)|^q ds \right)^{\frac{1}{q}} \leq$$

$$\leq \|x\|_p \left(\int_0^1 \int_0^1 |k(s,t)|^q dtds \right)^{\frac{1}{q}} = \|x\|_p \|k\|_q$$

de donde se sigue la continuidad de $T_k : L_p[0,1] \longrightarrow L_q[0,1]$.

Veamos algún ejemplo de operador lineal no continuo

EJEMPLO 2.8. *Sea $C^1[0,1]$ el espacio de las funciones con derivada continua, dotado de la norma del supremo, y sea*

$$D : C^1[0,1] \longrightarrow C[0,1]$$

el operador derivación. Entonces D es lineal pero no continuo, ya que, si $f_n(t) = t^n$, entonces $\|f_n\| = 1$ pero

$$\|D(f_n)\| = \|f_n'\| = \sup_t \|nt^{n-1}\| = n$$

Ejemplo 2.9. *Análogamente, podemos definir la forma lineal*

$$x' : C^1[0,1] \longrightarrow \mathbb{K}$$

dada por

$$x'(f) = f'(1)$$

y ver que **no** *es continua. En efecto, veamos que $\ker x' = x'^{-1}\{0\}$ no es cerrado: sea $f_n = t - \frac{t^n}{n}$. Entonces $f_n'(1) = 1 - \frac{n}{n} = 0$ para todo $n \in \mathbb{N}$, y por tanto $f_n' \in \ker x'$. Claramente f_n tiende a $f(t) = t$ pero $f \notin \ker x'$.*

Ya sabemos del Álgebra Lineal que el espacio de operadores lineales entre espacios vectoriales tiene a su vez estructura de espacio vectorial. Es muy fácil ver que los operadores lineales y continuos forman un subespacio vectorial de este (sin más que notar que $\|(T+S)(x)\| = \|T(x) + S(x)\| \leq \|T(x)\| + \|S(x)\| \leq C_1\|x\| + C_2\|x\| = (C_1 + C_2)\|x\|$ y que $\|\alpha T(x)\| = |\alpha|\|T(x)\|$). Denotaremos a este último espacio por $\mathcal{L}(X;Y)$. Veamos que $\mathcal{L}(X;Y)$ hereda de X e Y una estructura natural de espacio normado.

Definición 2.10. *Dados dos espacios normados X e Y y un operador $T : X \longrightarrow Y$ definimos su norma como*

$$\|T\| = \sup_{x \in B_X} \|T(x)\|$$

Es fácil ver que esa función es efectivamente una norma en $\mathcal{L}(X;Y)$.

Lema 2.11. *Dados dos espacios normados X e Y y un operador $T : X \longrightarrow Y$ se tiene*

$$\|T\| = \inf \{C \geq 0, \text{ tal que } \|T(x)\| \leq C\|x\| \text{ para todo } x \in X\} = \\ = \sup_{x \in S_X} \|T(x)\|.$$

Demostración. Llamemos $I := \inf\{C \geq 0$, tal que $\|T(x)\| \leq C\|x\|$ para todo $x \in X\}$. Si C es tal que

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

para todo $x \in X$, entonces

$$\sup_{x \leq 1} \|T(x)\| \leq C$$

y por tanto $\|T\| \leq I$.

Por otro lado, si $0 \neq x \in X$, entonces

$$\|T(x)\| = \|x\| \|T\left(\frac{x}{\|x\|}\right)\| \leq \|x\| \sup_{z \in S_X} \|T(z)\|,$$

de donde

$$I \leq \sup_{z \in S_X} \|T(z)\| \leq \|T\| \leq I.$$

□

Ya hemos visto que $\mathcal{L}(X;Y)$ con la norma de operadores es un espacio normado. Veamos además que es un espacio de Banach si Y lo es

Proposición 2.12. *Dados dos espacios normados X e Y, si Y es un espacio de Banach entonces $\mathcal{L}(X;Y)$ también es un espacio de Banach.*

Demostración. La demostración es esencialmente la misma de siempre: consideramos una sucesión de Cauchy, le buscamos un candidato al límite y finalmente verificamos que dicho candidato es efectivamente un elemento del espacio y que la sucesión converge a él.

Sea $(T_n)_{n \in \mathbb{N}} \subset \mathcal{L}(X;Y)$ una sucesión de Cauchy. Es fácil ver que, para todo $x \in X$, $(T_n(x))_n$ es una sucesión de Cauchy en Y. Por tanto, puesto que Y es completo, existe $y \in Y$ tal que $(T_n(x))_n$ tiende a y. Consideramos la aplicación $T : X \longrightarrow Y$ definida por

$$T(x) = \lim_n T_n(x).$$

Debemos ver que T es lineal y continua: la linealidad se sigue del hecho de que la suma y el producto son continuos; veamos por ejemplo la aditividad

$$T(x+y) = \lim_n T_n(x+y) = \lim_n (T_n(x) + T_n(y)) = T(x) + T(y)$$

Para ver que T es continuo, por ser (T_n) una sucesión de Cauchy, para todo $\epsilon > 0$ existe n_0 tal que para todo $n, m \geq 0$, $\|T_n - T_m\| < \epsilon$.

Entonces, por la continuidad de la norma, para todo $m \geq n_0$

$$\begin{aligned} \|T(x) - T_m(x)\| &= \lim_n \|T_n(x) - T_m(x)\| \leq \\ &\leq \lim_n \|T_n - T_m\| \, \|x\| \leq \epsilon \|x\| \end{aligned} \tag{1}$$

y por tanto, tomando $\epsilon = 1$,

$$\|T(x)\| \leq \|T(x) - T_m(x)\| + \|T_m(x)\| \leq (1 + \|T_m\|)\|x\|$$

lo que implica que T es continuo. La convergencia de (T_n) a T se sigue de (1). $\square$

Probablemente sea interesante recalcar en algún momento que es en general *muy* difícil calcular la norma de un operador, incluso para muchos operadores aparentemente sencillos. Considérese por ejemplo el caso de operadores $T : \ell_2^n \longrightarrow \ell_2^n$, dados por una matriz. Dada una matriz, resulta un problema muy complejo y a menudo importante calcular, o al menos estimar, su norma como operador entre esos espacios.

La siguiente proposición es muy usada.

PROPOSICIÓN 2.13. *Sean X, Y, Z espacios normados, y sean $T \in \mathcal{L}(X;Y)$ y $S \in \mathcal{L}(Y;Z)$. Entonces $S \circ T \in \mathcal{L}(X;Z)$ y $\|S \circ T\| \leq \|S\|\|T\|$*

DEMOSTRACIÓN. Basta observar que

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

$\square$

Dado un espacio normado X, llamamos $X^* := \mathcal{L}(X;\mathbb{K})$. Del Teorema 2.12 se sigue que X^* es siempre completo. Llamamos a X^* el *dual topológico*, o simplemente *dual* de X. Del Teorema 2.3 se sigue que en los espacios de dimensión infinita el dual algebraico y el topológico no coinciden. Nosotros usaremos exclusivamente el dual topológico, puesto que es el que nos permitirá utilizar la estructura topológica del espacio.

Para poder estudiar en el próximo capítulo las aplicaciones geométricas del Teorema de Hahn-Banach necesitaremos estar familiarizados con la noción de hiperplano. Aprovechamos este momento para introducir su definición y probar algunos resultados elementales.

Sea X un espacio vectorial. Un *hiperplano* $M \subset X$ es un subespacio vectorial de codimensión 1, es decir, tal que $\dim(X/M) = 1$. Existe una relación biunívoca entre hiperplanos y (rectas de) formas lineales (no necesariamente continuas).

PROPOSICIÓN 2.14. *Un subespacio vectorial $M \subset X$ es un hiperplano si y sólo si existe una forma lineal no nula $x' : X \longrightarrow \mathbb{K}$ tal que $M = \ker x'$.*

DEMOSTRACIÓN. Si M es un hiperplano, sea $Q : X \longrightarrow X/M$ la aplicación cociente canónica, y sea $\theta : X/M \longrightarrow \mathbb{K}$ un isomorfismo. Entonces $x' = \theta \circ Q : X \longrightarrow \mathbb{K}$ es una aplicación lineal tal que $\ker x' = M$.

Recíprocamente, si $x' : X \longrightarrow \mathbb{K}$ es una forma lineal no nula, llamando $M = \ker x'$ tenemos que la aplicación canónica

$$\hat{x'} : X/\ker x' \longrightarrow \mathbb{K}$$

dada por

$$\hat{x'}([x]) = x'(x)$$

está bien definida, es lineal, inyectiva y sobreyectiva, y por tanto establece un isomorfismo lineal entre $X/\ker x'$ y $\mathbb{K}$, es decir,

$$\dim(X/\ker x') = 1.$$

□

Proposición 2.15. *Si X es un espacio normado y $M \subset X$ es un hiperplano, entonces M es cerrado o denso.*

Demostración. Sea $\overline{M}$, la clausura topológica de M. Entonces $\overline{M}$ es un subespacio vectorial (Proposición 1.3) y se tiene que

$$M \subset \overline{M} \subset X$$

Al ser $\dim X/M = 1$, se tiene que, o bien $M = \overline{M}$ (M es cerrado) o bien $\overline{M} = X$ (M es denso). □

A la vista de la división de los hiperplanos en densos y cerrados, y de la relación entre hiperplanos y formas lineales, la siguiente proposición resulta natural.

Proposición 2.16. *Sea X un espacio de Banach y $x' : X \longrightarrow \mathbb{K}$ una forma lineal. Entonces $\ker x'$ es cerrado si y sólo si x' es continua.*

Demostración. Claramente si x' es continua entonces $\ker x' = x'^{-1}(\{0\})$ es cerrado.

Recíprocamente, supongamos que $\ker x'$ es cerrado. Sea

$$Q : X \longrightarrow X/\ker x'$$

la aplicación cociente, que es continua (Ejercicio 2.2). Sea

$$\theta : X/\ker x' \longrightarrow \mathbb{K}$$

un isomorfismo algebraico, que siempre es continuo por ser una aplicación lineal entre espacios de dimensión finita (Corolario 2.2). Entonces

$$f = \theta \circ Q : X \longrightarrow \mathbb{K}$$

es continua y $\ker f = \ker x'$. Veamos finalmente que existe $\alpha \in \mathbb{K}$ tal que $x' = \alpha f$, y por tanto x' es continua.

Sea $x_0 \in X$ tal que $f(x_0) = 1$. Entonces $x'(x_0) = \alpha \neq 0$ (porque los núcleos de f y x' coinciden). Sea ahora $x \in X$ y sea $a = f(x)$. Entonces

$$f(x - ax_0) = f(x) - a1 = 0$$

y por tanto

$$x - ax_0 \in \ker f = \ker x',$$

es decir

$$x'(x - ax_0) = 0$$

de donde $x'(x) = ax'(x_0)$, es decir

$$x'(x) = f(x)x'(x_0) = \alpha f(x) \text{ para todo } x \in X.$$

□

EJEMPLO 2.17. *Veamos un ejemplo de un hiperplano denso: consideremos la sucesión* $(e_n) \subset c_0$. *Sea* $x^0 \in c_0$, $x^0 = (1, \frac{1}{2}, \frac{1}{3}, \ldots)$, *es decir* $x^0_n = \frac{1}{n}$. *Es fácil ver que* $\{x_0\} \cup \{e_n; n \in \mathbb{N}\}$ *es un sistema linealmente independiente. Sea entonces* $B = \{x_0\} \cup \{e_n; n \in \mathbb{N}\} \cup \{b_i; i \in I\}$ *una base de Hamel. Entonces todo* $x \in c_0$ *se puede escribir en forma única como*

$$x = \alpha_0 x_0 + \sum_n \alpha_n e_n + \sum_i \alpha_i e_i$$

(donde sólo una cantidad finita de coeficientes son no nulos). Definimos $x' : c_0 \longrightarrow \mathbb{K}$ *como* $x'(x) = \alpha_0$. *Así definido* x' *es una forma lineal obviamente no nula puesto que* $x'(x_0) = 1$. *Claramente* $c_{00} \subset \ker x'$ *y puesto que* c_{00} *es denso en* c_0 *se tiene que* $\ker x'$ *es denso.*

Prácticas sugeridas

EJERCICIO 2.1. *Dos espacios normados* X, Y *son isomorfos si y sólo si existe una aplicación lineal y biyectiva* $\theta : X \longrightarrow Y$ *y existen* $\alpha, \beta > 0$ *tales que, para todo* $x \in X$,

$$\alpha\|x\| \leq \|\theta(x)\| \leq \beta\|x\|.$$

EJERCICIO 2.2. *Sea* X *un espacio normado,* $Y \subset X$ *un subespacio vectorial cerrado. Entonces la aplicación cociente*

$$Q : X \longrightarrow X/Y$$

es continua.

EJERCICIO 2.3. *Sea* $1 \leq q < p \leq \infty$. *Probar que la identidad formal* $i : \ell_q \longrightarrow \ell_p$ *es continua.*

EJERCICIO 2.4. [**22**, p. 103] *Si* $(F_n) \subset L(C[0,1])$ *es una sucesión de operadores* positivos *tal que*

$$\lim_n F_n(f) = f$$

para $f = 1, x, x^2$ *entonces*

$$\lim_n F_n(f) = f$$

para todo $f \in C[0,1]$.

EJERCICIO 2.5. *Sea* $H = L_2$ *y sea* $C^{(1)}$ *el conjunto de las funciones con derivada continua. Sea* $t \in [0,1]$ *y sea* $D_t : C^{(1)} \longrightarrow \mathbb{K}$ *la forma dada por*

$$D_t(f) = f'(t).$$

Demostrar que D_t *no se puede extender con continuidad a* $L_2[0,1]$.

EJERCICIO 2.6 (Matriz de Hilbert). *Probar que la matriz* $A = (a_{i,j})_{i,j\in\mathbb{N}}$ *con*

$$a_{i,j} = \frac{1}{i+j}$$

define un operador acotado $T : \ell_2 \longrightarrow \ell_2$ *con* $\|T\| \leq \pi$.

Capítulo 3

Teoremas de Hahn-Banach

El Teorema de Hahn-Banach (o más bien *los* Teoremas de Hahn-Banach) juegan un papel fundamental en la teoría de Espacios Normados, así como en la teoría de Espacios Localmente Convexos. En su forma analítica, el Teorema de Hahn-Banach nos permite extender formas lineales definidas sobre subespacios y dominadas por funcionales sublineales o seminormas. Esto nos garantiza que los espacios duales son, en primer lugar, no vacíos, y en segundo lugar, suficientemente ricos como para determinar importante información estructural acerca de los espacios. En su forma geométrica, el Teorema de Hahn-Banach nos permite separar por medio de hiperplanos determinados pares de conjuntos del espacio y esto tiene una importancia fundamental, como se verá en las aplicaciones. Las dos formas del teorema están relacionadas ya que, como vimos en el capítulo anterior, hay una estrecha relación entre formas lineales e hiperplanos. Nosotros demostraremos en primer lugar la forma analítica, y de ahí deduciremos la forma geométrica. El lector interesado en leer una presentación de estos teoremas en otro orden puede consultar por ejemplo [**44**], [**26**] o [**31**].

La demostración de la versión analítica del Teorema que presentamos es la más habitual, y creemos que también la más elegante. En todo el capítulo hay que distinguir entre el caso real y el caso complejo, y así lo hacemos. Como consecuencia de la forma analítica del teorema obtenemos la existencia de *formas normantes*

Antes de la presentación de las formas geométricas del teorema, hemos incluido una breve presentación de los Espacios Localmente Convexos y de los Espacios Vectoriales Topológicos. No es estrictamente necesario hacerlo así, y de hecho en este curso sólo utilizamos una versión del Teorema de separación de Hahn-Banach en Espacios Localmente Convexos en la demostración del Teorema de Goldstine en el Capítulo 8, demostración que no es indispensable para la asignatura como comentaremos en su momento.

El motivo de que hayamos incluido esa pequeña sección sobre Espacios Localmente Convexos es doble: por un lado nos permite presentar el Teorema de separación de Hahn-Banach en su contexto más general, y por otro lado nos sirve de excusa para informarle al alumno de la existencia de tales espacios y que no todo espacio funcional es normable. De todas formas, podemos prescindir de esa sección y hacer una presentación algo menos general de los Teoremas de separación de Hahn-Banach tal y como se hace por ejemplo en [**24**] o [**31**].

Así como el Teorema de extensión de Hahn-Banach es esencialmente único en su enunciado, existen numerosas versiones más o menos generales o especializadas del Teorema de separación. La aparentemente más general y de la que se siguen muchas de las otras es nuestro Teorema 3.22, debido a Mazur. A continuación hemos presentado algunos de sus corolarios más útiles, pero hay muchas otras versiones posibles algunas de las cuales se pueden sugerir como ejercicios.

Seguimos con un estudio de la *reflexividad* en espacios normados y el estudio del dual del espacio cociente.

En este capítulo comienzan a ser muy relevantes las aplicaciones. Las de los Teoremas de Hahn-Banach son *muy* numerosas. Hemos incluido algunas de las “clásicas”, pero se podrían añadir muchas otras. Debido simplemente a nuestro interés personal en este momento hemos hecho una pequeña reseña de algunas aplicaciones a la Teoría Económica de estos teoremas, pero probablemente no serían las más adecuadas para contar en clase dado que requieren bastantes definiciones adicionales.

Para la presentación del Teorema de extensión de Hahn-Banach hemos seguido principalmente [**8**]. La breve introducción a los Espacios Localmente Convexos sigue [**13**] aunque también nos hemos servido de [**44**] y [**26**]. Seguimos de nuevo [**8**] para introducir las versiones geométricas del Teorema de Hahn-Banach. El lector interesado esencialmente en versiones geométricas del teorema para espacios normados puede consultar [**24**, Ejercicios al final del Capítulo 2].

El Teorema de Extensión de Hahn-Banach

Aunque en este curso usaremos el Teorema de Hahn-Banach mas frecuentemente en el contexto de los espacios normados, el teorema sigue siendo cierto, y se utiliza a menudo en el contexto de espacios vectoriales topológicos localmente convexos. Incluso en el contexto de los espacios normados, las versiones geométricas del teorema necesitan de las nociones de seminorma y funcional sublineal. Es por ello que antes de enunciar y demostrar el teorema de Hahn-Banach debemos definir seminormas y funcionales sublineales.

DEFINICIÓN 3.1. *Sea X un espacio vectorial. Una función*

$$p : X \longrightarrow [0, +\infty)$$

es una seminorma si verifica

1. $p(\lambda x) = |\lambda| p(x)$, *para todo* $x \in X$ *y* $\lambda \in \mathbb{K}$.
2. $p(x + y) \leq p(x) + p(y)$, *para todos* $x, y \in X$.

Es decir, una seminorma es exactamente igual a una norma salvo en que puede haber elementos $x \neq 0$ tales que $p(x) = 0$.

Aunque volveremos sobre ello más adelante, adelantamos que un espacio localmente convexo será un espacio vectorial cuya topología vendrá definida por medio de una familia de seminormas.

Necesitamos también definir funcionales sublineales.

DEFINICIÓN 3.2. *Sea X un espacio vectorial sobre $\mathbb{R}$ o $\mathbb{C}$. Un funcional sublineal p sobre X es una aplicación $p : X \longrightarrow \mathbb{R}$ positivamente homogénea y subaditiva, es decir p verifica*

- $p(\lambda x) = \lambda p(x)$ *para todo* $\lambda \in \mathbb{R}_+$, *y para todo* $x \in X$
- $p(x + y) \leq p(x) + p(y)$ *para todos* $x, y \in X$

Claramente, toda seminorma es un funcional sublineal, aunque no al revés. En particular una norma es un funcional sublineal.

Puesto que el Lema de Zorn es parte esencial de la demostración del Teorema de Hahn-Banach, incluimos aquí su enunciado.

LEMA 3.3 (Lema de Zorn). *Sea $A \neq \emptyset$ un conjunto ordenado, en el que todo subconjunto totalmente ordenado tiene una cota superior (en ese caso se dice que A es inductivo). Entonces todo elemento de A está mayorado por un elemento maximal (un elemento a es maximal si $a \leq b$ implica $a = b$)*

Las demostraciones (y los enunciados) del Teorema de Hahn-Banach son levemente distintos en el caso real y el caso complejo. Merece la pena mencionar que si bien el teorema en el caso real lo probaron independientemente Hahn (1926) y Banach (1927), la demostración del caso complejo no apareció hasta 1938 en un artículo de Bohnenblust y Sobczyk.

TEOREMA 3.4 (Hahn-Banach). *Sea X un espacio vectorial real, $p : X \longrightarrow \mathbb{R}$ un funcional sublineal, $M \subset X$ un subespacio vectorial de*

X y $f : M \longrightarrow \mathbb{R}$ una aplicación lineal dominada por p, es decir, para todo $x \in M$,

$$f(x) \leq p(x).$$

Entonces existe una aplicación lineal $\overline{f} : X \longrightarrow \mathbb{R}$ que extiende a f (es decir, tal que $\overline{f}(x) = f(x)$ para todo $x \in M$) y que sigue estando dominada por p, es decir, para todo $x \in X$,

$$\overline{f}(x) \leq p(x).$$

Demostración. Demostraremos primero que podemos extender f a "una dimensión" más. A continuación el Lema de Zorn operará su magia.

Sea $e \in X \setminus M$, sea $G = M \oplus [e]$. Entonces todo elemento $y \in G$ admite una única descomposición $y = x + \lambda e$, con $x \in M$, $\lambda \in \mathbb{R}$. Para todo $y \in G$ descompuesto en la forma indicada definimos

$$\overline{f}(y) = f(x) + \lambda c$$

Claramente $\overline{f} : G \longrightarrow \mathbb{R}$ es lineal y extiende a f. Sólo nos falta demostrar que podemos elegir $c = \overline{f}(e)$ de manera que $\overline{f} \leq p$ en G, es decir de manera que

$$\overline{f}(y) = f(x) + \lambda c \leq p(x + \lambda e) = p(y) \tag{2}$$

para todo $y \in G$.

Sabemos por hipótesis que (2) es cierta para $\lambda = 0$. Para $\lambda > 0$, dividiendo por λ, la ecuación (2) se puede escribir como

$$f\left(\frac{x}{\lambda}\right) + c \leq p\left(\frac{x}{\lambda} + e\right)$$

y por tanto queremos que

$$c \leq p\left(\frac{x}{\lambda} + e\right) - f\left(\frac{x}{\lambda}\right)$$

es decir

$$c \leq \inf_{x_1 \in M} \{p(x_1 + e) - f(x_1)\}$$

Para $\lambda < 0$, dividiendo por $-\lambda$, la ecuación (2) se puede escribir como

$$f(x_2) - c \leq p(x_2 - e), \qquad \text{con } x_2 = \frac{-x}{\lambda},$$

y por tanto queremos que

$$c \geq \sup_{x_2 \in M} \{f(x_2) - p(x_2 - e)\}$$

Es decir, podremos elegir un c que cumpla lo pedido si y sólo si

$$\sup_{x_2 \in M} \{f(x_2) - p(x_2 - e)\} \leq \inf_{x_1 \in M} \{p(x_1 + e) - f(x_1)\}$$

Esto es lo mismo que pedir que, para todo $x_1, x_2 \in M$,

$$f(x_2) - p(x_2 - e) \leq p(x_1 + e) - f(x_1)$$

es decir,

$$f(x_1) + f(x_2) \leq p(x_1 + e) + p(x_2 - e).$$

Pero por hipótesis sabemos que

$$f(x_1) + f(x_2) = f(x_1 + x_2) \leq p(x_1 + x_2) = p(x_1 + e + x_2 - e) \leq$$
$$\leq p(x_1 + e) + p(x_2 - e)$$

Por tanto efectivamente existe algún c (en general más de uno) que verifica lo pedido. Nótese que la no unicidad de c hace que tampoco la extensión sea única en general.

Ahora es cuando el lema de Zorn entra en acción: Consideramos el conjunto de pares

$$A := \{(H, f_H), \text{ donde } H \subset X \text{ es un subespacio vectorial}$$
$$\text{que contiene a } M \text{ y } f_H : H \longrightarrow \mathbb{R} \text{ es una aplicación lineal}$$
$$\text{que extiende a } f \text{ y está dominada por } p\}.$$

Podemos definir un orden natural para nuestro problema en A como

$$(H_1, f_{H_1}) \leq (H_2, f_{H_2}) \text{ si } H_1 \subset H_2 \text{ y } f_{H_2} \text{ extiende a } f_{H_1}.$$

Con este orden A es inductivo: en efecto, sea $(H_i, f_{H_i})_{i \in I}$ un conjunto totalmente ordenado. En ese caso el par $(\mathcal{H}, f_{\mathcal{H}})$ definido como

$$\begin{cases} \mathcal{H} = \cup_i H_i \\ f_{\mathcal{H}}(x) = f_{H_i}(x) \qquad \text{si } x \in H_i \end{cases}$$

pertenece a A y es una cota superior de $(H_i, f_{H_i})_{i \in I}$

Por tanto, puesto que A es inductivo el elemento $(M, f) \in A$ está mayorado por un elemento maximal, llamémosle (L, f_L). Si L no fuera el espacio total X, usando la primera parte de la demostración podríamos extender f a un subespacio L' estrictamente mayor que L de manera que (L, f_L) no sería maximal, una contradicción. Por tanto $L = X$ y $\overline{f} = f_L$ es la extensión buscada de f a todo X. $\square$

OBSERVACIÓN 3.5. *Merece la pena hacer aquí un par de observaciones. En primer lugar que en la demostración del Teorema no hemos usado ninguna topología, por lo que en particular f no tiene por qué ser continua ni F tiene por qué ser cerrado en X (puesto que ni siquiera hemos definido ninguna topología en X)*

En segundo lugar mencionar que este Teorema, con toda su potencia que tendremos ocasión de aplicar más adelante, tiene también una limitación que no conviene olvidar. El procedimiento de obtención de $\overline{f}$ es no constructivo, y además $\overline{f}$ no es única, ni tenemos forma de elegir "bien" una extensión entre todas las posibles (en particular no podemos en general seleccionar linealmente una extensión). Cuando veamos las versiones geométricas del teorema, será natural plantearse la relación entre la unicidad de la extensión y la geometría del espacio. Hay varios resultados en esta dirección. Un ejemplo de lo que se puede esperar es el Teorema de Taylor-Foguel que nos dice que, dado un espacio de Banach X, para todo $Y \subset X$ y para todo $y^* \in Y^*$ y^* tiene una extensión de Hahn-Banach única si y sólo si X^* es estrictamente convexo, es decir, si dados $f_1, f_2 \in S_{X^*}$, $\left\|\frac{f_1+f_2}{2}\right\| < 1$.

Hemos seguido la presentación más habitual del Teorema. Para ciertas aplicaciones puede ser interesante observar que el Teorema, con (esencialmente) la misma demostración, sigue siendo cierto si a p le pedimos tan sólo que sea convexa ([**1**]*). Es fácil ver que todo funcional sublineal es convexo.*

Damos a continuación como corolario la versión compleja del teorema.

TEOREMA 3.6 (Hahn-Banach). *Sea X un espacio vectorial real o complejo, $p : X \longrightarrow [0, +\infty)$ una seminorma, M un subespacio vectorial de X y $f : M \longrightarrow \mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ ó $\mathbb{C}$ según X sea real o complejo) una forma lineal dominada por p en módulo, es decir, para todo $x \in M$,*

$$|f(x)| \leq p(x).$$

Entonces existe una extensión $\hat{f} : X \longrightarrow \mathbb{K}$ de f también dominada por p en módulo.

DEMOSTRACIÓN. Supongamos primero que X es un espacio vectorial real. Puesto que las seminormas son funcionales sublineales, el teorema anterior nos dice que existe una extensión de f, $\hat{f} : X \longrightarrow \mathbb{R}$ tal que $\hat{f}(x) \leq p(x)$ para todo $x \in X$. Entonces, cambiando x por $-x$ se tiene que

$$\hat{f}(-x) \leq p(-x)$$

es decir
$$-\hat{f}(x) \leq p(x)$$
y por tanto
$$|\hat{f}(x)| \leq p(x)$$
para todo $x \in X$.

Si X es un espacio vectorial complejo, también lo es real, de manera que si $f : M \longrightarrow \mathbb{C}$ es una forma lineal, podemos definir la forma real
$$g(x) := \Re(f)(x) : M \longrightarrow \mathbb{R}$$
que podemos extender a $\hat{g}$ con $|g(x)| \leq p(x)$ para todo $x \in X$ por los razonamientos anteriores.

Ahora, la linealidad compleja de f nos dice que si $f(x) = \alpha + \imath\beta$ entonces
$$f(\imath x) = -\beta + \imath\alpha,$$
es decir, que para todo $x \in M$,
$$\Im(f)(x) = -g(\imath x) = -\Re(f)(\imath x)$$

Por tanto, la forma
$$\hat{f}(x) := \hat{g}(x) - \imath\hat{g}(\imath x) : X \longrightarrow \mathbb{C}$$
extiende a f. Para ver que $\hat{f}$ está dominada en módulo por p, dado $x \in X$, sea $\theta = \arg \hat{f}(x)$. Entonces
$$|\hat{f}(x)| = e^{-\imath\theta}\hat{f}(x) = \hat{f}(e^{-\imath\theta}x) = \Re\hat{f}(e^{-\imath\theta}x) =$$
$$= \hat{g}(e^{-\imath\theta}x) \leq p(e^{-\imath\theta}x) = p(x)$$
lo que demuestra el teorema. $\square$

El teorema de Hahn-Banach tiene muchísimas consecuencias y aplicaciones en al Análisis Funcional. La primera de ellas es que, puesto que los espacios de dimensión finita tienen funcionales controlados por la norma, todo espacio normado (y también todo espacio localmente convexo) tiene dual topológico no vacío. El siguiente corolario nos dice que no sólo el dual es no vacío, sino que hay suficientes elementos para normar cada vector de X. Enunciamos y demostramos el resultado en el caso más general de seminormas.

Corolario 3.7. *Sea X un espacio vectorial real o complejo, p una seminorma en X y $x \in X$. Entonces existe una forma lineal $f : X \longrightarrow \mathbb{K}$ tal que, para todo $y \in X$,*
$$|f(y)| \leq p(y)$$

y tal que

$$f(x) = p(x).$$

En particular, si X es un espacio normado entonces dado $x \in X$ existe una forma lineal y continua f de norma 1 tal que

$$f(x) = \|x\|$$

DEMOSTRACIÓN. Tomamos $M = [x]$ y definimos $f : M \longrightarrow \mathbb{K}$ como

$$f(\lambda x) = \lambda p(x)$$

para todo $\lambda \in \mathbb{K}$. Entonces el Teorema 3.6 nos dice que existe $\hat{f} : X \longrightarrow \mathbb{K}$ que cumple lo pedido. □

El siguiente resultado también es una consecuencia del teorema. Nos proporciona una "dualidad" entre la norma de los elementos de X y los de X^*.

COROLARIO 3.8. *Sea X un espacio normado. Entonces, para todo $x \in X$,*

$$\|x\| = \sup_{f \in B_{X^*}} |f(x)|$$

DEMOSTRACIÓN. En efecto,

$$|f(x)| \leq \|f\| \|x\|$$

y por tanto

$$\sup_{f \in B_{X^*}} |f(x)| \leq \|x\|$$

Por otro lado, tomando f_0 un funcional de norma 1 normante de x, como en el corolario anterior, tenemos

$$\|x\| = |f_0(x)| \leq \sup_{f \in B_{X^*}} |f(x)|$$

□

Espacios Localmente Convexos

El contexto natural para las versiones geométricas de los teoremas de Hahn-Banach es el de los espacios localmente convexos. Por ello antes de enunciar y probar dichas versiones geométricas dedicaremos esta sección a la definición y breve estudio de los Espacios Localmente Convexos (ELC) en particular.

Los espacios vectoriales topológicos (EVT) son la generalización natural de los espacios normados en el siguiente sentido: a menudo es necesario estudiar un espacio vectorial con una topología “bien relacionada” con la estructura lineal (en el sentido que vamos a precisar a continuación) pero dicho espacio *no* es normado. Buena parte -pero no toda- de la teoría (y de la intuición) de espacios normados se puede trasladar sin excesiva dificultad a este nuevo contexto

Empezamos con una definición.

Definición 3.9. *Un espacio vectorial topológico X es un espacio vectorial en el que se ha definido una topología que hace que*

1. *La aplicación suma* $+ : X \times X \longrightarrow X$ *es continua.*
2. *La aplicación producto* $\cdot : \mathbb{K} \times X \longrightarrow X$ *es continua.*

Ejercicio 3.1. *Sea X un EVT, y sea $x \in X$. Probar que la aplicación $T_x : X \longrightarrow X$ dada por*

$$T_x(y) = x - y$$

es un homeomorfismo (es decir, es biyectiva y continua en ambos sentidos).

Ya hemos visto que los espacios normados son EVT's.

No vamos a estudiar en esta memoria los EVT's en general, puesto que entendemos que dicho estudio es excesivamente especializado y que probablemente sea mejor que el contacto inicial de los alumnos con el Análisis Funcional ocurra en el contexto más “razonable” de los espacios de Banach. Sí queremos en esta sección estudiar sucintamente un tipo especial de EVT's, los Espacios Localmente Convexos (ELC), ya que son el contexto natural para los Teoremas geométricos de Hahn-Banach. Además, dado un espacio de Banach X, tanto (X, w) como (X^*, w^*), que estudiaremos más adelante, son ELC's. Por ello definimos a continuación Espacios Localmente Convexos, y a continuación probamos algunas propiedades básicas de los ELC que usamos más adelante. Cuando estas propiedades lo sean de todos los EVT, no pediremos la convexidad en el enunciado.

Los ELC's son aquellos EVT en los que cada punto posee una base de entornos formada por conjuntos convexos. Veremos más adelante cómo a partir de un conjunto convexo se puede definir su *funcional de Minkowski* que será una *seminorma*, de manera que probaremos la equivalencia entre tener una base de entornos convexos y tener una topología definida por seminormas. Elegimos de momento esta última como la definición de ELC, por parecernos más intuitiva inicialmente.

DEFINICIÓN 3.10. *Sea X un espacio vectorial y sea $\mathcal{P}$ una familia de seminormas definidas sobre X. Sea $\mathcal{T}$ la topología definida en X que tiene como subbase los conjuntos*

$$W_{(x_0;p;\epsilon)} := \{x \in X \text{ tales que } p(x - x_0) < \epsilon\}$$

para todo $x_0 \in X$, $p \in \mathcal{P}$ y $\epsilon > 0$, de manera que un conjunto $U \subset X$ es abierto si y sólo si para todo $x_0 \in X$ existe un entorno de x_0

$$W_{(x_0;p_1,\ldots,p_n;\epsilon)} := \{x \in X \text{ tales que } p_i(x - x_0) < \epsilon \text{ para todo } 1 \leq i \leq n\}$$

de manera que $W_{(x_0;p_1,\ldots,p_n;\epsilon)} \subset U$. Si además la topología $\mathcal{T}$ es Hausdorff, decimos que X es un Espacio Localmente Convexo.

Es un ejercicio (y como tal lo proponemos) verificar que todo ELC es un EVT. También es un ejercicio verificar que si X es un espacio vectorial con la topología dada por una familia de seminormas $\mathcal{P}$, entonces X es Hausdorff si y sólo si

$$\cap_{p \in \mathcal{P}} \{x \in X \text{ tales que } p(x) = 0\} = \{0\}.$$

Parafraseando [**1**], "if we had our way, all topological vector spaces would be Hausdorff". Hacemos nuestra tal afirmación, puesto que no tendremos ocasión de trabajar con EVT's no separados. En nuestras definiciones, si bien no le pedimos a los EVT que sean separados, sí se lo pedimos a los ELC.

Recordemos la definición de conjunto convexo.

DEFINICIÓN 3.11. *Sea X un espacio vectorial. Un conjunto $A \subset X$ es convexo si para todos $a, b \in A$, el segmento*

$$[a, b] = \{\lambda a + (1 - \lambda)b; \lambda \in [0, 1]\}$$

está contenido en A

Es fácil ver que A es convexo si y sólo si para todos $a_1, \ldots, a_n \in A$ y para todos $\lambda_1, \ldots, \lambda_n \in [0, 1]$ con $\sum_{i=1}^n \lambda_i = 1$ se tiene

$$\sum_{i=1}^n \lambda_i a_i \in A.$$

También es fácil ver que la intersección de conjuntos convexos es convexo. Puesto que además un espacio vectorial X siempre es convexo, tiene sentido definir la envoltura convexa de un conjunto.

Definición 3.12. *Sea X un espacio vectorial y sea $A \subset X$. Se define la envoltura convexa de A, que denotamos $co(A)$, como la intersección de todos los conjuntos convexos que contienen a A. Si además X es un EVT, se define la envoltura convexa y cerrada de A, y la denotamos como $\overline{co}(A)$, como la intersección de todos los conjuntos convexos y cerrados que contienen a A.*

Ejercicio 3.2. *Dado $A \subset X$, demostrar que*

$$co(A) = \{\sum_{i=1}^{n} \lambda_i a_i;\ n \in \mathbb{N},\ a_1, \ldots, a_n \in A,\ \lambda_1, \ldots, \lambda_n \in [0,1]\}.$$

Ejercicio 3.3. *Sean X, Y espacios vectoriales. Si $T : X \longrightarrow Y$ es una aplicación lineal y $C \subset Y$ es convexo entonces $T^{-1}(C)$ también es convexo.*

Necesitamos dos definiciones más.

Definición 3.13. *Sea X un espacio vectorial. Un subconjunto $A \subset X$ se dice* equilibrado *si $\alpha A \subset A$ para todo $\alpha \in \mathbb{K}$ tal que $|\alpha| \leq 1$.*

Obviamente todos los conjuntos equilibrados contienen al origen. La idea de su definición es una cierta "simetría respecto al origen".

Definición 3.14. *Sea X un espacio vectorial. Un subconjunto $A \subset X$ se dice absorbente si para todo $x \in X$ existe un $\alpha > 0$ tal que $x \in \lambda A$ para todo $\lambda \in \mathbb{K}$ tal que $|\lambda| > \alpha$.*

Además, si $a \in A$, A se dice absorbente en a si $A - a$ es absorbente.

Ejercicio 3.4. *Dado un espacio vectorial X y una seminorma $p : X \longrightarrow [0, \infty)$, comprobar que el conjunto*

$$A = \{x \in X \text{ tales que } p(x) < 1\}$$

es convexo, absorbente y equilibrado y que además A es absorbente en todos sus puntos.

El ejercicio anterior muestra que la "bola unidad" abierta asociada a una seminorma es siempre un conjunto convexo, absorbente en todos sus puntos y equilibrado. Lo interesante es que el recíproco también es cierto.

Proposición 3.15. *Sea X un espacio vectorial y sea $A \subset X$ un subconjunto convexo, absorbente en todos sus puntos y equilibrado. Entonces existe una única seminorma $p : X \longrightarrow [0, \infty)$ tal que $A = \{x \in X \text{ tales que } p(x) < 1\}$.*

Demostración. Definimos $p : X \longrightarrow [0,\infty)$ como

$$p(x) = \text{ínf}\{t \text{ tales que } t \geq 0 \text{ y } x \in tA\}.$$

Tenemos que ver que p está bien definida, que es una seminorma y que $A = \{x \in X \text{ tales que } p(x) < 1\}$. Puesto que A es absorbente, para todo $x \in X$ existe $t > 0$ tal que $x \in tA$, por lo que p está bien definida. Veamos que p es una seminorma. Claramente $p(0) = 0$ puesto que $0 \in A$. Veamos que $p(\alpha x) = |\alpha| p(x)$. Podemos suponer que $\alpha \neq 0$. Usando que A es equilibrado tenemos que

$$\begin{aligned} p(\alpha x) = \text{ínf}\{t \geq 0; \alpha x \in tA\} = \text{ínf}\{t \geq 0; x \in \frac{t}{\alpha}A\} = \\ = \text{ínf}\{t \geq 0; x \in \left|\frac{t}{\alpha}\right| A\} = \text{ínf}\{t \geq 0; x \in \frac{t}{|\alpha|}A\} = \\ = |\alpha|\, \text{ínf}\{\frac{t}{|\alpha|} \geq 0; x \in \frac{t}{|\alpha|}A\} = |\alpha| p(x). \end{aligned}$$

Para terminar de comprobar que p es una seminorma nos resta ver la propiedad triangular. Observemos como paso previo que si $\alpha, \beta \geq 0$ y $a, b \in A$ entonces

$$\alpha a + \beta b = (\alpha + \beta)\left(\frac{\alpha}{\alpha+\beta}a + \frac{\beta}{\alpha+\beta}b\right) \in (\alpha+\beta)A$$

por la convexidad de A. Ahora, sean $x, y \in X$ con $p(x) = \alpha$ y $p(y) = \beta$. Sea $\epsilon > 0$. Por definición de p,

$$x \in (\alpha + \epsilon)A$$

e

$$y \in (\beta + \epsilon)A.$$

Por tanto

$$x + y \in (\alpha + \epsilon)A + (\beta + \epsilon)A = (\alpha + \beta + 2\epsilon)A$$

y de aquí se sigue que

$$p(x + y) \leq \alpha + \beta + 2\epsilon = p(x) + p(y) + 2\epsilon.$$

Puesto que esto es cierto para todo $\epsilon > 0$ se tiene la desigualdad triangular.

Veamos ahora que

$$A = \{x \in X \text{ tales que } p(x) < 1\}.$$

Si $p(x) = \alpha < 1$ entonces para todo $\alpha < \beta < 1$ se tiene que $x \in \beta A \subset A$ (la última inclusión se sigue de que A es equilibrado). Por tanto

$$\{x \in X \text{ tales que } p(x) < 1\} \subset A.$$

Para la otra inclusión, sea $a \in A$. Entonces $p(a) \leq 1$. Puesto que A es absorbente en todos sus puntos, en particular lo es en a y por tanto existe $\lambda > 0$ tal que

$$a + \lambda a = y \in A.$$

Entonces $a = \frac{y}{1+\lambda}$ y de aquí se sigue que

$$p(a) = \frac{1}{1+\lambda} p(y) \leq \frac{1}{1+\lambda} < 1.$$

Sólo falta por tanto ver la unicidad de p. Sea $q : X \longrightarrow [0, \infty)$ tal que $A = \{x \in X \text{ tales que } q(x) < 1\}$.

Veamos que $p(x) \leq q(x)$ para todo $x \in X$. La otra desigualdad se prueba análogamente. Sea $x \in X$ y sea $\alpha = q(x)$. Para todo $\epsilon > 0$

$$q\left(\frac{x}{\alpha+\epsilon}\right) = \frac{\alpha}{\alpha+\epsilon} < 1.$$

Por tanto $\frac{x}{\alpha+\epsilon} \in A$ y de aquí se sigue que

$$\frac{p(x)}{\alpha+\epsilon} = p\left(\frac{x}{\alpha+\epsilon}\right) < 1.$$

Por tanto

$$p(x) < \alpha + \epsilon = q(x) + \epsilon.$$

Puesto que esto se tiene para todo $\epsilon > 0$, se sigue que $p \leq q$. □

El resultado anterior nos invita a formular una definición. Nótese previamente que sólo hemos usado que A fuera absorbente en cada punto para obtener el menor estricto en la relación $A = \{x \in X \text{ tales que } p(x) < 1\}$. En particular nótese que p también es una seminorma si sólo le pedimos a A que sea absorbente, equilibradoy convexo.

Definición 3.16. *Sea $A \subset X$ un conjunto convexo, absorbente y equilibrado. Entonces podemos definir el funcional de Minkowski de A, que a menudo llamaremos j, como la función $j : X \longrightarrow [0, \infty)$ dada por*

$$j(x) = \inf\{\lambda \text{ tales que } \lambda \geq 0 \text{ y } x \in \lambda A\}.$$

Del resultado anterior se sigue que en ese caso el funcional de Minkowski es una seminorma.

Lema 3.17. *Si X es un EVT y $A \subset X$ es un abierto entonces A es absorbente en todos sus puntos.*

Demostración. Para todo $a \in A$, se tiene que $A - a$ es entorno de 0, y por tanto es absorbente. □

Tal y como prometimos al principio de la sección, probamos que un ELC es precisamente un EVT con una base de entornos convexos.

PROPOSICIÓN 3.18. *Sea X un EVT y sea*

$$\mathcal{U} = \{A \subset X \text{ tales que } A \text{ es abierto, convexo y equilibrado }\}.$$

Entonces X es un ELC si y sólo si $\mathcal{U}$ es una base de entornos del origen.

DEMOSTRACIÓN. Si X es un ELC, según la Definición 3.10, los entornos de la forma $W_{(x_0;p_1,\dots,p_n;\epsilon)} := \{x \in X \text{ tales que } p_i(x - x_0) < \epsilon \text{ para todo } 1 \leq i \leq n\}$, que son claramente abiertos, convexos y equilibrados forman una base entornos del origen.

La otra implicación es más fácil. □

Más adelante necesitaremos este resultado.

PROPOSICIÓN 3.19. *Sea X un EVT. Sean $A \subset X$ un subconjunto compacto y $B \subset X$ un subconjunto cerrado con $A \cap B = \emptyset$. Entonces existe un entorno abierto del origen U tal que $A + U$ y $B + U$ no se intersecan.*

PROPOSICIÓN 3.20. *Sea X un EVT y $f : X \longrightarrow \mathbb{K}$ una forma lineal. Entonces son equivalentes:*

(i) *f es continua.*
(ii) *f es continua en algún punto.*
(iii) *f es continua en el origen.*
(iv) *$\ker f$ es cerrado.*

DEMOSTRACIÓN. La equivalencia entre (i), (ii) y (iii) se sigue del hecho de que las traslaciones en los EVT's son homeomorfismos.

Claramente (iii) implica (iv).

Veamos ahora que (iv) implica (iii). Supongamos que $\ker f$ es cerrado y sea $(x_\alpha)_{\alpha\in A} \subset X$ una red tal que $x_\alpha \to 0$. Sea también $u \in X$ tal que $f(u) = 1$. Si $f(x_\alpha) \not\to 0$ entonces podemos suponer (pasando a una subred en caso necesario) que existe $\epsilon > 0$ tal que $|f(x_\alpha)| > \epsilon$ para todo α. Sea

$$y_\alpha = u - \frac{f(u)}{f(x_\alpha)} x_\alpha$$

y observemos que para todo α se tiene que $y_\alpha \in \ker f$. Además claramente $y_\alpha \to u$. Por ser $\ker f$ cerrado se tiene que $u \in \ker f$ en contradicción con que $f(u) = 1$. Por tanto se ha de tener que $f(x_\alpha) \to 0$, lo que nos dice que f es continua en el origen. $\square$

Como ya vimos en la forma analítica del Teorema de Hahn-Banach, las seminormas y los funcionales sublineales juegan un papel destacado en su enunciado y demostración. Ya hemos visto que dado un EVT, las seminormas están asociadas a los conjuntos convexos, absorbentes y equilibrados. El siguiente resultado, cuya demostración no incluimos por ser totalmente análoga a la de la Proposición 3.15 nos dice que si no le pedimos al conjunto que sea equilibrado lo que obtenemos no es una seminorma sino un funcional sublineal. Recuérdese que ya hemos visto que los conjuntos abiertos son absorbentes.

PROPOSICIÓN 3.21. *Sea X un EVT y sea $A \subset X$ un subconjunto abierto y convexo que contiene al origen. Entonces el funcional de Minkowski de A, $j_A : X \longrightarrow [0, \infty)$ definido como*

$$j_A(x) = \inf\{\lambda \text{ tales que } \lambda \geq 0 \text{ y } x \in \lambda A\}$$

es un funcional sublineal, y $A = \{x \in X \text{ tales que } j_A(x) < 1\}$.

Antes de enunciar el teorema de Hahn-Banach en forma geométrica, debemos comprobar que los resultados básicos referidos a hiperplanos que probamos para espacios normados siguen siendo ciertos en el contexto de EVT's. En particular se prueba análogamente al caso de espacios normados que si X es un EVT y $H \subset X$ es un hiperplano entonces H es cerrado o denso. Que un hiperplano es el núcleo de una forma lineal sigue siendo obviamente cierto puesto que ahí los únicos razonamientos involucrados eran algebraicos, no topológicos. Por último necesitaremos que si H es un hiperplano núcleo de una forma lineal f entonces H es cerrado si y sólo si f es continua. Esto ya lo hemos probado en la Proposición 3.20.

Teoremas de separación de Hahn-Banach

El siguiente resultado se debe a Mazur.

TEOREMA 3.22 (Hahn-Banach, forma geométrica). *Sea X un espacio vectorial topológico (real o complejo), $A \subset X$ un conjunto convexo abierto no vacío, L un subespacio afín de X que no interseca a A. Entonces existe un hiperplano afín cerrado H que contiene a L y no interseca a A.*

Demostración. Demostramos primero el caso real. Suponemos que $0 \in A$ (si no es así, podemos considerar $A - a$ y $L - a$ para cualquier $a \in A$).

Por ser A un abierto convexo que contiene al origen es absorbente, y podemos definir su funcional de Minkowski $p : X \longrightarrow [0, \infty)$. Ya hemos visto que p es un funcional sublineal y que $A = \{x \in X \text{ tales que } p(x) < 1\}$.

Sea $x_0 \in L$ y consideremos el subespacio vectorial $F = L \oplus x_0$. Entonces F es el subespacio vectorial generado por L. Puesto que, por hipótesis, $0 \notin L$, L es un hiperplano afín de F.

Por tanto existe una forma (no necesariamente continua) $f : F \longrightarrow \mathbb{R}$ tal que $L = \{x \in F; f(x) = 1\}$. Por tanto $f(x) \leq p(x)$ para todo $x \in L$ (porque para todo $x \in L$ se tiene que $x \notin A$ y por tanto $p(x) \geq 1$). Veamos que también $f(x) \leq p(x)$ para todo $x \in F$: Sea $x \in F$. Si $f(x) \leq 0$, automáticamente se sigue que $f(x) \leq p(x)$. Si, en cambio $f(x) > 0$ consideramos el elemento $\frac{x}{f(x)} \in F$. Puesto que $f\left(\frac{x}{f(x)}\right) = 1$, se tiene que $\frac{x}{f(x)} \in L$ y entonces

$$\frac{p(x)}{f(x)} = p\left(\frac{x}{f(x)}\right) \geq 1,$$

de donde

$$f(x) \leq p(x).$$

Entonces podemos aplicar la forma analítica del Teorema de Hahn-Banach, caso real, puesto que p es un funcional sublineal, y tenemos que existe una forma lineal $\hat{f} : X \longrightarrow \mathbb{R}$ que extiende f y que está dominada por p ($\hat{f}(x) \leq p(x)$ para todo $x \in X$).

Sea ahora el hiperplano afín $H := \{x \in X; \hat{f}(x) = 1\}$. Claramente contiene a L, puesto que $\hat{f}$ extiende a f. Además $H \cap A = \emptyset$ porque si $y \in H$ entonces $p(y) \geq \hat{f}(y) = 1$ y por tanto $y \notin A$. Para acabar, veamos que H es cerrado: Es un ejercicio comprobar que, al igual que ocurría con los hiperplanos vectoriales, un hiperplano afín es cerrado o denso en X. Pero H no puede ser denso puesto que no interseca al abierto A.

Ahora supongamos que X es un espacio vectorial complejo. Usando los razonamientos anteriores podemos hallar un hiperplano afín *real* H_0 que contiene a L y que no interseca a A. Sabemos que existe $x_0 \in L$ tal que $H_0 = x_0 + H_0'$, con H_0' un hiperplano *vectorial* real. Entonces $H' := H_0' \cap \imath H_0'$ es un hiperplano vectorial complejo y $H = x_0 + H'$ es un hiperplano afín complejo que no interseca a A y que contiene a L, ya que $L = x_0 + L'$, donde L' es un espacio vectorial complejo y por tanto $\imath L' = L'$. $\square$

Se pueden dar ahora una amplia variedad de corolarios de la forma geométrica del Teorema. Propondremos alguno a continuación y otros irán en los ejercicios. Es conveniente mencionar que muy a menudo es necesario fabricarse teoremas de separación "tipo Hahn-Banach" a medida para el problema en que se está trabajando. Esto ocurre a menudo, por ejemplo, en Economía Matemática, donde los teoremas geométricos de Hahn-Banach se usan profusamente pues son parte esencial de la demostración de los Teoremas del Estado del Bienestar sobre los que se basa casi toda la Teoría Económica y Financiera moderna. Un ejemplo (entre muchos) de esta construcción *ad hoc* de un teorema no trivial "tipo Hahn-Banach" se puede ver en [**43**].

COROLARIO 3.23. *Sea X un EVT real, $A \subset X$ un subconjunto abierto convexo no vacío y $B \subset X$ un subconjunto convexo no vacío tal que $A \cap B = \emptyset$. Entonces existe una forma lineal y continua $f \in X^*$ y un $\alpha \in \mathbb{R}$ tales que, para todos $x \in A$, $y \in B$,*

$$f(x) < \alpha \leq f(y),$$

es decir, existe un hiperplano cerrado afín ($f^{-1}(\{\alpha\})$) que separa A y B.

DEMOSTRACIÓN. El conjunto

$$C = A - B := \{a - b; a \in A, b \in B\} = \cup_{b \in B}(A - b)$$

(no $A \setminus B$!!) es abierto, por ser unión de abiertos, y convexo (si $c_1 = x_1 - y_1 \in C$, $c_2 = x_2 - y_2 \in C$, entonces $\lambda c_1 + (1-\lambda)c_2 = \lambda x_1 + (1-\lambda)x_2 - \lambda y_1 + (1-\lambda)y_2 \in C$ por la convexidad de A y B). Además C es no vacío y no contiene al origen. Usando el teorema podemos encontrar un hiperplano afín cerrado que pasa por el origen (es decir, un hiperplano vectorial cerrado) que no corta a C. Usando la convexidad de C podemos encontrar $f \in X^*$ tal que $f(x) < 0$ para todo $c \in C$. Por tanto, para todo $x \in A$, $y \in B$

$$f(x - y) = f(x) - f(y) < 0$$

y por tanto, si llamamos

$$\alpha = \inf_{y \in B} f(y)$$

tenemos, para todo $x \in A$, $y \in B$

$$f(x) \leq \alpha \leq f(y)$$

pero ahora, como A es abierto, se tiene que $f(x) < \alpha$ para todo $x \in A$, ya que si no fuera así, existiría $x_0 \in A$ tal que $f(x_0) = \alpha$. Supongamos por ejemplo que $\alpha > 0$. Por la continuidad del producto, existiría $\epsilon > 0$ tal que $(1+\epsilon)x_0 \in A$, y tendríamos que $f((1+\epsilon)x_0) \geq \alpha$. Si $\alpha < 0$ se razona análogamente. □

El mismo corolario admite una versión compleja

Corolario 3.24. *Sea X un EVT complejo, $A \subset X$ un subconjunto abierto convexo y equilibrado y $B \subset X$ un subconjunto convexo no vacío tal que $A \cap B = \emptyset$. Entonces existe una forma (compleja) lineal y continua $f \in X^*$ y un $\alpha > 0$ tales que, para todos $x \in A$, $y \in B$,*

$$|f(x)| < \alpha \leq |f(y)|,$$

es decir, existe un hiperplano cerrado afín ($f^{-1}(\{\alpha\})$) que separa A y B.

Demostración. [**8**, Corollary 4, p. 32] □

Finalmente damos las versiones real y compleja de un teorema de separación bastante utilizado. Lo enunciamos y probamos en el caso de espacios localmente convexos.

Corolario 3.25. *Sea X un ELC real. Sea $A \subset X$ un subconjunto compacto, convexo no vacío y $B \subset X$ un subconjunto cerrado convexo no vacío que no interseca a A. Entonces existe un hiperplano afín cerrado que separa estrictamente A y B, esto es, existe $f \in X^*$ y $\alpha \in \mathbb{R}$ tales que*

$$\sup_{x\in A} f(x) < \alpha < \sup_{y\in B} f(y)$$

Demostración. Por la Proposición 3.19 sabemos que existe un entorno abierto del origen U tal que $A+U$ y $B+U$ cumplen las hipotesis del Corolario 3.23 y por tanto existen $f \in X^*$ y $\alpha \in \mathbb{R}$ tales que, para todos $x \in A+U$, $y \in B+U$,

$$f(x) < \alpha < f(y),$$

la última desigualdad es estricta porque B es abierto, razonando como lo hicimos en el Corolario 3.23 para A. □

La versión compleja del Corolario es la siguiente

Corolario 3.26. *Sea X un ELC complejo, Hausdorff. Sea $A \subset X$ un subconjunto compacto, convexo equilibrado no vacío y $B \subset X$ un subconjunto cerrado convexo no vacío que no interseca a A. Entonces* *existe un hiperplano afín cerrado que separa estrictamente A y B, esto es, existe $f \in X^*$ y $\alpha > 0$ tales que para todos $x \in A$, $y \in B$,*

$$|f(x)| < \alpha < |f(y)|$$

DEMOSTRACIÓN. [**8**, Corollary 6 p. 33] □

Reflexividad

Notemos que si tenemos un espacio normado X podemos definir una aplicación canónica

$$J : X \longrightarrow X^{**}$$

dada por la relación

$$J(x)(x^*) = x^*(x)$$

para todos $x \in X$, $x^* \in X^*$. Es un sencillo ejercicio comprobar que J está bien definida (es decir, que para todo $x \in X$ la aplicación $J(x) : X^* \longrightarrow \mathbb{K}$ es lineal y continua). Además, del Teorema de Hahn-Banach se sigue que

$$\|J(x)\| = \sup_{x^* \in B_{X^*}} |J(x)(x^*)| = \sup_{x^* \in B_{X^*}} |x^*(x)| = \|x\|$$

y por tanto J es una isometría (no necesariamente sobreyectiva).

Si X es un espacio de Banach se sigue fácilmente que $J(X) \subset X^{**}$ es un espacio cerrado. Si X no es un espacio de Banach, entonces $\overline{J(X)} \subset X^{**}$ es un subespacio cerrado en X^{**}. Puesto que X^{**} *sí* es un espacio de Banach se sigue que $\overline{J(X)}$ es la compleción de X (en X^{**}). Esta compleción es única en el siguiente sentido:

Si X_1 es otro espacio de Banach tal que $X \subset X_1$ y X es denso en X_1 entonces X_1 y $\overline{J(X)}$ son isométricos. Veámoslo. Sea

$$\theta : J(X) \longrightarrow X_1$$

la aplicación dada por

$$\theta(J(x)) = x.$$

θ es claramente una isometría y por tanto se extiende por densidad a una isometría

$$\tilde{\theta} : \overline{J(X)} \longrightarrow X_1.$$

Sólo nos falta comprobar que $\tilde{\theta}$ es sobreyectiva. Sea $x_1 \in X_1$. Puesto que X es denso en X_1, existe una sucesión $(x_n)_n \subset X$ tal que $x_n \to x_1$.

Entonces la sucesión $(J(x_n))_n \subset J(X)$ es de Cauchy y por tanto existe $y \in \overline{J(X)}$ tal que $J(x_n) \to y$. Entonces

$$\tilde{\theta}(J(x_n)) \to \tilde{\theta}(y)$$

y también

$$\tilde{\theta}(J(x_n)) \to x_1$$

por lo que $x_1 = \tilde{\theta}(y)$ lo que termina la demostración.

Decimos que un espacio X es *reflexivo* si la aplicación canónica $J : X \longrightarrow X^{**}$ es sobreyectiva. Veremos más adelante que si $1 < p < \infty$ entonces ℓ_p y $L_p[0,1]$ son espacios reflexivos. En cambio, veremos también que c_0 y ℓ_1 no son reflexivos.

Un espacio es reflexivo si la inyección canónica $J : X \longrightarrow X^{**}$ es sobreyectiva. En **[27]**, James muestra un ejemplo de un espacio no reflexivo X tal que existe una isometría sobreyectiva (no la canónica, naturalmente) $\theta : X \longrightarrow X^{**}$.

Dual de un espacio cociente

En la siguiente sección probamos dos resultados útiles a menudo. Ambos pueden ser considerados simples ejercicios y, con alguna sugerencia, pueden ser resueltos por los propios alumnos.

Proposición 3.27. *Sea X un espacio de Banach y $M \subset X$ un subespacio vectorial cerrado. Si $M^\perp = \{x^* \in X^*$ tales que $x^*(x) = 0$ para todo $x \in M\}$ entonces*

$$X^*/M^\perp = M^*$$

donde el igual en la línea de arriba denota que ambos espacios son linealmente isométricos.

Demostración. Sea $\theta : X^*/M^\perp \longrightarrow M^*$ la aplicación dada por

$$\theta([x^*]) = x^*_{|M},$$

es decir, para todo $m \in M$,

$$\theta([x^*])(m) = x^*(m).$$

θ está bien definida y es lineal.

Si $\theta([x^*]) = 0$ entonces $x^* \in M^\perp$ y por tanto $[x^*] = 0$, por lo que θ es inyectiva.

Además, dado $m^* \in M^*$, podemos extenderlo por Hahn-Banach a $x^* \in X^*$. Es claro que $\theta([x^*]) = m^*$ por lo que θ es sobre.

Veamos que es una isometría. Sea $x^* \in X^*$ y sea $m^\perp \in M^\perp$. Entonces

$$\|\theta([x^*])\| = \|\theta([x^* + m^\perp])\| = \sup_{m \in B_M} \|x^*(m) + m^\perp(m)\| \leq \|x^* + m^\perp\|,$$

y tomando ínfimos en $m^\perp \in M^\perp$ se tiene que

$$\|\theta([x^*])\| \leq \|[x^*]\|.$$

Para la otra desigualdad, dado $x^* \in X^*$, sea $m^* \in M^*$ dado por $m^* = X^*_{|_M}$ y sea y^* una extensión de Hahn-Banach de m^* con la misma norma ($\|y^*\| = \|m^*\|$). Entonces $[x^*] = [y^*]$, $m^* = \theta([x^*]) = \theta([y^*])$ y

$$\|\theta([x^*])\| = \|m^*\| = \|y^*\| \geq \|[y^*]\| = \|[x^*]\|,$$

lo que termina la demostración. $\square$

Proposición 3.28. *Sea X un espacio de Banach y $M \subset X$ un subespacio vectorial cerrado. Entonces*

$$(X/M)^* = M^\perp$$

donde de nuevo el signo igual denota que los espacios son linealmente isométricos.

Demostración. Sea $Q : X \longrightarrow X/M$ la aplicación cociente y sea

$$\theta : (X/M)^* \longrightarrow M^\perp$$

la aplicación dada por

$$\theta(y^*) = y^* \circ Q.$$

Veamos en primer lugar que θ termina en $M^\perp$: Para todo $m \in M$ y para todo $y^* \in (X/M)^*$,

$$\theta(y^*)(m) = y^* \circ Q(m) = y^*(0) = 0.$$

Es claro que θ es lineal. Veamos que es una isometría:

Para todo $y^* \in (X/M)^*$,

$$\|\theta(y^*)\| = \|y^* \circ Q\| \leq \|Q\|\|y^*\| = \|y^*\|$$

lo que nos da una de las desigualdades.

Para la otra, dado $y^* \in (X/M)^*$ y dado $\epsilon > 0$ sea $[x] \in X/M$ tal que $\|[x]\| < 1$ y tal que

$$|y^*[x]| \geq \|y^*\| - \epsilon.$$

Por la definición de la norma cociente existe $m \in M$ tal que

$$\|x + m\| \leq 1$$

y entonces

$$\|\theta(y^*)\| = \|y^* \circ Q\| \geq |(y^* \circ Q)(x + m)\| =$$

$$= |(y^* \circ Q)(x)\| = |y^*([x])| \geq \|y^*\| - \epsilon,$$

y de aquí se sigue que, para todo $y^* \in (X/M)^*$,

$$\|\theta(y^*)\| \geq \|y^*\|$$

por lo que θ es una isometría (y, en consecuencia, inyectiva).

Sólo falta ver que θ es sobreyectiva. Para ello, sea $m^\perp \in M^\perp$. Sea $y^* : X/M \longrightarrow \mathbb{K}$ la forma dada por

$$y^*([x]) = m^\perp(x).$$

y^* está bien definida, es lineal, y es continua porque, para todo $m \in M$,

$$|y^*([x])| = |m^\perp(x)| = |m^\perp(x + m)| \leq \|m^\perp\| \|x + m\|$$

y por tanto, tomando ínfimos en m, se tiene que

$$|y^*(x)| \leq \|m^\perp\| \|[x]\|.$$

Por tanto $y^* \in (X/M)^*$, y claramente $\theta(y^*) = m^\perp$, lo que termina la demostración. $\square$

Aplicaciones

Las aplicaciones de los Teoremas de Hahn-Banach tanto a problemas del Análisis Funcional como a problemas más "aplicados" son demasiado numerosos para incluir en esta memoria ni siquiera una cantidad significativa de ellos. Nos limitaremos a mencionar algunas de estas aplicaciones.

Límites de Banach: Demostrar que existe un funcional lineal L definido sobre ℓ_∞ tal que

- $\|L\| = 1$
- Si $x = (x_n) \in c$ entonces $L(x) = \lim_{n\to\infty} x_n$
- Si $x = (x_n) \in \ell_\infty$ y $x_n \geq 0$ para todo $n \in \mathbb{N}$ entonces $L(x) \geq 0$.
- Si $x = (x_n) \in \ell_\infty$ y llamamos $x' = (x_2, x_3, \ldots)$ entonces $L(x') = L(x)$

Esta es una de las aplicaciones más clásicas del Teorema de Hahn-Banach y se puede consultar en muchos libros. Dos presentaciones adecuadas para este curso pueden ser las de [**13**, p. 82] o [**24**, Ejercicio 19, p.39].

Existencia de medidas sobre el círculo unidad invariantes por rotaciones: S. Banach estaba intentando resolver este problema cuando probó el Teorema de Hahn-Banach (y a continuación lo aplicó para resolver el problema). Se puede ver una presentación por ejemplo en [**22**].

Sea $\mathbb{T} = \{z \in \mathbb{C} \text{ tales que } |z| = 1\}$. Entonces se tiene.

TEOREMA 3.29. *Existe una medida finitamente aditiva invariante por rotaciones definida sobre la σ-álgebra de todas las partes de $\mathbb{T}$. Es decir, existe una función*

$$m : \mathcal{P}(\mathbb{T}) \longrightarrow [0,1]$$

tal que

- $m(\mathbb{T}) = 1$.
- *Para todos $S, T \in \mathcal{P}(\mathbb{T})$ con $S \cap T = \emptyset$ se tiene*
$$m(S \cup T) = m(S) + m(T).$$
- *Para toda rotación $f : \mathbb{T} \longrightarrow \mathbb{T}$ (dada por $f(z) = ze^{\imath\theta}$) y para todo $S \in \mathcal{P}(\mathbb{T})$ se tiene*
$$m(f(S)) = m(S).$$

Aplicaciones Económicas Aunque probablemente no se puedan desarrollar en clase porque nos separaría excesivamente del núcleo de la asignatura, si los alumnos mostraran interés en ello se podrían comentar aquí algunas aplicaciones de los Teoremas de Hahn-Banach a la Matemática Económica. Hay una gran cantidad de tales aplicaciones, mencionaremos aquí brevemente las dos que consideramos principales. La primera son los Teoremas del Estado del Bienestar, o Teoremas de Arrow-Debreu, demostrados en 1954 ([**3**] que merecieron un Premio Nobel y que forman la base de la Teoría Económica moderna. La segunda es el llamado Teorema Fundamental de la Valoración de Activos, uno de los pilares de la Matemática Financiera, y pieza clave en los trabajos sobre la valoración de opciones de Black y Scholes [**9, 10**] y [**34**], que también resultaron premiados con un Premio Nobel.

- **Teoremas del estado del bienestar.** Se pueden consultar en cualquier texto riguroso de Microeconomía. Una presentación atractiva para un matemático se puede ver por ejemplo en [**17**]. No podemos introducir ahora toda la notación necesaria para enunciar y demostrar los teoremas, pero mencionaremos simplemente que en uno de los pasos se separan dos conjuntos por el Teorema de separación de Hahn-Banach, y la forma del dual asociada al hiperplano separador se puede interpretar como un "vector de precios" que a cada "vector de bienes" le asocia su valor.
- **Teorema fundamental de la valoración de activos.** Este es uno de los resultados más utilizados en las técnicas de valoración de Derivados Financieros, y nos dice que la ausencia de arbitraje, una hipótesis muy razonable económicamente sobre los mercados es equivalente a la existencia de una medida en cierto espacio de probabilidad con respecto de la cual ciertos procesos de precios forman una martingala. Si bien las formas finito-dimensionales del teorema son razonablemente simples, las formas más generales son de una gran complejidad, no económica sino matemática. El lector interesado en este tema puede consultar por ejemplo [**14**] y la bibliografía que allí aparece.

Teorema bipolar Sea X un espacio normado y sea $M \subset X$ y $N \subset X^*$ dos conjuntos no vacíos. Definimos la polar de M como

$$M^0 = \{x^* \in X^* \text{ tales que } |x^*(x)| \leq 1 \text{ para todo } x \in M\}$$

y la polar de N como

$$N^0 = \{x \in X \text{ tales que } |x^*(x)| \leq 1 \text{ para todo } x^* \in N\}.$$

Obviamente hay una cierta incoherencia en la definición tal y como la hemos presentado, puesto que N^0 debiera ser un subconjunto de X^{**}; para ser más estrictos debiéramos haber llamado a N^0 algo así como "prepolar", y haber utilizado una notación como ${}^0N$ o algo así, como se hace en algunos textos. En realidad el contexto natural para la noción de polar de un conjunto es el de los *pares duales*, y en ese contexto se reduce la ambigüedad, pero no vamos a incluir en esta memoria la noción de par dual, por lo que nos limitamos a presentar estas ideas en el contexto de los espacios normados.

EJERCICIO 3.5. *Para todos M y N como en la definición M^0 y N^0 son conjuntos cerrados, convexos y equilibrados.*

La noción de polar está muy relacionada con la noción de *anulador* de un subespacio (que se corresponde con el *subespacio ortogonal* en espacios de Hilbert).

DEFINICIÓN 3.30. *Sea $M \subset X$ un subespacio. Entonces definimos su anulador $M^{\perp}$ como*

$$M^{\perp} = \{x^* \in X^* \text{ tales que } x^*(x) = 0 \text{ para todo } x \in M\}$$

EJERCICIO 3.6. *Si $M \subset X$ es un subespacio entonces $M^0 = M^{\perp}$*

EJERCICIO 3.7. *Si $B \subset A$ entonces $A^0 \subset B^0$.*

Con esto ya podemos probar el siguiente teorema, conocido como Teorema (de la) bipolar. También es válido en el contexto más amplio de par dual.

TEOREMA 3.31. *Sea X un espacio normado, $M \subset X$. Entonces*

$$M^{00} := (M^0)^0 = \overline{coe}(M)$$

donde $\overline{coe}(M)$ es la envoltura convexa equilibrada y cerrada de M.

DEMOSTRACIÓN. Sea $A = \cap_{i \in I} A_i$ donde $\{A_i; i \in I\}$ son todos los conjuntos cerrados convexos y equilibrados que contienen a M. Hay que probar que $M^{00} = A$.

En primer lugar, puesto que M es cerrado convexo y equilibrado se tiene que $A \subset M^{00}$.

Para el otro contenido, sea $x_0 \notin A$. Como A es cerrado y convexo, usando el Teorema de Hahn-Banach se tiene que existe $x^* \in X^*$, $\alpha \in \mathbb{R}$ y $\epsilon > 0$ tales que, para todo $a \in A$

$$\Re x^*(a) < \alpha < \alpha + \epsilon < \Re x^*(x_0).$$

Por ser A convexo y equilibrado sabemos que $0 \in A$, y por tanto

$$0 = x^*(0) < \alpha.$$

Entonces, sustituyendo x^* por $\frac{x^*}{\alpha}$ se tiene que existe $\varepsilon > 0$ tal que, para todo $a \in A$,

$$\Re x^*(a) < 1 < 1 + \varepsilon < \Re x^*(x_0).$$

Si $a \in A$ y $x^*(a) = |x^*(a)| e^{\imath\theta}$ entonces $e^{-\imath\theta} a \in A$ (por ser A equilibrado) y por tanto

$$|x^*(a)| = \Re x^*(e^{-\imath\theta} a) < 1 < 1 + \varepsilon < \Re x^*(x_0).$$

Puesto que eso ocurre para todo $a \in A$, se sigue que $x^* \in A^0$. Notemos que $M \subset A$. Por tanto $A^0 \subset M^0$, así que $x^* \in M^0$. Pero esto implica que $x_0 \notin M^{00}$. Es decir, hemos probado que el complementario de A está contenido en el complementario de M^{00}, es decir, que $M^{00} \subset A$. □

El lector de estas notas interesado en aplicaciones económicas de los teoremas de Hahn-Banach, puede consultar [**4**] para ver una curiosa aplicación del Teorema Bipolar a lo que Delbaen llama "coherent risk measures", medidas del riesgo de un agente económico "coherentes" en cierto sentido.

Prácticas sugeridas

Ejercicio 3.8. *Comprobar que todo ELC es un EVT.*

Ejercicio 3.9. *Sea X un EVT y $x_0 \in X$. Entonces la aplicación $t_{x_0} : X \longrightarrow X$ dada por $t_{x_0}(x) = x + x_0$ es biyectiva, continua y de inversa continua (es decir, es un homeomorfismo). Como consecuencia, los trasladados de una base de entornos del origen forman una base de entornos de cualquier punto. Análogamente, para todo $0 \neq \alpha \in \mathbb{K}$ la aplicación $h_\alpha : X \longrightarrow X$ definida por $h_\alpha(x) = \alpha x$ también es un homeomorfismo.*

Ejercicio 3.10. *Sea X un espacio vectorial complejo. Si $F_1 : X \longrightarrow \mathbb{R}$ es una forma lineal considerando X como espacio vectorial real, demostrar que $F : X \longrightarrow \mathbb{C}$ definida como*

$$F(x) = F_1(x) - \imath F_1(\imath x)$$

es una forma lineal sobre X considerado como espacio vectorial complejo.

Ejercicio 3.11. *Si X es un espacio de Banach de dimensión infinita, demostrar que existen conjuntos convexos C_1 y C_2 tales que $C_1 \cup C_2 = X$, $C_1 \cap C_2 = \emptyset$ y tanto C_1 como C_2 son densos en X.* **Sugerencia:** *Considérese una forma lineal no continua y los dos semiespacios en que divide a X.*

Ejercicio 3.12. *Demostrar que un espacio de Banach X es reflexivo si y sólo si X^* es reflexivo.*

Ejercicio 3.13. *Demostrar que si un espacio de Banach X es reflexivo entonces todo subespacio cerrado $Y \subset X$ también es reflexivo.*

Capítulo 4

El Teorema de Baire y sus consecuencias: El Principio de Acotación Uniforme y el Teorema de la Gráfica Cerrada.

Los Teoremas de la Aplicación Abierta, de la Gráfica Cerrada, el de Banach-Steinhaus y el principio de Acotación uniforme, junto con el Teorema de Hahn-Banach ya estudiado forman la base de la Teoría de Espacios de Banach y son herramientas indispensables para el estudio de los operadores entre espacios de Banach. Los estudiamos en este capítulo. Todos ellos tienen en común el que su demostración más usual los deriva del Teorema de Baire. Quizás merezca la pena mencionar que varias de las consecuencias de estos teoremas se pueden demostrar directamente por métodos de joroba deslizante.

La importancia de estos resultados se manifiesta en que en todo el resto de la memoria se usarán frecuentemente.

Una posibilidad a la hora de presentar estos resultados es utilizar el lenguaje de las *categorías* de Baire para obtener con poco más esfuerzo versiones algo más finas de estos resultados. Sin embargo, dado que no vamos a necesitar esas versiones más finas en este curso, hemos optado por no usar dicho lenguaje de categorías para mantener una presentación más simple.

El Principio de Acotación Uniforme nos informa acerca de la posibilidad de garantizar que una familia de operadores es "uniformemente acotada" sabiendo que es "puntualmente acotada". Antes de enunciar formalmente ese teorema veamos otro resultado muy conocido y útil, el Teorema de Arzela-Ascoli, que también nos habla, aunque en otro contexto, de la posibilidad de obtener una acotación "uniforme" partiendo de una acotación "puntual".

Seguimos el capítulo estudiando las otras dos aplicaciones antes mencionadas del Teorema de Baire al Análisis Funcional, los *Teoremas de la Gráfica Cerrada y de la Aplicación Abierta.*

Como una primera aplicación de estos resultados estudiamos las *proyecciones* en espacios de Banach, lo que nos da pie a definir *subespacios complementados*. No es sencillo en general demostrar si un determinado subespacio está o no complementado, ni es obvio presentar ejemplos de subespacios no complementados sin usar resultados profundos. Por ello, mostramos en este capítulo que c_0 no está complementado en ℓ_∞.

Terminamos el capítulo viendo otro de los teoremas fundamentales del Análisis Funcional, aunque su demostración no está relacionada con los anteriores, el Teorema de Stone-Weierstrass. Por su importancia en muchas aplicaciones, creemos necesario enunciarlo y probarlo en este curso. Su principal aplicación en este curso la veremos en el Capítulo 11, cuando se utilizará para estudiar Análisis de Fourier en el marco de la Teoría de espacios de Hilbert.

Con el paso del tiempo han ido apareciendo diferentes demostraciones de este teorema fundamental. Entre ellas, una demostración "directa", utilizando los polinomios de Bernstein, una demostración utilizando el Teorema de Korovkin (véase por ejemplo [**31**, p. 40]), o una demostración razonablemente corta, pero que utiliza el Teorema de Alaoglu, el Teorema de Krein-Milman y la descripción del dual de $C(K)$ como el espacio de medidas de Radon (véase por ejemplo [**13**, p. 145]). Finalmente hemos elegido la demostración seguida en [**12**, p. 137] puesto que no utiliza resultados profundos adicionales y a la vez resulta, creemos, razonablemente instructiva y sencilla.

Al igual que ocurría con el Teorema de Hahn-Banach, son innumerables las aplicaciones de los resultados de este capítulo al Análisis Funcional y a otras áreas de la matemática. Presentamos brevemente al final del capítulo algunas de las aplicaciones que nos han parecido más adecuadas para este curso.

Para la preparación de este capítulo hemos seguido sobre todo [**31**] y [**24**], con la adición de [**12**] para el estudio del Teorema de Stone-Weierstrass.

El Teorema de Ascoli-Arzela

Antes de enunciar y probar el Teorema de Ascoli-Arzela y el Principio de Acotación Uniforme veamos que responde a una pregunta natural. Observemos que una familia de funciones continuas de un espacio métrico en otro puede ser puntualmente acotada sin serlo uniformemente. Por ejemplo, consideremos las funciones $f_n \in C([0,1])$

$$f_n(t) = \begin{cases} n^2 t\,, & \text{si } 0 \leq t \leq \frac{1}{n} \\ \frac{1}{t}\,, & \text{si } \frac{1}{n} < t \leq 1 \end{cases}$$

Esta familia está puntualmente acotada ya que $f_n(0) = 0$ para cada $n \in \mathbb{N}$ y $|f_n(t)| \leq \frac{1}{t}$ para todo $t \in (0,1]$. Pero en cambio no están uniformemente acotadas (es decir, no admiten una misma cota valida para todo n y para todo t) ya que $f_n(\frac{1}{n}) = n$ para todo $n \in \mathbb{N}$.

Precisamente el Teorema de Ascoli-Arzela nos da condiciones suficientes para poder garantizar la acotación uniforme.

Vamos a trabajar en $C(T)$, el espacio de las funciones continuas definidas sobre un compacto T que suponemos *metrizable*.

Previamente necesitamos unas definiciones.

Ya sabemos que toda función $f \in C(T)$ es de hecho *uniformemente continua*, donde la "uniformidad" es en los $t \in T$. A un subconjunto $E \subset C(T)$ le podemos pedir que sea *equicontinuo* en t, y esto va a ser una continuidad uniforme en $f \in E$.

Definición 4.1. *Sea T un espacio compacto metrizable y sea $E \subset C(T)$. E se dice* equicontinuo *en $t \in T$ si para todo $\epsilon > 0$ existe $\delta > 0$ tal que para todo $f \in E$, si $d(s,t) < \delta$ entonces $|f(t) - f(s)| < \epsilon$.*

Definición 4.2. *$E \subset C(T)$ se dice* acotado *en $t \in T$ si*

$$\sup_{f \in E} |f(t)| < \infty.$$

Teorema 4.3 (Ascoli-Arzela). *Sea T un espacio compacto metrizable y sea $E \subset C(T)$. Supongamos que E es un conjunto acotado y equicontinuo en t para todo $t \in T$. Entonces $E \subset C(T)$ es precompacto (en particular E está acotado en la norma del supremo, es decir E está* uniformemente acotado*). Por tanto, toda sucesión de E contiene una subsucesión uniformemente convergente.*

Demostración. Tenemos que probar que E es precompacto. Sea $\epsilon > 0$. Por ser E equicontinuo en todo punto, para todo $t \in T$ existe $\delta_t > 0$ tal que para todo $f \in E$ y para todo $s \in B(t, \delta_t)$

$$|f(s) - f(t)| < \epsilon.$$

El conjunto $\{B(t,\delta_t); t \in T\}$ es un recubrimiento abierto de T y por ser T compacto existen $t_1, \ldots, t_n \in T$ tales que, llamando $\delta_i = \delta_{t_i}$,

$$T = \cup_{i=1}^n B(t_i, \delta_i).$$

Por ser E acotado en todo punto, existen $\alpha_1, \ldots, \alpha_n > 0$ tales que

$$|f(t_i)| \leq \alpha_i \text{ para todo } f \in E, \text{ y para todo } 1 \leq i \leq n.$$

Sea ahora $\alpha = \text{máx}_i\{\alpha_i\} + \epsilon$. Entonces para todo $t \in T$ existe $i \in \{1, \ldots, n\}$ tal que para todo $f \in E$

$$|f(t)| = |f(t) - f(t_i) + f(t_i)| \leq \epsilon + |f(t_i)| \leq \alpha$$

y por tanto E está uniformemente acotado en T, es decir E está acotado en $C(T)$. Veamos que de hecho E es *precompacto*. Sea

$$D_\alpha = \{k \in \mathbb{K} \text{ tales que } |k| \leq \alpha\}$$

y para toda $f \in E$ sea

$$e(f) = (f(t_1), \ldots, f(t_n)) \in D_\alpha^n.$$

El Teorema de Heine-Borel nos dice que D_α^n es un conjunto compacto de $\mathbb{K}^n$ (con cualquiera de sus normas). Consideramos por ejemplo la norma ∞ en $\mathbb{K}^n$. Por ser D_α^n compacto existe una cantidad finita de bolas $V_1, \ldots, V_m$ de radio ϵ que cubren D_α^n.

Ahora, para cada $1 \leq j \leq m$, si

$$V_j \cap \{e(f); f \in E\} \neq \emptyset$$

elijamos un $f_j \in E$ tal que $e(f_j) \in V_j$.

Vamos entonces a mostrar que las bolas de centros f_j y radio 5ϵ forman un recubrimiento de E:

Sea $f \in E$. Puesto que los V_j forman un recubrimiento de D_α^n sabemos que existe $j \in \{1, \ldots, m\}$ tal que $e(f) \in V_j$. Puesto que $e(f_j) \in V_j$ y dado que el radio de V_j es ϵ, sabemos que

$$\|e(f) - e(f_j)\|_\infty < \epsilon$$

es decir, para todo $i \in \{1, \ldots, n\}$

$$|f(t_i) - f_j(t_i)| < 2\epsilon.$$

Ahora, dado $t \in T$, existe $i \in \{1, \ldots, n\}$ tal que $t \in B(t_i, \delta_i)$ y por lo tanto

$$|f(t) - f_j(t)| \leq |f(t) - f(t_i)| + |f(t_i) - f_j(t_i)| + |f_j(t_i) - f_j(t)| < \epsilon + 2\epsilon + \epsilon.$$

Es decir

$$\|f - f_j\|_\infty \leq 4\epsilon < 5\epsilon,$$

lo que prueba que E es precompacto.

La última afirmación es inmediata: al ser $C(T)$ un espacio de Banach, la clausura de E, $\overline{E}$ es un subconjunto compacto. Por tanto toda sucesión en E tiene una subsucesión convergente (a un elemento de $\overline{E}$). $\square$

El Teorema de Banach-Steinhaus y el Principio de Acotación Uniforme

Antes de estudiar el Principio de Acotación Uniforme y el Teorema de Banach-Steinhaus, empezaremos recordando a los alumnos el Teorema de Baire, que suponemos ya conocido. Si no lo fuera, podríamos incluir una demostración.

TEOREMA 4.4 (Baire, 1899). *Sea X un espacio métrico. Entonces la intersección de una familia finita de subconjuntos abiertos densos de X es denso (y abierto) en X.*

Si además X es completo, entonces la intersección de una familia contable de subconjuntos abiertos densos de X en densa en X.

Notemos que, tomando complementarios, el Teorema de Baire dice que en un espacio métrico completo la unión contable de cerrados con interior vacío tiene interior vacío.

Podemos pensar el teorema de Ascoli-Arzela como una herramienta que nos da condiciones suficientes para que en una familia de funciones continuas escalares definidas sobre un compacto metrizable, la acotación puntual implique la acotación uniforme.

Nos preguntamos ahora cuál es la situación para una familia de aplicaciones lineales entre espacios de Banach. La respuesta nos la da el

TEOREMA 4.5 (Principio de Acotación Uniforme). *Sea X un espacio de Banach, Y un espacio normado y sea $\mathcal{T}$ un subconjunto de $\mathcal{L}(X;Y)$ puntualmente acotado, es decir, tal que para todo $x \in X$ el conjunto $\{T(x); T \in \mathcal{T}\} \subset Y$ está acotado. Entonces existe $C > 0$ tal que*

$$\sup_{T \in \mathcal{T}} \|T\| = \sup_{T \in \mathcal{T}, x \in B_X} \|T(x)\| \leq C.$$

DEMOSTRACIÓN. Para todo $n \in \mathbb{N}$, sea

$$D_n = \{x \in X; \|T(x)\| > n \text{ para algún } T \in \mathcal{T}\}.$$

Puesto que, toda $T \in \mathcal{T}$ es continua, y $\|\cdot\|$ también es una función continua, tenemos que $\|\cdot\| \circ T : X \longrightarrow [0, +\infty)$ es continua, por lo que

$$D_n^T := (\|\cdot\| \circ T)^{-1}(n, \infty) = \{x \in X \text{ tales que } \|T(x)\| > n\}$$

es un abierto. Puesto que $D_n = \cup_{T \in \mathcal{T}} D_n^T$, se sigue que D_n es un abierto.

Sea $x \in X$. Por hipótesis, existe $n_0 \in \mathbb{N}$ tal que $\|T(x)\| \leq n_0$ para todo $T \in \mathcal{T}$. Por tanto $x \notin D_{n_0}$ y se sigue que $\cap_{n\in\mathbb{N}} D_n = \emptyset$, en particular $\cap_{n\in\mathbb{N}} D_n$ no puede ser denso en X.

Ahora el Teorema de Baire nos dice que existe $m_0 \in \mathbb{N}$ tal que D_{m_0} no es denso. Por tanto existe $a \in X$ y $r > 0$ tal que $B(a,r) \cap D_{m_0} = \emptyset$. Esto quiere decir que $T(B(a,r)) \leq m_0$ para todo $T \in \mathcal{T}$, y veremos que a partir de ahí es fácil acabar sin mas que "trasladar" la bola al origen.

Sea $x \in B_X$, $T \in \mathcal{T}$. Entonces

$$\|T\,(rx + a)\| \leq m_0$$

(puesto que $rx + a \in B(a,r)$) de forma que

$$\|T(x)\| = \frac{1}{r}\,\|T\,(rx)\| = \frac{1}{r}\,\|T\,(rx + a) - T(a)\| \leq$$

$$\leq \frac{1}{r}\,(\|T\,(rx + a)\| + \|T(a)\|) \leq \frac{2m_0}{r}.$$

Es decir

$$\sup_{x\in B_X, T\in\mathcal{T}} \|T(x)\| = \sup_{T\in\mathcal{T}} \|T\| \leq \frac{2m_0}{r}.$$

□

Observación 4.6. *Probablemente sea conveniente hacer algunos comentarios al teorema.*

En primer lugar, obsérvese que el conjunto $E_0 = \{x \in X;$ *tales que* $\mathcal{T}$ *está acotado en* $x\}$ *es un subespacio vectorial. Por tanto, puesto que todo subespacio vectorial de* X *o bien es* X *o bien tiene interior vacío, se tiene que o bien el interior de* E_0 *es vacío (es decir, el complementario de* E_0 *es denso), o bien* $E_0 = X$. *Por lo tanto, o bien* $\mathcal{T}$ *está acotado en todo punto, o para todo* x *en un conjunto denso de* X, $\mathcal{T}$ *no está acotado en* x. *Esta es la formulación del Teorema que sigue por ejemplo* [**8**].

En realidad, usando la notación de categorías de Baire, se puede ver fácilmente que la demostración vale igual suponiendo que X *es un espacio normado y* $\mathcal{T}$ *está puntualmente acotada en un conjunto de* segunda categoría.

Corolario 4.7. *Sea* X *un espacio de Banach,* Y *un espacio normado y sea* $(T_n)_{n\in\mathbb{N}} \subset \mathcal{L}(X;Y)$ *una sucesión tal que para todo* $x \in X$ *la sucesión* $(T_n(x))_n$ *es convergente en* Y. *Definamos la aplicación* $T : X \longrightarrow Y$ *como*

$$T(x) = \lim_n T_n(x)$$

entonces

(i) **(Teorema de Banach Steinhaus, 1927)** $T \in \mathcal{L}(X;Y)$ *y*

$$\|T\| \leq \sup_n \|T_n\| < \infty.$$

(ii) *Sea $B \subset X$ un conjunto compacto. Entonces $(T_n(x))_n$ converge a $T(x)$ uniformemente en $x \in B$*

Demostración. (i) T es lineal, de nuevo por la continuidad de la suma y el producto y el escalares. Para ver T es continuo, empecemos notando que, para todo $x \in X$, el conjunto $\{\|T_n(x)\|; n \in \mathbb{N}\}$ está acotado, porque $(T_n(x))_n$ es una sucesión convergente. El Principio de Acotación Uniforme nos dice ahora que el conjunto $\{\|T_n\|; n \in \mathbb{N}\}$ está acotado y por tanto existe el $\sup_n \|T_n\|$. Entonces, para todo $x \in B_X$,

$$\|T(x)\| = \lim_n \|T_n(x)\| \leq \sup_n \|T_n\| < \infty$$

De forma que T es continua y

$$\|T\| \leq \sup \|T_n\| < \infty$$

(ii) Sea $\epsilon > 0$. Puesto que B es precompacto existen $x_1, \ldots, x_m \in B$ tales que

$$B \subset B(x_1, \epsilon) \cup \cdots \cup B(x_m, \epsilon)$$

Puesto que $T_n(x_i) \to T(x_i)$ para todo $1 \leq i \leq m$, existe $n_0 \in \mathbb{N}$ tal que

$$\|T_n(x_i) - T(x_i)\| < \epsilon$$

para todo $n \geq n_0$ y para todo $1 \leq i \leq m$. Sea ahora $x \in B$ y sea x_i tal que $\|x - x_i\| < \epsilon$. Entonces, para todo $n \geq n_0$, se tiene

$$\|T_n(x) - T(x)\| \leq \|T_n(x - x_i)\| + \|(T_n - T)(x_i)\| + \|T(x_i - x)\| \leq$$

$$\leq (\|T_n\| + \|T\|)\|x - x_i\| + \|T_n(x_i) - T(x_i)\| \leq (2 \sup_n \|T_n\| + 1)\epsilon.$$

□

Observación 4.8. *Observemos en primer lugar con un ejemplo que no podemos prescindir de la hipótesis de la completitud de X ni en el Principio de Acotación Uniforme ni en el Teorema de Banach-Steinhaus. Consideremos $X = c_{00}$ con la norma del supremo (en ese caso su completado es c_0) y para todo $n \in \mathbb{N}$ sea, para todo $x \in c_{00}$*

$$T_n(x) = \sum_{j=1}^{n} x(j)$$

Entonces $T_n \in c_{00}^* = \mathcal{L}(c_{00};\mathbb{K})$ *y* $\|T_n\| = n$ *para todo* $n \in \mathbb{N}$*. Sea* $x \in c_{00}$ *y sea* $i_x \in \mathbb{N}$ *tal que* $x(j) = 0$ *para todo* $j \geq i_x$*. Entonces, para todo* $n \in \mathbb{N}$*,*

$$|T_n(x)| \leq i_x \|x\|_\infty < \infty$$

y en cambio el conjunto $\{\|T_n\|; n \in \mathbb{N}\}$ *no está acotado. Esto muestra que el resultado del Principio de Acotación Uniforme no es cierto en este caso. Para ver que el Teorema de Banach-Steinhaus también falla, definiendo la aplicación lineal* $T_\infty : c_{00} \longrightarrow \mathbb{K}$ *como*

$$T_\infty(x) = \sum_{j=1}^{\infty} x(j)$$

se tiene que $T_n(x) \to T_\infty(x)$ *para todo* $x \in c_{00}$*, pero* T_∞ *no es continuo porque* $T_\infty(1, \overset{(n)}{\ldots}, 1, 0, 0, \ldots) = n$ *y* $(1, \overset{(n)}{\ldots}, 1, 0, 0, \ldots) \in B_{c_{00}}$ *para todo* $n \in \mathbb{N}$*.*

Ahora es fácil probar como consecuencia que un conjunto $B \subset E$ está acotado si y sólo si está débilmente acotado, y que un operador es continuo si y sólo si es débilmente continuo, donde el sentido de débilmente acotado y débilmente acotado es el dado por el enunciado siguiente.

TEOREMA 4.9. *Sean* X *e* Y *espacios normados y sean* $B \subset X$ *un subconjunto y* $T : X \longrightarrow Y$ *una aplicación lineal. Entonces*

(i) B *es acotado si y solo si para todo* $x^* \in X^*$ $x^*(B)$ *está acotado en* $\mathbb{K}$

(ii) T *es continua si y sólo si* $y^* \circ T$ *es continua para todo* $y^* \in Y^*$*.*

DEMOSTRACIÓN. Es fácil ver que si B está acotado también está débilmente acotado. Para la otra implicación, supongamos que para todo $x^* \in X^*$, $x^*(B) \subset \mathbb{K}$ está acotado. Sea $J : X \hookrightarrow X^{**}$ la inmersión canónica de X en su bidual. Entonces el conjunto $\{J(x); x \in B\}$ es un subconjunto de $X^{**} = \mathcal{L}(X^*;\mathbb{K})$ donde X^* es un espacio de Banach. Además, para todo $x^* \in X^*$ el conjunto $\{|J(x)(x^*)|; x \in B\} = \{|x^*(x)|; x \in B\}$ está acotado. Por lo tanto, el Principio de Acotación Uniforme nos dice que el conjunto $\{\|J(x)\|; x \in B\}$ está acotado, y como $\|J(x)\| = \|x\|$, se sigue que B está acotado.

Para la parte (ii), de nuevo una implicación está clara. Para la otra, supongamos que T es débilmente continua. Sea $B = \{T(x); x \in B_X\} \subset Y$. Entonces para todo $y^* \in Y^*$, $y^*(B) = (y^* \circ T)(B_X)$ está acotado, de forma que la parte (i) nos dice que B está acotado, es decir $T(B_X)$ está acotado, es decir T es continuo. $\square$

Los Teoremas de la Gráfica Cerrada y de la Aplicación Abierta

Al igual que hicimos al introducir el Principio de Acotación Uniforme y el Teorema de Banach-Steinhaus, antes de introducir los Teoremas de la Gráfica Cerrada y de la Aplicación Abierta reflexionaremos brevemente acerca de la pregunta natural a que los Teoremas dan respuesta.

Sean X e Y espacios métricos y sea $T : X \longrightarrow Y$ una aplicación. T es *continua* si y sólo si es continua por sucesiones, es decir, si para toda sucesión $(x_n) \subset X$ tal que $x_n \to x \in X$ se tiene que $T(x_n) \to T(x)$.

Por otro lado, T se dice *cerrada* si para toda sucesión $(x_n) \subset X$ tal que $x_n \to x \in X$ y $T(x_n) \to y \in Y$ se tiene que $y = T(x)$.

Se sigue inmediatamente de la definición que una aplicación continua es cerrada. En cambio el recíproco no es cierto, como lo prueba el siguiente ejemplo. Sea $X = Y = \mathbb{R}$ y sea

$$T(x) = \begin{cases} \frac{1}{x}\,, & \text{si } x \neq 0 \\ 0\,, & \text{si } x = 0 \end{cases}$$

Ejercicio 4.1. *Demostrar que T es cerrada si y sólo si la gráfica de T, esto es, el conjunto*

$$Gr(T) := \{(x, T(x)) \in X \times Y; x \in X\}$$

es cerrado en $X \times Y$ con la topología producto.

Notemos también que si una aplicación cerrada T es biyectiva, entonces su inversa T^{-1} también es cerrada, ya que que su gráfica es la misma (una está en $X \times Y$ y la otra en $Y \times X$, pero ambos espacios son claramente homeomorfos). En cambio la inversa de una aplicación biyectiva y continua no tiene por qué ser continua, como lo prueba el siguiente ejemplo: Sea $X = [0, 2\pi)$, $Y = \{z \in \mathbb{C}; |z| = 1\}$ y sea $T : X \longrightarrow Y$ la aplicación dada por

$$T(x) = e^{ix}$$

T es continua y biyectiva pero T^{-1} no es continua.

Vamos a ver ahora cuál es la situación para aplicaciones lineales entre espacios de Banach. Seguimos **[8]** para la demostración.

TEOREMA 4.10 (Gráfica Cerrada). *Sean X, Y espacios de Banach, y sea $T : X \longrightarrow Y$ una aplicación lineal. Entonces T es cerrada si y sólo si T es continua.*

DEMOSTRACIÓN. Ya hemos comentado que si T es continua entonces es cerrada. Veamos la otra implicación. Probamos primero el siguiente Aserto, en el que sólo usaremos la linealidad de T

Aserto: Para todo $\epsilon > 0$ existe $\delta > 0$ tal que

$$\delta B_X \subset \overline{T^{-1}(\epsilon B_Y)}$$

donde $\delta B_X = \{x \in X, \text{ tales que } \|x\| \leq \delta\}$ y análogamente ϵB_Y

Demostración del Aserto: Sea $\epsilon > 0$. Consideramos los conjuntos

$$G_n = nT^{-1}\left(\frac{\epsilon}{2}B_Y\right).$$

La unión $\cup_{n\in\mathbb{N}} G_n$ es un recubrimiento de X, y por tanto también lo es $\cup_{n\in\mathbb{N}} \overline{G_n}$. (Observese que todavía no sabemos que T sea continua, lo único que usamos es que T está bien definida). Por el Teorema de Baire (tomando complementarios) existe $n_0 \in \mathbb{N}$ tal que $\overline{G_{n_0}}$ tiene interior no vacío, y por tanto contiene una bola cerrada de radio $\eta > 0$ y centro x_0, lo que implica que si $\|x - x_0\| \leq \eta$ entonces $x \in \overline{n_0 T^{-1}\left(\frac{\epsilon}{2}B_Y\right)}$. Es fácil ver que esto es equivalente a que

$$\frac{x}{n_0} \in \overline{T^{-1}\left(\frac{\epsilon}{2}B_Y\right)}$$

y por lo tanto también se verifica, para todo $n \geq n_0$,

$$\frac{x}{n} \in \overline{T^{-1}\left(\frac{\epsilon}{2}B_Y\right)}. \tag{3}$$

Además, existe $n_1 \in \mathbb{N}$ tal que, para todo $n \geq n_1$,

$$\frac{x_0}{n} \in T^{-1}\left(\frac{\epsilon}{2}B_Y\right) \tag{4}$$

(esto simplemente porque $T(x_0) \in Y$)

Sea entonces $n_2 = \max\{n_0, n_1\}$ y sea $\delta = \frac{\eta}{n_2}$. Entonces para todo $y \in \delta B_Y$ se tiene

$$y = \frac{n_2 y + x_0 - x_0}{n_2} = \frac{n_2 y + x_0}{n_2} - \frac{x_0}{n_2}.$$

Ahora, puesto que $\|n_2 y + x_0 - x_0\| \leq \eta$, la ecuación (3) implica que

$$\frac{n_2 y + x_0}{n_2} \in \overline{T^{-1}\left(\frac{\epsilon}{2}B_Y\right)}.$$

Por otro lado, de la ecuación (4) se sigue que

$$\frac{x_0}{n_2} \in T^{-1}\left(\frac{\epsilon}{2}B_Y\right)$$

y ahora es fácil ver que

$$y \in \overline{T^{-1}(\epsilon B_Y)},$$

es decir

$$\delta B_Y \subset \overline{T^{-1}(\epsilon B_Y)}$$

lo que termina la demostración del Aserto.

Dado un $\epsilon > 0$, llamemos $\delta(\epsilon)$ a (uno de los) δ que nos proporciona el Aserto, y observemos que para cada $\epsilon > 0$ siempre podemos suponer que $\delta(\epsilon) < \epsilon$.

Continuamos con la demostración del Teorema. Suponemos que T es cerrada, y vamos a probar que, para todo $\epsilon > 0$

$$T\left(\delta\left(\frac{\epsilon}{2}\right)B_X\right) \subset \epsilon B_Y$$

lo que probará que T es continua en el origen, y por tanto en todo punto.

Así, fijamos $\epsilon > 0$ y $x \in \delta\left(\frac{\epsilon}{2}\right)B_X$. Elijamos $x_1 \in T^{-1}(\frac{\epsilon}{2}B_Y)$, con

$$x - x_1 \in \delta(\frac{\epsilon}{4})B_X,$$

(esto es, $\|x - x_1\| \le \delta(\frac{\epsilon}{4})$). A continuación elegimos $x_2 \in T^{-1}(\frac{\epsilon}{4}B_Y)$ con

$$x - x_1 - x_2 \in \delta(\frac{\epsilon}{8})B_X,$$

y procedemos inductivamente para construir una sucesión $(x_i)_{i\in\mathbb{N}}$ tal que $x_i \in T^{-1}(\frac{\epsilon}{2^i}B_Y)$ y

$$x - \sum_{j=1}^{i} x_j \in \delta(\frac{\epsilon}{2^{i+1}})B_X,$$

y por tanto,

$$x - \sum_{j=1}^{i} x_j \in T^{-1}\left(\frac{\epsilon}{2^{i+1}}B_X\right).$$

Entonces $T(x_i) \in \frac{\epsilon}{2^i}B_Y$ y por tanto la sucesión $z_n = \sum_{i=1}^{n} T(x_i)$ es una sucesión de Cauchy que converge a un punto $z \in Y$. Puesto que $z_n \in \epsilon B_Y$ para todo n, se tiene que $z \in \epsilon B_Y$.

Además $x - \sum_{j=1}^{n} x_j \in \delta(\frac{\epsilon}{2^{n+1}})B_X \subset \frac{\epsilon}{2^{n+1}}B_X$ (porque $\delta(\epsilon) < \epsilon$) y por tanto la sucesión $t_n = \sum_{j=1}^{n} x_j$ converge a x. Puesto que $T(t_n) = z_n$

y puesto que T es cerrada, tenemos que $T(x) = z$, por lo que $x \in T^{-1}(\epsilon B_Y)$, lo que prueba que $\delta(\frac{\epsilon}{2})B_X \subset T^{-1}\epsilon B_Y$ y termina la demostración del teorema. $\square$

OBSERVACIÓN 4.11. *Nótese que el teorema dice que si X e Y cumplen las hipótesis, entonces para probar que una aplicación T sea continua basta comprobar que si $(x_n)_{n\in\mathbb{N}} \subset X$ es una sucesión convergente a 0 tal que la sucesión $(T(x_n))_{n\in\mathbb{N}} \subset Y$ converge a $y \in Y$ entonces $y = 0$. Esta es la forma en que se usa el Teorema en muchas aplicaciones.*

El siguiente corolario es de gran importancia en múltiples aplicaciones.

COROLARIO 4.12. *Sean X, Y dos espacios de Banach y sea $T : X \longrightarrow Y$ un operador (aplicación lineal y continua) biyectiva. Entonces $T^{-1} : Y \longrightarrow X$ es continua.*

DEMOSTRACIÓN. Si T es continua, su gráfica es cerrada. Pero la gráfica de T^{-1} es la misma que la de T (una en $X \times Y$ y la otra en $Y \times X$) por lo que la gráfica de T^{-1} también es cerrada. Ahora el Teorema de la Gráfica Cerrada nos garantiza que T^{-1} es continua. $\square$

Continuamos el capítulo estudiando el Teorema de la Aplicación Abierta.

DEFINICIÓN 4.13. *Dados dos espacios topológicos X e Y, una aplicación $T : X \longrightarrow Y$ se dice abierta si la imagen de cualquier abierto es un abierto.*

Claramente, si la aplicación T es inyectiva, T es abierta si y sólo si T^{-1} es continua.

Entonces tenemos el siguiente resultado.

TEOREMA 4.14. *Sean X, Y espacios normados y $T : X \longrightarrow Y$ una aplicación lineal. Entonces T es una aplicación abierta si y sólo si existe $\gamma > 0$ tal que*

$$B_Y \subset T(\gamma B_X),$$

es decir, para todo $y \in Y$ existe $x \in X$ con $\|x\| \leq \gamma\|y\|$ de manera que $T(x) = y$. En particular, las aplicaciones abiertas son sobreyectivas.

DEMOSTRACIÓN. Supongamos que T es una aplicación abierta. Entonces transforma entornos abiertos de 0 en entornos abiertos de 0. Sea

$$U_X = \{x \in X \text{ tales que } \|x\| < 1\}.$$

Entonces existe $\delta > 0$ tal que

$$\delta U_Y \subset T(U_X) \subset T(B_X).$$

Sea $0 < \delta' < \delta$. Entonces

$$\delta' B_Y \subset \delta U_Y \subset T(U_X) \subset T(B_X).$$

Ahora es fácil ver que

$$\gamma = \frac{1}{\delta'}$$

cumple lo pedido.

Para la otra implicación, supongamos que

$$\gamma B_Y \subset T(B_X)$$

y sea $G \subset X$ un abierto. Veamos que $T(G)$ es un abierto.

Sea $a \in G$. Por ser G abierto existe $\delta > 0$ tal que

$$a + \delta B_X \subset G.$$

Entonces

$$T(a + \delta B_X) = T(a) + \delta T(B_X) \subset T(G)$$

y por lo tanto

$$T(a) + \delta\gamma B_Y \subset T(G)$$

por lo que $T(G)$ es un abierto. □

Teorema 4.15. *Sean X, Y espacios normados.*

1. *Si $Z \subset X$ es un subespacio cerrado, entonces la aplicación cociente*

$$Q : X \longrightarrow X/Z$$

es continua y abierta.

2. *Sea $T : X \longrightarrow Y$ una aplicación lineal tal que* $\ker(T) \subset E$ *es cerrado. Sea $\overline{T} : X/\ker(T) \longrightarrow Y$ la aplicación lineal dada por*

$$\overline{T}([x]) = T(x).$$

Entonces T es abierta si y sólo si $\overline{T}$ es abierta.

Demostración. Para probar (1), en primer lugar Q es continuo porque $|||Q(x)||| = |||[x]||| \leq \|x\|$. Para ver que Q es abierta, utilizamos el resultado anterior. Sea un $\epsilon > 0$ cualquiera y sea $[x] \in X/Z$. Entonces

$$\inf\{\|x + z\|; z \in Z\} = |||[x]||| < (1 + \epsilon)|||[x]|||$$

por tanto existe $z_0 \in Z$ tal que $\|x + z_0\| < (1 + \epsilon)|||[x]|||$. Puesto que $Q(x + z_0) = [x]$, podemos hacer $\gamma = 1 + \epsilon$ en el teorema anterior y obtenemos que Q es abierta.

Para probar (2), puesto que $\ker(T)$ es cerrado, sea $Q : X \longrightarrow X/\ker(T)$ la aplicación cociente. Entonces para todo $A \subset X$,

$$T(A) = \overline{T}(Q(A)).$$

Puesto que Q es abierta, se sigue trivialmente que si $\overline{T}$ es abierta también lo es T.

Por otro lado, para todo $\overline{A} \subset X/\ker(T)$, se tiene que

$$\overline{T}(\overline{A}) = T(Q^{-1}(\overline{A}))$$

ya que Q es sobre.

Puesto que Q es continuo por (1), se sigue que si T es abierta $\overline{T}$ también lo es. $\square$

Finalmente ya tenemos todas las herramientas para probar

TEOREMA 4.16 (Teorema de la Aplicación Abierta, Banach 1932). *Sean X, Y espacios de Banach y sea $T : X \longrightarrow Y$ una aplicación cerrada y sobre. Entonces T es continua y abierta.*

DEMOSTRACIÓN. Por el Teorema de la Gráfica Cerrada, T es continua. Por tanto $\ker(T) = T^{-1}(0)$ es cerrado y la aplicación

$$\overline{T} : X/\ker(T) \longrightarrow Y$$

definida por

$$\overline{T}([x]) = T(x)$$

es continua (esto lo vimos en el Teorema 2.1, punto (v)). Podríamos también no incluirlo en aquel Teorema y demostrarlo ahora.) Por tanto $\overline{T}$ es cerrada. $\overline{T}$ es inyectiva (esto es álgebra) y, puesto que T es sobre también lo es $\overline{T}$ (más álgebra). Por tanto $\overline{T}$ es una aplicación lineal cerrada y biyectiva. Así pues la aplicación inversa $\overline{T}^{-1}$ también es lineal cerrada y biyectiva (es cerrada porque la gráfica es la misma cambiando el orden de los pares). Puesto que Y y $X/\ker(T)$ son espacios de Banach, tenemos que $\overline{T}^{-1}$ es continua, es decir $\overline{T}$ es abierta. Por el Teorema anterior T es abierta. $\square$

Proyecciones en un espacio de Banach

DEFINICIÓN 4.17. *Sea X un espacio normado e $Y \subset X$ un subespacio vectorial cerrado. Un operador (lineal y continuo) $P : X \longrightarrow X$ es una* proyección sobre Y *si $P_{|_Y} : Y \longrightarrow Y$ es la identidad y si $P(x) \in Y$ para todo $x \in X$.*

DEFINICIÓN 4.18. *Sea X un espacio normado e $Y \subset X$ un subespacio vectorial cerrado. Y se dice* complementado *en X si existe una proyección de X sobre Y.*

EJERCICIO 4.2. *Sea X un espacio normado, $Y \subset X$ un subespacio vectorial cerrado y $P : X \longrightarrow X$ un operador (lineal y continuo). Probar que P es una proyección si y sólo si $P^2 = P$, $P(x) \in Y$ para todo $x \in X$ y $P : X \longrightarrow Y$ es sobreyectiva.*

PROPOSICIÓN 4.19. *Sea X un espacio de Banach e $Y \subset X$ un subespacio vectorial cerrado. Y está complementado en X si y sólo si existe un subespacio vectorial cerrado $Z \subset X$ tal que $X = Y \oplus Z$ en el sentido algebraico, es decir, $X = Y + Z$ e $Y \cap Z = \{0\}$. En ese caso se dice que Z es un complementario de Y en X.*

DEMOSTRACIÓN. Supongamos que existe Z como en el enunciado. Tanto Y como Z son espacios de Banach y por tanto $(Y \oplus Z, \|\cdot\|_1)$ es un espacio de Banach, donde

$$\|(y,z)\|_1 = \|y\| + \|z\|.$$

Definimos

$$\varphi : Y \oplus Z \longrightarrow X$$

como

$$\varphi(y,z) = y + z.$$

φ es claramente lineal. También es continua (considerando la norma 1 en $Y \oplus Z$) ya que

$$\|\varphi(y,z)\| = \|y+z\| \leq \|y\| + \|z\| = \|(y,z)\|_1.$$

Además φ es inyectiva: $\varphi(y,z) = 0$ implica $y + z = 0$ lo que a su vez implica que $y = z = 0$ por ser $X = Y \oplus Z$.

También φ es sobreyectiva: para todo $x \in X$, x se puede escribir como $x = y + z$ (de forma única) y $\varphi(y,z) = x$.

Por los Teoremas de la Gráfica Cerrada (o de la Aplicación abierta) sabemos que φ^{-1} es continua. De hecho se tiene, aunque no lo usamos explícitamente, que X es isomorfo a $(Y \oplus Z, \|\cdot\|_1)$; es un ejercicio comprobar que podemos sustituir la norma 1 por, por ejemplo, cualquier norma p.

Notemos además que la aplicación $\pi_1 : Y \oplus Z \longrightarrow Y \oplus Z$ dada por $\pi_1(y,z) = y$ es claramente una proyección.

Definimos entonces la aplicación

$$P = \varphi \circ \pi_1 \circ \varphi^{-1} : X \longrightarrow Y \oplus Z \longrightarrow y \oplus Z \longrightarrow X$$

Es ahora muy fácil comprobar que P es una proyección sobre Y.

Supongamos ahora que $Y \subset X$ está complementado, y sea $P : X \longrightarrow X$ una proyección. Sea $Z = \ker P$. Z es claramente cerrado.

Además, si $y \in Y \cap Z$ se tiene que $P(y) = y$ (porque $y \in Y$) y que $P(y) = 0$ (porque $y \in Z$) y por tanto $y = 0$, luego $Y \cap Z = \{0\}$.

Por otro lado, sea $x \in X$, sea $y = P(x) \in Y$ y sea $z = x - y$. Entonces

$$P(z) = P(x) - P(y) = y - y = 0,$$

por tanto $z \in Z$ y $x = y + z$, por lo que $X = Y + Z$. □

EJERCICIO 4.3. *Sea $Y \subset X$ un subespacio complementado, sea $P : X \longrightarrow X$ una proyección y sea $Z = \ker P$ un complementario de Y en X. Entonces Z está complementado en X y*

$$Id - P : X \longrightarrow X$$

es una proyección. **Sugerencia:** *Usar el Ejercicio 4.2.*

Desde el punto de vista puramente algebraico, todo subespacio vectorial está complementado. Se puede probar, dado un espacio de Banach X, que todo subespacio cerrado $Y \subset X$ está complementado si y sólo si X es Hilbert. En cambio, dado un subespacio cerrado Y de un espacio de Banach X, no es en general nada sencillo averiguar si Y está complementado o no. El artículo **[38]** (al menos partes de él) puede ser leído por no especialistas que deseen aprender algo acerca de este problema.

En esta línea, con técnicas elementales se puede probar que c_0 no está complementado en ℓ_∞. Por claridad, extraemos parte de la demostración como lema previo.

LEMA 4.20. *Existe una familia $\mathcal{F}$ de subconjuntos de $\mathbb{N}$ tal que*

- *$card(\mathcal{F}) = card(\mathbb{R})$*
- *Para todo $F \in \mathcal{F}$ se tiene que $card(F) = \aleph_0$*
- *Para todos $F_1, F_2 \in \mathcal{F}$*

$$card(F_1 \cap F_2) < \aleph_0$$

DEMOSTRACIÓN. Sea $i : \mathbb{N} \longrightarrow \mathbb{Q}$ una biyección. Para todo $r \in \mathbb{R}$ elijamos una sucesión $(q_n^r)_n \subset \mathbb{Q}$ tal que $q_n^r \to r$. Sea

$$\varphi : \mathbb{R} \longrightarrow \mathbb{N}^{\mathbb{N}}$$

la aplicación dada por

$$\varphi(r) = \left(i^{-1}(q_n^r)\right)_n .$$

Es fácil ver que φ es inyectiva. Sea ahora

$$\mathcal{F} = \{\varphi(r); r \in \mathbb{R}\}.$$

No es difícil ver que $\mathcal{F}$ cumple lo pedido. $\square$

TEOREMA 4.21. *c_0 no está complementado en ℓ_∞*

DEMOSTRACIÓN. [**24**, Thm. 95]. $\square$

Veamos que los cocientes por subespacios complementados son exactamente lo que esperaríamos que fueran.

PROPOSICIÓN 4.22. *Sea X un espacio de Banach, $Y \subset X$ un subespacio complementado de manera que $X = Y \oplus Z$. Entonces $(X/Y) = Z$, donde la igualdad quiere decir que ambos espacios son isomorfos (pero no necesariamente isométricos).*

DEMOSTRACIÓN. Sea $\theta : Z \longrightarrow X/Y$ dada por $\theta(z) = [z]$. Fácilmente se ve que θ es lineal y continua. Es inyectiva porque si $[z] = 0$ entonces $z \in Y$, y puesto que $Y \cap Z = \{0\}$, se sigue que $z = 0$. Es sobreyectiva porque si $[x] \in X/Y$ y $x = y + z$, con $y \in Y$ y $z \in Z$, entonces $\theta(z) = [x]$. Por tanto es biyectiva. Los Teoremas de este capítulo nos dicen ahora que θ es un homeomorfismo lineal $\square$

Veamos ahora que los subespacios finito dimensionales están siempre complementados. Aprovechamos la ocasión para definir las Bases de Auerbach.

DEFINICIÓN 4.23. *Dado un espacio de Banach Y con $\dim Y = n < \infty$ una base de Auerbach de Y es un sistema $\{e_1, \ldots, e_n\}, \{e_1^*, \ldots, e_n^*\}$ tal que $\{e_1, \ldots, e_n\}$ es una base de Y, $\{e_1^*, \ldots, e_n^*\} \subset Y^*$, para todos $1 \leq i, j \leq n$*

$$e_i^*(e_j) = \delta_{ij}$$

y, para todos $1 \leq i, j \leq n$,

$$\|e_i\| = \|e_j^*\| = 1.$$

(Obsérvese que de aquí se sigue en particular que $\{e_1^*, \ldots, e_n^*\}$ son linealmente independientes, y por tanto una base de Y^*, ya que si

$$\sum_i \alpha_i e_i^* = 0$$

entonces para todo j

$$\sum_i \alpha_i e_i^*(e_j) = \alpha_j = 0.$$

Ejercicio 4.4. [**24**, Thm 89] *Sea X un espacio de Banach, $Y \subset X$ un subespacio n-dimensional. Entonces existen vectores $\{e_1, \dots, e_n\} \subset Y$, $\{e_1^*, \dots, e_n^*\} \subset X^*$ tales que $\{e_1, \dots, e_n\}, \{e^*_{1|_Y}, \dots, e^*_{n|_Y}\}$ es una base de Auerbach. Por lo tanto Y está complementado en X y existe una proyección P de X sobre Y con $\|P\| \leq n$.*

Existe una versión de este resultado garantizando la existencia de una forma débil de Bases de Auerbach (no de proyecciones acotadas) en todo espacio de Banach separable [**36**].

Por otro lado, del Teorema de Kadets-Snobar se sigue que todo espacio n-dimensional admite una proyección de norma menor o igual que $\sqrt{n}$ (ver, por ejemplo, [**16**, Theorem 4.18]).

El Teorema de Stone-Weierstrass

Sea K un espacio compacto, y $C(K, \mathbb{R})$ el espacio de Banach de las aplicaciones continuas $f : C(K) \longrightarrow \mathbb{R}$, dotado de la norma del supremo. Observemos en primer lugar que $C(K, \mathbb{R})$ tiene además de la estructura de espacio de Banach, una estructura natural de álgebra.

Recordamos al lector la definición de álgebra.

Definición 4.24. *Sea A un espacio vectorial sobre $\mathbb{K}$. Decimos que A es un álgebra si existe un producto $\cdot : A \times A \longrightarrow A$ definido en él tal que $(A, +, \cdot)$ es un anillo y además*

$$\alpha(x \cdot y) = (\alpha x) \cdot y \text{ para todos } x, y \in A,\ \alpha \in \mathbb{K}.$$

Siguiendo la costumbre habitual escribiremos xy en lugar de $x \cdot y$ siempre que no haya riesgo de confusión. Claramente $C(K, \mathbb{R})$ con el producto punto a punto es un álgebra.

Decimos que un conjunto $A \subset C(K, \mathbb{R})$ es *reticulado* si para todos $f, g \in A$ las funciones $\sup(f, g)$ e $\inf(f, g)$ pertenecen a A, donde estas funciones se definen como

$$\sup(f, g)(t) := \max\{f(t), g(t)\} \text{ para todo } t \in K,$$

y análogamente la función $\inf(f, g)$. Es un ejercicio verificar que si $f, g \in C(K, \mathbb{R})$ entonces $\sup(f, g) \in C(K, \mathbb{R})$ e $\inf(f, g) \in C(K, \mathbb{R})$.

Antes de probar el Teorema de Stone-Weierstrass necesitamos dos lemas. El primero es

Lema 4.25. *Sea $A \subset C(K, \mathbb{R})$ un conjunto reticulado. Entonces $f \in \overline{A}$ (la clausura topológica de A) si y sólo si para todos $s, t \in K$, f*

es el límite sobre $\{s,t\}$ *de funciones de* A*, es decir, dados* $s,t \in K$ *y dado* $\epsilon > 0$ *existe* $g_{s,t} \in A$ *tal que*

$$|f(s) - g_{s,t}(s)| < \epsilon \quad y \quad |f(t) - g_{s,t}(t)| < \epsilon.$$

Demostración. Una de las implicaciones está clara. Para la otra, supongamos que para todos $s,t \in K$, f es el límite sobre $\{s,t\}$ de funciones de A. Sea $\epsilon > 0$ y sea $g_{s,t} \in A$ tal que

$$|f(s) - g_{s,t}(s)| < \epsilon \quad \text{y} \quad |f(t) - g_{s,t}(t)| < \epsilon. \tag{5}$$

Definamos

$$W_{s,t} = \{u \in K \text{ tales que } g_{s,t}(u) < f(u) + \epsilon\}.$$

Como la función $(g_{s,t} - f)$ es continua, se tiene que $W_{s,t} \subset K$ es abierto y usando (5) es muy fácil ver que contiene a t. Por tanto, fijado $s \in K$ la familia de conjuntos $\{W_{s,t}; t \in K\}$ forma un recubrimiento abierto de K. Al ser K compacto podemos extraer de este recubrimiento un subrecubrimiento finito $\{W_{s,t_i}; 1 \leq i \leq n\}$.

Definamos ahora

$$g_s = \operatorname*{ínf}_i (g_{s,t_i}).$$

Puesto que A es un conjunto reticulado se tiene que $g_s \in A$. Para todo $u \in K$, existe $i \in \{1, \ldots, n\}$ tal que $u \in W_{s,t_i}$, es decir $g_{s,t_i}(u) < f(u) + \epsilon$ y por tanto

$$g_s < f + \epsilon, \text{ es decir } g_s(u) < f(u) + \epsilon \text{ para todo } u \in K.$$

Además

$$g_s(s) > f(s) - \epsilon,$$

ya que, para todo $t \in K$, se tiene que $|f(s) - g_{s,t}(s)| < \epsilon$.

Definamos ahora

$$W_s = \{u \in K \text{ tales que } g_s(u) < f(u) - \epsilon\}.$$

Como la función $(g_s - f)$ es continua, se tiene que $W_s \subset K$ es abierto y es fácil ver que contiene a s. Por tanto la familia de conjuntos $\{W_s; s \in K\}$ forman un recubrimiento abierto de K. De nuevo por ser K compacto podemos extraer de este recubrimiento un subrecubrimiento finito $\{W_{s_j}; 1 \leq j \leq m\}$.

Definamos ahora

$$g = \sup_j (g_{s_j}).$$

De nuevo por ser A es un conjunto reticulado se tiene que $g \in A$. Además se tiene

$$f - \epsilon < g < f + \epsilon$$

y puesto que esto se puede hacer para todo $\epsilon > 0$ se sigue que $f \in \overline{A}$. $\square$

Necesitamos un lema más. Previamente recordemos que si A es un álgebra entonces decimos que $A' \subset A$ si A' con las operaciones inducidas también es un álgebra. Obviamente, al igual que ocurría con subespacios vectoriales, basta con que sea cerrada para las operaciones del álgebra.

LEMA 4.26. *Sea $A \subset C(K, \mathbb{R})$ una subálgebra cerrada (en el sentido topológico). Entonces A es un conjunto reticulado.*

DEMOSTRACIÓN. Observando que

$$\sup(f,g) = \frac{f+g+|f-g|}{2}$$

y

$$\text{ínf}(f,g) = \frac{f+g-|f-g|}{2}$$

queda claro que basta demostrar que si $f \in A$ entonces $|f| \in A$

Para ello vamos a ver que $|f|$ es el límite uniforme de polinomios en f de la forma $\sum_{n=1}^{m} a_n f^n$.

Es fácil ver que podemos suponer sin perdida de generalidad que $\|f\| \leq 1$.

Notemos que para todo $\epsilon > 0$ y para todo $x \in \mathbb{R}$, se tiene

$$0 \leq (x^2 + \epsilon^2)^{\frac{1}{2}} - |x| \leq \epsilon$$

Por otro lado, para todo $x \in [-1, 1]$ se tiene

$$x^2 + \epsilon^2 = 1 + \epsilon^2 + (x^2 - 1) = (1 + \epsilon^2)(1 + u),$$

donde

$$u = \frac{x^2 - 1}{1 + \epsilon^2}.$$

Además, para todo $x \in [-1, 1]$

$$|u| < \frac{1}{1 + \epsilon^2} < 1.$$

Por lo tanto, la serie de Taylor de $(1+u)^{\frac{1}{2}}$ converge uniformemente a $(1+u)^{\frac{1}{2}}$ en el intervalo $[-1, 1]$. Es decir, existe un polinomio $P(x)$ tal que, para todo $x \in [-1, 1]$,

$$|(x^2 + \epsilon^2)^{\frac{1}{2}} - P(x)| \leq \epsilon.$$

En particular se tiene que

$$|P(0)| \leq 2\epsilon.$$

Finalmente, si hacemos $Q = P - P(0)$, se tiene
$$\|Q(f) - |f|\| \leq 4\epsilon.$$
$\square$

Necesitamos una definición más.

DEFINICIÓN 4.27. *Sean K, Y dos conjuntos y sea $A \subset Y^K$ un conjunto de aplicaciones de K en Y. Decimos que A separa puntos de K si para todos $s, t \in K$ existe $f \in A$ tal que $f(s) \neq f(t)$.*

Finalmente podemos enunciar y probar el resultado principal de este capítulo. Probamos primeramente el caso real.

TEOREMA 4.28 (Stone-Weierstrass). *Sea $A \subset C(K, \mathbb{R})$ una subálgebra tal que*
1. *A separa puntos de K.*
2. *Para todo $t \in K$ existe $f \in A$ tal que $f(t) \neq 0$.*

Entonces $\overline{A} = C(K, \mathbb{R})$.

DEMOSTRACIÓN. Por los lemas anteriores, basta ver que para todos $s, t \in K$ con $s \neq k$ y para todos $\alpha, \beta \in \mathbb{R}$ existe $f \in A$ tal que $f(s) = \alpha$ y $f(t) = \beta$. Puesto que A separa puntos sabemos que existe $g \in A$ tal que $g(s) \neq g(t)$. Supongamos inicialmente que $g(s) \neq 0 \neq g(t)$. Veamos entonces que existen $a, b \in \mathbb{R}$ tales que
$$f = ag + bg^2$$
cumple lo pedido. Para esto tenemos que solucionar el sistema
$$\begin{cases} \alpha = ag(s) + bg^2(s) \\ \beta = ag(t) + bg^2(t) \end{cases}$$
y sabemos que este sistema tiene solución si
$$g(s)g^2(t) - g(t)g^2(s) = g(s)g(t)(g(s) - g(t)) \neq 0.$$
Por hipótesis sabemos que esta condición se cumple, y por tanto existen a y b que cumplen lo pedido.

Si la función g que separa s y t cumpliera que, por ejemplo, $g(s) = 0$, entonces por las hipótesis sobre A existe otra función $g' \in A$ tal que $g'(s) \neq 0$. Entonces podemos elegir un $\epsilon > 0$ suficientemente pequeño de manera que
$$g'' = g + g'$$
separa s y t y no se anula en ninguno de ellos, y podemos aplicarle ahora a g'' los razonamientos anteriores. $\square$

Los siguientes corolarios son a menudo cómodos de aplicar.

COROLARIO 4.29. *Si* $\{f_i; i \in I\} \subset C(K, \mathbb{R})$ *es un conjunto que separa puntos de* K *y si para todo* $t \in K$ *existe* f_i *tal que* $f_i(t) \neq 0$*, entonces toda función* $f \in C(K, \mathbb{R})$ *se puede escribir como el límite uniforme de polinomios (sin término constante) en los elementos* f_i*.*

DEMOSTRACIÓN. Los polinomios en esas funciones forman una subálgebra que verifica las condiciones del Teorema 4.28 □

COROLARIO 4.30. *Sea* $A \subset C(K, \mathbb{R})$ *una subálgebra tal que*

1. A *separa puntos de* K*.*
2. $\mathbb{1} \in A$*, donde* $\mathbb{1}$ *es la función constantemente igual a* 1*.*

Entonces $\overline{A} = C(K, \mathbb{R})$*.*

COROLARIO 4.31 (Weierstrass, 1885). *En* $C([0,1], \mathbb{R})$ *los polinomios son densos.*

Veamos finalmente la versión compleja del Teorema.

COROLARIO 4.32. *Sea* K *un conjunto compacto y sea* $A \subset C(K, \mathbb{C})$ *una subálgebra (sobre* $\mathbb{C}$*) tal que*

1. A *separa puntos.*
2. *Para todo* $x \in K$ *existe* $f \in A$ *tal que* $f(x) \neq 0$
3. *Para todo* $f \in A$*, también* $\overline{f} \in A$*, donde* $\overline{f}(t) = \overline{f(t)}$*, el conjugado complejo de* $f(t)$*.*

Entonces $\overline{A} = C(K, \mathbb{C})$*, donde* $\overline{A}$ *denota la clausura topológica de* A*.*

DEMOSTRACIÓN. Sea

$$A_r = \{\Re(f); f \in A\}.$$

Es fácil ver que A_r es una subálgebra (sobre $\mathbb{R}$) de $C(K, \mathbb{R})$ y que verifica las condiciones 1 y 2 del Teorema 4.28 (Si $f(x) \neq f(y)$ entonces $\Re(f)(x) \neq \Re(f)(y)$ ó $\Re(\imath f)(x) \neq \Re(\imath f)(y)$; análogamente, si $f(x) \neq 0$ entonces $\Re(f)(x) \neq 0$ ó $\Re(\imath f)(x) \neq 0$). Por lo tanto $\overline{A_r} = C(K, \mathbb{R})$, donde $\overline{A_r}$ es la clausura topológica de A_r y se ve fácilmente que entonces

$$\overline{A} = \overline{A_r + \imath A_r} = C(K, \mathbb{C}).$$

□

Aplicaciones

Al igual que con el Teorema de Hahn-Banach, hay multitud de aplicaciones de los teoremas de este capítulo al Análisis Funcional o a otras áreas del Análisis. Citamos aquí alguna de ellas.

Existencia de solución de ecuaciones diferenciales: Una aplicación del Teorema de Ascoli-Arzela que se puede ver en [**22**, p. 30] es el siguiente teorema.

TEOREMA 4.33. *Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ una función continua en un abierto D. Entonces, para todo $(x_0, y_0) \in D$ la ecuación diferencial*

$$y' = f(x, y)$$

admite una solución local (no necesariamente única) que pasa por el punto (x_0, y_0).

Divergencia de la serie de Fourier de funciones continuas: Se tiene el siguiente resultado.

TEOREMA 4.34. *Existe un conjunto denso $E \subset C[-\pi, \pi]$ tal que para toda $f \in E$ la serie de Fourier de f diverge en el 0.*

Es bastante frecuente encontrar esta aplicación en los libros. Una buena referencia para este curso puede ser [**31**, p. 144] o [**24**, ej 39 p. 59].

En [**22**, p. 78] se puede ver una aplicación similar a la divergencia de polinomios interpoladores.

Métodos de Integración: En [**31**, p. 146] se puede ver una aplicación de los resultados de este capítulo a la convergencia de métodos de integración.

Condensación de singularidades: La idea general de un principio de condensación de singularidades es ver condiciones bajo las cuales se puede garantizar que, si para todo punto t_0 en un conjunto denso existe una función f_0 en un conjunto dado de funciones $\mathcal{F}$ que tiene una singularidad en $t = t_0$ entonces existe $f \in \mathbb{F}$ tal que f tiene una singularidad en todo $t \in S$. En el caso particular de operadores entre espacios de Banach se tiene el siguiente resultado.

TEOREMA 4.35. *Sea X un espacio de Banach y sea Y un espacio normado. Sea $(T_n^m)_{n,m\in\mathbb{N}} \subset \mathcal{L}(X; Y)$ una sucesión (indexada por n y m) tal que para todo $m \in \mathbb{N}$ existe $x_m \in X$ tal que*

$$\operatorname*{lím\,sup}_n \|T_n^m(x_m)\| = +\infty.$$

Entonces existe $x \in X$ tal que para todo $m \in \mathbb{N}$

$$\operatorname*{lím\,sup}_n \|T_n^m(x)\| = +\infty.$$

Se puede ver por ejemplo en [**22**].

Propiedad de lifting de ℓ_1. Se tiene el siguiente resultado

TEOREMA 4.36. *Sean X e Y espacios de Banach y $Q : X \longrightarrow Y$ un operador sobreyectivo. Entonces, para todo operador $T : \ell_1 \longrightarrow Y$ existe un operador $\tilde{T} : \ell_1 \longrightarrow X$ tal que $Q\overline{T} = T$.*

Se puede ver por ejemplo en [**8**, Ex. 3, p. 15].

Todo espacio de Banach separable es isomorfo a un cociente de ℓ_1, o todo espacio de Banach es isomorfo a un cociente de $\ell_1(\Gamma)$ para algún Γ. Se puede ver por ejemplo en [**24**, Theorem 91].

Soluciones de ecuaciones diferenciales. Sean X, Y espacios de Banach y $T \in \mathcal{L}(X; Y)$ un operador sobre, es decir para todo $y \in F$ la ecuación

$$T(x) = y \tag{6}$$

tiene solución en X. Por los Teoremas de la Aplicación Abierta, existe $\gamma \in \mathbb{R}$ tal que, para todo $y \in Y$ existe una solución x de (6) con $\|x\| \leq \gamma\|y\|$.

En muchas ocasiones la posibilidad de acotar la norma de una solución de la ecuación en términos de la norma del término independiente es importante. Por ilustrar esa importancia consideremos la siguiente situación. Supongamos, y este es el caso a menudo, que T no es sólo sobre sino biyectiva, de forma que los teoremas de este capítulo nos garantizan que T^{-1} también es lineal y continua. En ocasiones, aunque (6) tenga solución para todo término independiente y, sólo sabemos calcular una solución x cuando y pertenece a un cierto conjunto denso $D \subset F$. Si ahora $y \in F \setminus D$, podemos hallar una sucesión $(y_n) \subset D$ tal que $y_n \to y$ y una sucesión $(x_n) \subset E$ tal que $T(x_n) = y_n$. Obviamente la pregunta es si podemos garantizar que $x_n \to x$ donde x es el único elemento de X tal que $T(x) = y$.

La respuesta es sí, ya que

$$\|x - x_n\| = \|T^{-1}(y) - T^{-1}(y_n)\| \leq \|T^{-1}\|\|y - y_n\| \to 0$$

De forma que tiene sentido decir que x_n es una solución aproximada de (6) y además vemos que x varia continuamente con respecto a y. Estos razonamientos son los que están detrás de ciertas técnicas de perturbación utilizadas en la resolución de ecuaciones diferenciales: perturbamos el término independiente levemente (de y a y_n) y obtenemos una solución aproximada x_n que está suficientemente cerca de x (en cierta norma) si y_n está suficientemente cerca de y (en otra cierta norma).

Veámoslo con un ejemplo. Sea la ecuación diferencial lineal de orden m con coeficientes variables dada por

$$a_m(t)x^{(m)}(t) + \cdots + a_1(t)x'(t) + a_0(t)x(t) = y(t) \quad t \in [0,1] \tag{7}$$

donde $a_j \in C[0,1]$ para todo $1 \leq j \leq m$ y $a_m(t) \neq 0$ para todo $t \in [0,1]$.

Consideramos las condiciones iniciales

$$x(0) = x'(0) = \cdots = x^{(m-1)}(0) = 0 \tag{8}$$

Se demuestra en los cursos elementales de ecuaciones diferenciales que para todo $y \in C[0,1]$ existe una única solución de (7) que verifica (8). Consideramos los espacios vectoriales $Y = C[0,1]$, $X = \{x \in C^{(m)}[0,1]$ tales que $x(0) = x'(0) = \cdots = x^{(m-1)}(0) = 0\}$. Y es un espacio de Banach con la norma del supremo, y se puede comprobar que si dotamos a X de la norma

$$\|x\| = \|x\|_\infty + \|x'\|_\infty + \cdots + \|x^{(m)}\|_\infty$$

entonces también X es un espacio de Banach. Definimos ahora la aplicación

$$T : X \longrightarrow Y$$

dada por

$$T(x) = a_m x^{(m)} + \cdots + a_1 x' + a_0 x.$$

Es sencillo verificar que T es lineal y continua. Del teorema de existencia y unicidad antes mencionado se sigue que T es biyectiva.

Supongamos que tenemos un método para solucionar (7) (es decir, para calcular $T^{-1}(y)$) cuando y es un polinomio. Del Teorema de Weierstrass se sigue que los polinomios son densos en Y. Sea ahora $y_0 \in Y$ y sea $(p_n) \subset Y$ una sucesión de polinomios que converge a y_0. Entonces si x_n es la solución de (7) cuando el término independiente es p_n, tenemos que $x_n \to x_0$, donde x_0 es la única solución de (7) verificando las condiciones iniciales (8) cuando el término independiente es y_0. En particular esto implica que

$$x_n \overset{\|\cdot\|_\infty}{\to} x_0,$$

$$x_n' \overset{\|\cdot\|_\infty}{\to} x_0'$$

$$\vdots$$
$$x_n^{(m)} \overset{\|\cdot\|_\infty}{\longrightarrow} x_0^{(m)}$$

Prácticas sugeridas

EJERCICIO 4.5. *Sea X un espacio de Banach. Entonces una base algebraica suya tiene cardinal finito o no contable.*

Solución: *Supongamos que X admitiera una base contable $(e_i)_{i\in\mathbb{N}}$. Para cada $n \in \mathbb{N}$ sea F_n el subespacio generado por $\{e_1, \ldots, e_n\}$. F_n es cerrado por ser de dimensión finita y puesto que $(e_i)_{i\in\mathbb{N}}$ es una base algebraica, se tiene que $X = \cup_{n\in\mathbb{N}} F_n$. Por el Teorema de Baire tiene que existir $n_0 \in \mathbb{N}$ tal que F_{n_0} tiene interior no vacío. Pero si F_{n_0} contiene una bola abierta entonces contiene también una bola abierta centrada en el origen de manera que F_{n_0} es un entorno del origen, lo que implica que $X = F_{n_0}$, una contradicción con que la dimensión de X sea infinita.*

EJERCICIO 4.6. *Sean X, Y, Z espacios normados, con X o Y espacio de Banach. Sea*

$$F : X \times Y \longrightarrow Z$$

una aplicación bilineal separadamente continua (esto es, para cada $x \in X$ la aplicación lineal $F_x : Y \longrightarrow Z$ dada por

$$F_x(y) = F(x, y)$$

es continua, y análogamente en la segunda variable. Entonces F es continua, es decir existe $\alpha > 0$ tal que para todo $(x, y) \in X \times Y$

$$\|F(x, y)\| \leq \alpha \|x\| \|y\|$$

EJERCICIO 4.7. *Los Teoremas de la Aplicación Abierta y de la Gráfica Cerrada fallan si no hay completitud.* [**31**, Example 10.7, p.177]

EJERCICIO 4.8. *Sea X un espacio normado en el que consideramos dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$. Si X es un espacio de Banach con ambas y si existe $C \in \mathbb{R}_+$ tal que $\|\cdot\|_1 \leq C\|\cdot\|_2$ entonces las dos normas son equivalentes.*

EJERCICIO 4.9. *Sean X, Y espacios de Banach, $T : X \longrightarrow Y$ un operador sobreyectivo. Entonces Y es isomorfo a un cociente de X, en concreto a $X/\ker T$.* **Sugerencia** *Considerar el operador $\theta : X/\ker T \longrightarrow Y$ dado por $\theta([x]) = T(x)$ y verificar que es biyectivo y continuo. A continuación aplicar el Teorema de la Aplicación Abierta.*

EJERCICIO 4.10. *Sea E un espacio compacto y sea $(f_i)_{i=1}^n \subset C(E, \mathbb{R})$ un conjunto que separa puntos de E. Demostrar que E es homeomorfo a un subconjunto de $\mathbb{R}^n$*

EJERCICIO 4.11. *Demostrar que el álgebra generado por $\{1, x^2\}$ es densa en $C[0, 1]$ pero no en $C[-1, 1]$.*

Capítulo 5

Espacios duales y operadores traspuestos

En varios libros de Análisis Funcional se puede leer que el Análisis Funcional es sobre todo dualidad y teoría espectral. Sin entrar en la posible exactitud de esa frase, no cabe duda que la noción de espacio dual es fundamental en muchas de las aplicaciones del Análisis Funcional: muchas propiedades estructurales y geométricas sólo se entienden si se formulan o se estudian en términos del dual, y las nociones de convergencia débil y convergencia débil* necesitan de la definición de espacio dual para su misma definición. Una vez definido el dual de un espacio, dado un operador entre espacios normados tiene sentido definir su traspuesto, lo que hacemos en este capítulo puesto que necesitamos esa noción para poder desarrollar la Teoría Espectral de operadores compactos en los capítulos siguientes.

Dedicamos además buena parte de este capítulo a estudiar los duales de algunos espacios de Banach clásicos: c_0, ℓ_p, $L_p[0,1]$ $(1 \leq p < \infty)$ y $C[0,1]$. El estudio del dual de este último nos lleva a hacer una pequeña presentación de la integral de Riemann-Stieltjes y las funciones de variación acotada. Si los alumnos conocieran Teoría de la Medida estudiaríamos directamente la representación del dual de $C(K)$ en términos de medidas de Radon.

El estudio de los duales de ℓ_p y c_0 lo hemos basado sobre todo en [**24**], el del dual de $C[0,1]$ en [**5**] y el de los duales de L_p en [**42**] y [**31**].

Espacio dual

Comenzamos recordando la definición de espacio dual.

Definición 5.1. *Sea X un espacio normado. Entonces*

$$X^* := \mathcal{L}(X;\mathbb{K}) = \{T : X \longrightarrow \mathbb{K}; T \text{ lineal y continuo } \}$$

Como ya vimos anteriormente todo espacio dual es un espacio de Banach, por la completitud de $\mathbb{K}$.

En este sentido, veamos que si partimos de un espacio normado no completo o de su completado, el dual es el mismo. Notemos que si dos espacios de Banach X, Y, o dos espacios normados en general, son isométricos (es decir existe una aplicación $\theta : X \longrightarrow Y$ lineal, biyectiva y continua tal que $\|\theta(x)\| = \|x\|$ para todo $x \in X$) entonces todas las estructuras de X e Y que conocemos y utilizamos (su estructura conjuntista, su estructura vectorial y su norma, y junto con ella su estructura topológica) se conservan, por lo que no podemos distinguir un espacio del otro. Por tanto, a partir de ahora diremos que $X = Y$ si X e Y son isométricos.

PROPOSICIÓN 5.2. *Sea $X_0 \subset X$ un subespacio denso. Entonces $X_0^* = X^*$ con la identificación natural.*

DEMOSTRACIÓN. Sea $\theta : X^* \longrightarrow X_0^*$ el operador restricción, es decir

$$\theta(T) = T_{|X_0}$$

Claramente θ es lineal y continuo, con

$$\|\theta(T)\| \leq \|T\|$$

De la densidad de B_{X_0} en B_X se sigue que de hecho

$$\|\theta(T)\| = \|T\|.$$

Sólo nos falta ver que θ es sobre. Sea $T_0 \in X_0^*$. El Teorema de Hahn-Banach nos garantiza que existe $T \in X^*$ (con $\|T\| = \|T_0\|$) tal que $\theta(T) = T_0$, lo que termina la demostración. □

Estudiamos a continuación los duales de los espacios de Banach clásicos que hemos visto hasta ahora.

PROPOSICIÓN 5.3 (Riesz). $c_0^* = \ell_1$

DEMOSTRACIÓN. Dado $x^* \in c_0^*$, para cada $n \in \mathbb{N}$ definimos $x_n^* = x^*(e_n)$ donde $e_i = (0, \ldots, 0, \overset{(i)}{1}, 0, 0, \ldots,)$. A continuación Definimos el operador

$$\theta : c_0^* \longrightarrow \ell_1$$

de la siguiente forma

$$\theta(x^*) = (x_n^*)_{n \in \mathbb{N}}$$

Veamos que θ esta bien definido: si $x_n^* = x^*(e_n) = |x_n^*| e^{-\imath\varphi_n}$, para todo $m \in \mathbb{N}$, consideramos el vector

$$v_m = (e^{-\imath\varphi_1}, e^{-\imath\varphi_2}, \ldots, e^{-\imath\varphi_m}, 0, 0, \ldots)$$

. Puesto que $v_m \in B_{c_0^*}$ se tiene que

$$x^*(v_m) = x^* \left(\sum_{n=1}^{m} e^{-\imath\varphi_n} e_n \right) = \sum_{n=1}^{m} e^{-\imath\varphi_n} x^*(e_n) = \sum_{n=1}^{m} |x_n^*| \leq \|x^*\|.$$

Por tanto $\sum_{n=1}^{\infty} |x_n^*|$ existe y es menor o igual que $\|x^*\|$. Esto prueba que θ está bien definido y puesto que claramente θ es lineal, los razonamientos anteriores también prueban que θ es continuo y $\|\theta\| \leq 1$.

Vamos a usar ahora que el completado de $(c_{00}, \|\cdot\|_\infty)$ es c_0, por lo que para todo $x^* \in c_0^*$, se tiene que $\|x^*\| = \sup_{x \in B_{c_{00}}} x^*(x)$.

Dado $x^* \in c_0^*$, para todo $x \in B_{c_{00}}$, suponiendo $x_j = 0$ para todo $j \geq m$, se tiene

$$x^*(x) = x^*\left(\sum_{n=1}^{m} x_n e_n\right) = \sum_{n=1}^{m} x_n x_n^* \leq \|x\|_\infty \sum_{n=1}^{m} |x_n^*| \leq$$

$$\leq \sum_{n=1}^{m} |x_n^*| \leq \sum_{n=1}^{\infty} |x_n^*| = \|\theta(x^*)\|$$

por lo que $\|x^*\| \leq \|\theta(x^*)\|$ y esa era la desigualdad que nos faltaba para probar que θ es una isometría. Al ser isometría automáticamente θ es inyectiva y sólo nos resta comprobar que θ es sobre.

Sea $a = (a_n) \in \ell_1$. Consideramos la forma $x^* \in c_0^*$ dada por

$$x^*(x) = \sum_{n=1}^{\infty} a_n x_n$$

Claramente, para todo $x \in c_0$ la serie $\sum_{n=1}^{\infty} a_n x_n$ es absolutamente convergente, puesto que

$$\sum_{n=1}^{\infty} |a_n x_n| \leq \sup_n |x_n| \sum_{n=1}^{\infty} |a_n| = \|x\|_\infty \|a\|_1.$$

Por tanto x^* es efectivamente un elemento de c_0^* y claramente $\theta(x^*) = a$. □

Proposición 5.4 (Riesz). $\ell_1^* = \ell_\infty$

Demostración. Dado $x^* \in \ell_1^*$ definimos $x_n^* = x^*(e_n)$, y definimos el operador

$$\theta : \ell_1^* \longrightarrow \ell_\infty$$

como

$$\theta(x^*) = (x_n^*)$$

Veamos que θ está bien definido.

$$\|\theta(x^*)\| = \sup_n |x_n^*| = \sup_n |x^*(e_n)| \leq \|x^*\|$$

por lo que θ está bien definido y, puesto que es claramente lineal, también es continuo.

Análogamente a la demostración anterior, vamos a usar ahora que el completado de $(c_{00}, \|\cdot\|_1)$ es ℓ_1, por lo que para todo $x^* \in \ell_1^*$, se tiene que $\|x^*\| = \sup_{x \in B_{c_{00}}} x^*(x)$.

Dado $x^* \in \ell_1$, para todo $x = (x_n) \in B_{c_{00}}$, suponiendo $x_j = 0$ para todo $j \geq m$, se tiene

$$x^*(x) = x^*\left(\sum_{n=1}^{m} x_n e_n\right) = \sum_{n=1}^{m} x_n x_n^* \leq \sup_n |x_n^*| \sum_{n=1}^{m} |x_n| \leq$$

$$\leq \|\theta(x^*)\| \sum_{n=1}^{m} |x_n| \leq \|\theta(x^*)\| \sum_{n=1}^{\infty} |x_n| \leq \|\theta(x^*)\|$$

por lo que $\|x^*\| \leq \|\theta(x^*)\|$ y θ es una isometría. Sólo resta comprobar que θ es sobre.

Sea $a = (a_n) \in \ell_\infty$. Consideramos la forma $x^* \in \ell_1^*$ dada por

$$x^*(x) = \sum_{n=1}^{\infty} a_n x_n$$

Claramente, para todo $x \in \ell_1$ la serie $\sum_{n=1}^{\infty} a_n x_n$ es absolutamente convergente, puesto que

$$\sum_{n=1}^{\infty} |a_n x_n| \leq \sup_n |a_n| \sum_{n=1}^{\infty} |x_n| = \|a\|_\infty \|x\|_1.$$

Por tanto x^* es efectivamente un elemento de ℓ_1^* y claramente $\theta(x^*) = a$. □

PROPOSICIÓN 5.5 (Riesz). *Sean $1 < p, q < \infty$ tales que $\frac{1}{p} + \frac{1}{q} = 1$. Entonces $\ell_p^* = \ell_q$.*

DEMOSTRACIÓN. Dado $x^* \in \ell_p^*$, para cada $n \in \mathbb{N}$ definimos $x_n^* = x^*(e_n)$.

Entonces definimos el operador

$$\theta : \ell_p^* \longrightarrow \ell_q$$

de la siguiente forma

$$\theta(x^*) = (x_n^*)_{n \in \mathbb{N}}$$

Veamos que θ esta bien definido: Dado $x^* \in \ell_p^*$, para todo $m \in \mathbb{N}$, consideramos el vector $v_m = (|x_1^*|^{q-1} e^{-\imath\varphi_1}, \ldots, |x_m^*|^{q-1} e^{-\imath\varphi_m}, 0, 0, \ldots)$.

Notemos que si $\frac{1}{p} + \frac{1}{q} = 1$ entonces $\frac{1}{p} = \frac{q-1}{q}$ por lo que $p(q-1) = q$. Así

$$\sum_{n=1}^{m} |x_n^*|^q = \sum_{n=1}^{m} |x_n^*|^{q-1}|x_n^*| = \sum_{n=1}^{m} |x_n^*|^{q-1}e^{-\imath\varphi_n}x^*(e_n) =$$

$$= x^*\left(\sum_{n=1}^{m} |x_n^*|^{q-1}e^{-\imath\varphi_n}e_n\right) = x^*(v_m) \leq \|x^*\|\|v_m\|_p =$$

$$= \|x^*\|\left(\sum_{n=1}^{m} |x_n^*|^{p(q-1)}\right)^{\frac{1}{p}} = \|x^*\|\left(\sum_{n=1}^{m} |x_n^*|^{q}\right)^{\frac{1}{p}}$$

Hemos probado que

$$\sum_{n=1}^{m} |x_n^*|^q \leq \|x^*\|\left(\sum_{n=1}^{m} |x_n^*|^{q}\right)^{\frac{1}{p}}$$

y de aquí se sigue que

$$\left(\sum_{n=1}^{m} |x_n^*|^q\right)^{\frac{1}{q}} = \left(\sum_{n=1}^{m} |x_n^*|^{q}\right)^{1-\frac{1}{p}} \leq \|x^*\|.$$

Por tanto θ está bien definido y es lineal y continuo con $\|\theta\| \leq 1$.

Vamos a usar ahora que el completado de $(c_{00}, \|\cdot\|_p)$ es ℓ_p, por lo que para todo $x^* \in \ell_p^*$, se tiene que $\|x^*\| = \sup_{x\in B_{c_{00}}} x^*(x)$.

Dado $x^* \in \ell_p^*$, para todo $x \in B_{c_{00}}$, suponiendo $x_j = 0$ para todo $j \geq m$, se tiene, utilizando la desigualdad de Hölder, que

$$x^*(x) = x^*\left(\sum_{n=1}^{m} x_n e_n\right) = \sum_{n=1}^{m} x_n x_n^* \leq \|x\|_p\|\theta(x^*)\|_q$$

por tanto $\|x^*\| \leq \|\theta(x^*)\|$ y θ es una isometría. Veamos que es sobreyectiva:

Sea $a = (a_n) \in \ell_q$. Consideramos la forma $x^* \in \ell_p^*$ dada por

$$x^*(x) = \sum_{n=1}^{\infty} a_n x_n$$

De nuevo por la desigualdad de Hölder, para todo $x \in \ell_p$ la serie $\sum_{n=1}^{\infty} a_n x_n$ es absolutamente convergente, puesto que

$$\sum_{n=1}^{\infty} |a_n x_n| \leq \|(a_n)\|_q\|(x_n)\|_p.$$

Por tanto x^* es efectivamente un elemento de ℓ_p^* y claramente $\theta(x^*) = a$. $\square$

El dual de $C[0,1]$

Vamos a describir el dual de $C[0,1]$ como cierto subconjunto de las funciones de variación acotada, utilizando la Integral de Riemann-Stieltjes, que quizás nuestros alumnos ya conozcan. Si no la conocen, esta puede ser una buena oportunidad para presentarla y que se familiaricen con sus propiedades básicas.

Comenzamos definiendo funciones de variación acotada.

Definición 5.6. *Una función $g : [0,1] \longrightarrow \mathbb{R}$ se llama de* variación acotada *si su* variación $V(g)$ *es finita, es decir, si*

$$V(g) := \sup_{0<x_0 \ldots <x_n=1} \left\{ \sum_{i=1}^{n} |g(x_i) - g(x_{i-1})| \right\} < \infty.$$

Es muy fácil ver que con las operaciones puntuales, el conjunto de las funciones de variación acotada $BV[0,1]$ es un espacio vectorial. Dado que $V(g) = 0$ implica que g es constante, es también fácil ver que

$$\|g\|_v = |g(0)| + V(g)$$

define una norma sobre $BV[0,1]$.

Recordemos la definición de la Integral de Riemann-Stieltjes.

Dadas $f, g : [0,1] \longrightarrow \mathbb{R}$, para cada partición de $[0,1]$ $\pi = \{0 = x_0 < \cdots < x_n = 1\}$ y para cada vector $\xi = (\xi_1 \ldots, \xi_n)$, con $\xi_i \in [x_{i-1}, x_i]$ definimos una suma de Riemann-Stieltjes de f respecto de g como

$$S(f, \pi, \xi; g) = \sum_{i=1}^{n} f(\xi_i)(g(x_i) - g(x_{i-1}))$$

La función f se dice integrable Riemann-Stieltjes respecto de g, y escribimos $f \in \mathcal{R}(g)$, si existe el límite de la red $\{S(f, \pi, \xi_\pi; g)\}_\pi$ cuando π recorre el conjunto dirigido de todas las particiones de $[0,1]$ y ξ_π es una elección fija de ξ para cada π. A este límite si existe lo llamaremos

$$\int_0^1 f dg$$

y se puede demostrar que no depende de la elección de ξ.

Las siguientes propiedades de la integral de Riemann-Stieltjes son fáciles de probar y se pueden ver por ejemplo en [**2**].

- El conjunto $\mathcal{R}(g)$ de las funciones integrables Riemann-Stieltjes respecto de g es un espacio vectorial y la aplicación
$$T_g : \mathcal{R}(g) \longrightarrow \mathbb{R}$$
definida como
$$T_g(f) = \int_0^1 f dg$$
es una forma lineal.
- Aditividad respecto al intervalo: Para todo $c \in (0,1)$ y para todo $f \in \mathcal{R}(g)$,
$$\int_0^1 f dg = \int_0^c f dg + \int_c^1 f dg$$
- Linealidad respecto al integrador. Si $f \in \mathcal{R}(g) \cap \mathcal{R}(h)$, entonces $f \in \mathcal{R}(\alpha g + \beta h)$ para todo $\alpha, \beta \in \mathbb{R}$ y
$$\int_0^1 f d(\alpha g + \beta h) = \alpha \int_0^1 f dg + \beta \int_0^1 h dg$$
- Si g es de variación acotada entonces toda función continua f es integrable. En ese caso no es necesario tomar el límite en la red, sino que se puede considerar una *sucesión* $(\pi)_n$ de particiones que verifique
$$\|\pi_n\| := \text{máx}\{|x_i - x_{i-1}|; \, 1 \leq i \leq n\} \to 0.$$
Además siempre
$$S(f, \pi, \xi; g) \leq \|f\|_\infty \|g\|_v$$
por lo que
$$\left| \int_0^1 f dg \right| \leq \|f\|_\infty \|g\|_v$$

Por tanto, la aplicación $T_g : C[0,1] \longrightarrow \mathbb{R}$ es una forma lineal y continua para cada $g \in BV[0,1]$, y $\|T_g\| \leq V(g) \leq \|g\|_v$. Lo interesante ahora es que también podemos proceder en sentido opuesto.

Teorema 5.7 (Riesz). *Si $T : C[0,1] \longrightarrow \mathbb{R}$ es una forma lineal y continua, entonces existe $g \in BV[0,1]$ con $g(0) = 0$ tal que $T = T_g$ y* $\|T\| = \|g\|_v$

DEMOSTRACIÓN. Sea $T \in C[0,1]^*$. Si existiera $g \in BV[0,1]$ con $g(0) = 0$ tal que $T = T_g$, tendríamos

$$g(x) = g(x) - g(0) = \int_0^x dg = \int_0^1 \chi_{[0,x]} dg = T(\chi_{[0,x]})$$

y parece que bastará con definir

$$g(x) = T(\chi_{[0,x]}).$$

El problema es que en general $\chi_{[0,x]} \notin C[0,1]$, por lo que la expresión $T(\chi_{[0,x]})$ no tiene sentido. Será el Teorema de Hahn-Banach quien acuda en nuestra ayuda. Notemos que $C[0,1]$ es un subespacio vectorial (incluso cerrado) de $B[0,1]$, las funciones reales acotadas definidas sobre $[0,1]$. Podemos entonces extender T a una forma $\overline{T} \in B[0,1]^*$ con $\|T\| = \|\overline{T}\|$. Definamos las funciones φ_x como

$$\varphi_x = \chi_{[0,x]} \text{ si } x \neq 0$$

y

$$\varphi_0 = 0.$$

Ahora ya podemos definir $g : [0,1] \longrightarrow \mathbb{R}$ como

$$g(x) = \overline{T}(\varphi_x) \text{ para todo } x \in [0,1]$$

Veamos que g cumple lo que esperamos de ella.

En primer lugar, $g \in BV[0,1]$. En efecto, para toda partición

$$\pi = \{0 = x_0 < x_1 < \cdots < x_n = 1\}$$

tenemos

$$\sum_{i=1}^n |g(x_i) - g(x_{i-1})| = \sum_{i=1}^n \epsilon_i (g(x_i) - g(x_{i-1})) =$$

$$= \overline{T}\left(\sum_{i=1}^n \epsilon_i \left(\varphi_{x_i} - \varphi_{x_{i-1}}\right)\right) \leq \|T\| \left\|\sum_{i=1}^n \epsilon_i \left(\varphi_{x_i} - \varphi_{x_{i-1}}\right)\right\| =$$

$$= \|T\| \left\|\sum_{i=1}^n \epsilon_i \chi_{(x_{i-1},x_i]}\right\| = \|T\|,$$

donde $\epsilon_i = sgn(g(x_i) - g(x_{i-1}))$.

Por lo tanto $g \in BV[0,1]$ y $V(g) \leq \|T\|$. Como además $g(0) = 0$, tenemos

$$V(g) = \|g\|_v \leq \|T\|.$$

Veamos ahora que $T = T_g$. Sea $f \in C[0,1]$. Definamos

$$f_n(t) = \sum_{k=1}^n f\left(\frac{k}{n}\right) (\varphi_{\frac{k}{n}}(t) - \varphi_{\frac{(k-1)}{n}}(t)).$$

Por ser f uniformemente continua se tiene que f_n tiende a f en la norma uniforme ($\|\cdot\|_\infty$). Por lo tanto

$$\overline{T}(f_n) = \sum_{k=1}^{n} f\left(\frac{k}{n}\right)\left(g\left(\frac{k}{n}\right) - g\left(\frac{(k-1)}{n}\right)\right) \to \overline{T}(f) = T(f).$$

Pero, por las propiedades que hemos enunciado y la definición de la integral de Riemann-Stieltjes se tiene que también

$$\overline{T}(f_n) = \sum_{k=1}^{n} f\left(\frac{k}{n}\right)\left(g\left(\frac{k}{n}\right) - g\left(\frac{(k-1)}{n}\right)\right) \to \int_0^1 f dg.$$

En consecuencia

$$T(f) = \int_0^1 f dg,$$

es decir, $T = T_g$ y por lo visto anteriormente $\|T\| = \|g\|_v$. □

Ya hemos visto que todo elemento de $C[0,1]^*$ se puede ver como una función $g \in BV[0,1]$. Nos gustaría ahora para que la identificación fuera perfecta tener unicidad en la representación. Eso no es posible, ya que como sabemos las extensiones por Hahn-Banach en general no son únicas, y para cada extensión de T tenemos una g diferente. Lo que vamos a hacer es identificar un cierto subespacio de $BV[0,1]$ como el dual de $C[0,1]$. Para ello necesitaremos algunos resultados previos

Proposición 5.8. *Sea $g \in BV[0,1]$. Entonces $T_g = 0$ si y sólo si, para todo $c \in (0,1)$,*

$$g(0) = g(1) = g(c^+) = g(c^-),$$

donde

$$g(c^+) = \lim_{x\to c^+} g(x)$$

y análogamente $g(c^-)$.

Demostración. Sea g tal que $T_g = 0$ y sea $c \in (0,1)$. Para cada $h > 0$ tal que $c + h \leq 1$ definimos

$$f_h(t) = \begin{cases} 1 & \text{si } 0 \leq t \leq c \\ 0 & \text{si } c+h \leq t \leq 1 \\ \text{y lineal en el resto} & \end{cases}$$

Entonces

$$T_g(f_h) = 0 = \int_0^1 f_h dg = g(c) - g(0) + \int_c^{c+h} f_h dg$$

Integrando por partes se tiene

$$\int_c^{c+h} f_h dg = -g(c) + \frac{1}{h}\int_c^{c+h} g(t)dt$$

luego

$$g(0) = \frac{1}{h}\int_c^{c+h} g(t)dt \to g(c^+)$$

cuando $h \to 0$.

Análogamente se prueba que $g(1) = g(c^-)$. Finalmente tomando $f = \mathbb{1}$ se tiene que $0 = g(1) - g(0)$.

Recíprocamente, si g es como en la hipótesis, entonces g es constante en sus puntos de continuidad, que forman un conjunto denso de $[0,1]$ (pues su complementario es a lo sumo numerable). Tomando una sucesión de particiones con norma tendiendo a 0 formadas exclusivamente por puntos de continuidad de g, las correspondientes sumas son todas nulas, y por tanto su límite $\int_0^1 f dg$ también lo es. □

Definición 5.9. *Una función $g \in NBV[0,1]$ se llama* normalizada *si $g(0) = 0$ y g es continua por la derecha en todo punto $c \in (0,1)$. Llamamos $NBV[0,1]$ al subespacio vectorial de $BV[0,1]$ formado por las funciones normalizadas.*

Proposición 5.10. *Si $g_1, g_2 \in NBV[0,1]$ cumplen que $T_{g_1-g_2} = 0$, entonces $g_1 = g_2$*

Demostración. Por ser $g_1, g_2 \in NBV[0,1]$ se tiene

$$g_1(0) = g_2(0) = 0.$$

Por lo tanto $(g_1 - g_2)(0) = 0$. Por la Proposición 5.8 se tiene $g_1(1) - g_2(1) = g_1(0) - g_2(0)$ y por tanto $g_1(1) = g_2(1)$, y análogamente para todo $c \in (0,1)$ usando la continuidad por la derecha de g_i □

Proposición 5.11. *Para cada $g \in BV[0,1]$ existe una única función $\tilde{g} \in NBV[0,1]$ tal que $T_{g-\tilde{g}} = 0$. Además $V(\tilde{g}) \leq V(g)$.*

Demostración. La unicidad se sigue de la proposición anterior. En cuanto a la existencia, basta con definir

$$\tilde{g}(c) = \begin{cases} 0 & \text{si } c = 0 \\ g(1) - g(0) & \text{si } c = 1 \\ g(c^+) - g(0) & \text{si } c \in (0,1) \end{cases}$$

Veamos que $\tilde{g}$ cumple lo pedido. Sea $c \in (0,1)$. Sea $\epsilon > 0$. Existe $\delta > 0$ tal que si $c < s < c + \delta$ entonces

$$|g(c^+) - g(s)| \leq \epsilon.$$

Por lo tanto, tomando $(t_n) \subset (s, c+\delta)$ una sucesión decreciente que converge a s (esto lo escribimos $(t_n) \searrow s$) y dejando que n tienda a infinito se tiene que, para todo $s \in (c, c+\delta)$,

$$|g(c^+) - g(s^+)| \leq \epsilon.$$

De aquí se sigue que

$$\lim_{s_n \searrow c} \tilde{g}(s_n) = \tilde{g}(c)$$

y por lo tanto $\tilde{g}$ es continua a la derecha en todo $c \in [0,1)$.

De forma análoga se prueba que

$$\tilde{g}(c^-) = g(c^-) - g(0),$$

y por lo tanto la función $\tilde{g} - g$ verifica las condiciones de la Proposición 5.8. Por tanto ya sólo falta probar que $V(\tilde{g}) < \infty$. Para esto, sea

$$0 = t_0 < t_1 < \cdots < t_n = 1$$

una partición de $[0,1]$ y sea $\epsilon > 0$. Sean $s_i \in (t_i, t_{i+1})$ $(0 \leq i \leq n)$ tales que

$$|g(t_i^+) - g(s_i)| < \frac{\epsilon}{n}.$$

Entonces, haciendo $s_n = t_n = 1$ se tiene

$$\sum_{i=1}^{n} |\tilde{g}(t_i) - \tilde{g}(t_{i-1})| \leq 2\epsilon + \sum_{i=1}^{n} |g(s_i) - g(s_{i-1})| \leq 2\epsilon + V(g),$$

de donde

$$V(\tilde{g}) \leq V(g)$$

y se concluye la demostración.

□

Finalmente podemos enunciar

Teorema 5.12 (Representación de Riesz). $C[0,1]^* = NBV[0,1]$.

Demostración. Sea

$$\theta : NBV[0,1] \longrightarrow C[0,1]^*$$

la aplicación definida como

$$\theta(g) = T_g.$$

Ya hemos visto que θ es lineal, continua y que

$$\|\theta(g)\| = \|T_g\| \leq \|g\|_v.$$

Por otro lado, si $T \in C[0,1]^*$ hemos visto que existe $g \in BV[0,1]$ tal que $T_g = T$ y que existe una única $\tilde{g} \in NBV[0,1]$ tal que $T_{\tilde{g}-g} = 0$, es decir $T_{\tilde{g}} = T_g = T$. Por lo tanto θ es sobreyectiva y una isometría ya que

$$\|T\| = \|\theta(\tilde{g})\| \leq \|\tilde{g}\|_v \leq \|g\|_v = \|T\|.$$

□

Los duales de los $L_p[0,1]$

Utilizamos en esta sección nuestros conocimientos de la Teoría Integral de Lebesgue para estudiar los duales de $L_p[0,1]$.

TEOREMA 5.13. *Sea $1 \leq p, q \leq \infty$ con $\frac{1}{p} + \frac{1}{q} = 1$. Dado $g \in L_q[0,1]$ sea $T_g : L_p[0,1] \longrightarrow \mathbb{K}$ la aplicación definida como*

$$T_g(f) = \int_0^1 fgd\lambda \text{ para todo } f \in L_p[0,1].$$

Entonces $T_g \in (L_p[0,1]^*$ y $\|T_g\| = \|g\|_q$. Por lo tanto la aplicación

$$\theta : L_q[0,1] \longrightarrow (L_p[0,1])^*$$

dada por $\theta(g) = T_g$ es una isometría (aún no afirmamos que sea sobreyectiva).

DEMOSTRACIÓN. La demostración es bastante similar a la de los duales de los ℓ_p. Los detalles se pueden ver por ejemplo en [**31**, Theorem 14.1]. □

Se puede probar además que en el caso $1 \leq p < \infty$ la isometría θ del teorema anterior es sobreyectiva y por lo tanto $(L_p[0,1])^* = L_q[0,1]$. La demostración más habitual de este hecho utiliza el Teorema de Radon-Nikodym, un resultado profundo de Teoría de la Medida (o de la Teoría de la Integral de Lebesgue) que preferimos no tener que exponer. Los alumnos interesados pueden leer una demostración laboriosa pero elemental (sin Radon-Nikodym) de la sobreyectividad de θ en [**31**, Theorem 14.3].

Traspuesto de un operador

Relacionado con la idea de espacio dual está la noción de *traspuesto* de un operador.

Definición 5.14. *Sean X, Y espacios normados, y sea $T : X \longrightarrow Y$ un operador. Definimos el traspuesto de T, y lo denotamos T^*, como el operador*

$$T^* : Y^* \longrightarrow X^*$$

definido por

$$T^*(y^*)(x) = y^*(T(x)).$$

Para todo $x \in B_X$, $y^* \in B_{Y^*}$ se tiene que

$$|T^*(y^*)(x)| = |y^*(T(x)| \leq \|T(x)\| \leq \|T\|$$

de donde se obtiene que T^* está bien definido.Claramente es lineal y de la misma desigualdad de arriba se sigue que T^* es continuo y

$$\|T^*\| \leq \|T\|$$

Para obtener la igualdad de las normas podemos tomar supremos en la expresión de arriba, o considerar $T^{**} = (T^*)^* : X^{**} \longrightarrow Y^{**}$. Es fácil ver que

$$T^{**}(J(x)) = T(x),$$

(esto se puede proponer como ejercicio) por lo que

$$\|T\| \leq \|T^{**}\| \leq \|T^*\| \leq \|T\|$$

de donde se sigue la igualdad de todas las normas involucradas.

La notación de traspuesto de un operador se justifica con el siguiente ejemplo que proponemos como ejercicio.

Ejercicio 5.1. *Sea $X = \mathbb{K}^n$, $Y = \mathbb{K}^m$. Consideramos en ambos espacios por ejemplo las bases canónicas. Sabemos que todo operador $T : X \longrightarrow Y$ se representa de forma única (respecto de las bases elegidas) como una matriz $A = (a_{ij})$ donde para todo $1 \leq j \leq n$*

$$T(e_j) = \sum_{i=1}^{m} a_{ij} e_i$$

Entonces si consideramos también las bases canónicas en los espacios duales $(\mathbb{K}^n)^* = \mathbb{K}^n$ y $(\mathbb{K}^m)^* = \mathbb{K}^n$ demostrar que la matriz traspuesta A^T es la matriz que representa respecto de esas bases al operador $T^* : Y^* \longrightarrow X^*$

Ejercicio 5.2. $(T + S)^* = T^* + S^*$, $(T \circ S)^* = S^* \circ T^*$

Definición 5.15. *Sean X, Y espacios de Banach. Un operador $T : X \longrightarrow Y$ se dice que es un* isomorfismo inyectivo *si existen dos constantes $c, C > 0$ tales que para todo $x \in X$*

$$c\|x\| \leq \|T(x)\| \leq C\|T(x)\|.$$

EJERCICIO 5.3. *Sean* X, Y *espacios de Banach. Si* $T : X \longrightarrow Y$ *isomorfismo inyectivo entonces* $T(X) \subset Y$ *es un subespacio cerrado y* $T(X)$ *y* X *son isomórfos.*

PROPOSICIÓN 5.16. *Sean* X, Y *espacios de Banach,* $T : X \longrightarrow Y$ *un operador. Entonces* T *es un isomorfismo inyectivo si y sólo si* T^* *es sobre. Análogamente,* T *es sobre si y sólo si* T^* *es un isomorfismo inyectivo.*

DEMOSTRACIÓN. Supongamos que T es un isomorfismo. Entonces $T(X) \subset Y$ es un subespacio cerrado isomorfo a X. Por tanto $T(X)^*$ es isomorfo a X^*. Para ver que $T^* : Y^* \longrightarrow X^*$ es sobreyectiva, dado $x^* \in X^*$ podemos considerar el elemento de $e^* \in T(X)^*$ dado por $e^*(T(x)) = x^*(x)$. Extendemos e^* por Hahn-Banach a un elemento $y^* \in Y^*$ con $\|y^*\| = \|e^*\|$. Sea entonces $T^* : Y^* \longrightarrow X^*$. Se tiene que, para todo $x \in X$

$$T^*(y^*)(x) = y^*T(x) = e^*T(x) = x^*(x)$$

y por tanto $T^*(y^*) = x^*$.

Recíprocamente, si $T^* : Y^* \longrightarrow X^*$ es sobreyectivo, para todo $x \in X$, utilizando el Teorema de la Aplicación Abierta se tiene que

$$\begin{aligned} \|x\| = \sup_{x^* \in B_{X^*}} \|x^*(x)\| \leq \sup_{y^* \in \gamma B_{Y^*}} \|T^*(y^*)(x)\| = \\ = \sup_{y^* \in \gamma B_{Y^*}} \|y^*(T(x))\| \leq \|\gamma\| \|T\| \|x\| \end{aligned}$$

y esta es la condición que necesitamos para que T sea un isomorfismo inyectivo.

La demostración de la segunda mitad del teorema es análoga y se deja como ejercicio. $\square$

Prácticas sugeridas

EJERCICIO 5.4. *Sean* X, Y *espacios de Banach,* $T : X \longrightarrow Y$ *un operador. Entonces* $T(X)$ *es denso si y sólo si* T^* *es inyectivo.*

EJERCICIO 5.5. *Demostrar que la identidad* $T : \ell_1 \longrightarrow \ell_2$ *es inyectivo, pero* T^* *no es sobre (porque* ℓ_2 *es separable y* ℓ_∞ *no lo es).*

EJERCICIO 5.6. *Demostrar que si* $1 \leq p < \infty$ *y* $\frac{1}{p} + \frac{1}{q} = 1$ *entonces* $L_p^*[0,1] = L_q[0,1]$.

EJERCICIO 5.7. *Probar que* $\|\cdot\|_v$ *es una norma en* $BV[0,1]$.

Capítulo 6

Topologías débil y débil*

Como ya dijimos al comentar el Teorema de Hahn-Banach, dado un espacio normado (completo o no) X, dicho teorema nos garantiza la existencia de un dual X^* con una estructura "suficientemente rica". Resulta entonces que el estudio de X^*, y de X con la topología inducida por X^* (la topología débil que definimos en este capítulo) tienen gran importancia en muchos resultados fundamentales del Análisis Funcional. Con razonamientos similares se puede estudiar en X^* la topología inducida por X (la topología débil* que definimos más adelante) y veremos que ésta también tiene múltiples aplicaciones.

Además de las definiciones y propiedades básicas de las topologías débil y débil*, estudiamos en este capítulo los duales de X y X^* con las topologías débil y débil* respectivamente. También presentamos en Teorema de Alaoglu y el de Goldstine, y un breve estudio de la reflexividad en espacios de Banach.

Opcionalmente se podrían estudiar también el Teorema de Eberlein y el Teorema de Mazur acerca de la coincidencia de las clausuras en las topología débil y de la norma para espacios convexos. Este último teorema resulta además una bonita aplicación del Teorema de Hahn-Banch, al igual que el Teorema de Goldstine.

Este capítulo tal y como lo hemos planteado es algo más "topológico" que los anteriores, aunque los únicos conocimientos topológicos realmente necesarios son la noción de *red*, la misma noción de *topología* y la de *base* de una topología. Si los alumnos no estuvieran en posesión de estas nociones topológicas básicas, se podría plantear el estudio de este capítulo como se hace en [**31**], definiendo tan sólo la convergencia débil y débil* de *sucesiones*. Con eso es suficiente para probar, por ejemplo, una versión débil del Teorema de Alaouglu-Bourbaki, que dice que la bola unidad del dual de un espacio separable es débil* secuencialmente compacta. También se puede probar el Teorema de Eberlein, la propiedad de Schur de ℓ_1 y varias aplicaciones interesantes de la convergencia débil y débil*.

Nosotros hemos basado nuestra presentación de este capítulo sobre todo en [**15**] y [**24**].

La topología débil

DEFINICIÓN 6.1. *Sea X un espacio normado con dual X^*. Se define la topología débil de X como la topología generada por la siguiente base de entornos: dados $x_0 \in X$, $n \in \mathbb{N}$, $x_1^*, \ldots, x_n^* \in X^*$, y $\epsilon > 0$ un entorno débil de x_0 viene definido por*

$$W(x_0; x_1^*, \ldots, x_n^*; \epsilon) = \{x \in X \text{ tales que } |x_i^*(x - x_0)| < \epsilon$$

$$\text{para todo } 1 \leq i \leq n\}$$

Veamos que la topología débil es separada (Hausdorff). Dados $x \neq y \in X$, $\|x-y\| \neq 0$ por lo que el Teorema de Hahn-Banach nos garantiza que existe $x^* \in X^*$ tal que

$$x^*(x - y) > 0,$$

es decir, $x^*(x) - x^*(y) > 0$. Si $\epsilon = x^*(x) - x^*(y)$ entonces

$$x \in W(x; x^*; \frac{\epsilon}{3}),\ y \in W(y; x^*; \frac{\epsilon}{3}) \text{ y } W(x; x^*; \frac{\epsilon}{3}) \cap W(y; x^*; \frac{\epsilon}{3}) = \emptyset$$

La topología débil es una topología vectorial, es decir, la suma y el producto son continuos. Veámoslo para la suma, el producto es similar. Sean $x, y \in X$. Sea $W(x+y; x_1^*, \ldots, x_n^*; \epsilon)$ un entorno de $x+y$. Entonces tomando los entornos $W(x; x_1^*, \ldots, x_n^*; \frac{\epsilon}{2})$ y $W(y; x_1^*, \ldots, x_n^*; \frac{\epsilon}{2})$ se tiene que $W(x; x_1^*, \ldots, x_n^*; \frac{\epsilon}{2}) + W(y; x_1^*, \ldots, x_n^*; \frac{\epsilon}{2}) \subset W(x+y; x_1^*, \ldots, x_n^*; \epsilon)$, lo que prueba que la suma es continua.

Nótese que si X tiene dimensión infinita entonces los entornos débiles (por ejemplo del 0) son bastante "grandes". Por ejemplo, un entorno $W(0; x_1^*, \ldots, x_n^*; \epsilon)$ contiene al subespacio

$$\cap_{i=1}^n \ker x_i^*,$$

un subespacio de codimensión finita. Y eso es *muy* grande.

Como consecuencia de este comentario se tiene la siguiente proposición.

TEOREMA 6.2. *Sea X un espacio de Banach. La topología débil coincide con la topología de la norma si y sólo si* $\dim X < \infty$.

DEMOSTRACIÓN. Supongamos que la topología débil coincide con la topología de la norma. Entonces la bola unidad abierta $U_X = \{x \in$

X tales que $\|x\| < 1\}$ es abierta para la topología débil. Si X tuviera dimensión infinita entonces por el comentario anterior existiría un subespacio no trivial dentro de U_X, algo claramente imposible.

Recíprocamente, supongamos que $\dim X = n < \infty$. Puesto que ya hemos visto que en dimensión finita todas las normas son equivalentes, X es isomorfo a ℓ_∞^n. Veamos ahora que la topología débil y la topología de la norma coinciden sobre X.

Si $U \subset X$ es un abierto débil, ya hemos visto que U es abierto en la topología de la norma.

Por otro lado, si $U \subset X$ es un abierto en la topología de la norma, entonces también $U \subset \ell_\infty^n$ es abierto (en la topología de la norma ∞). Sea $x_0 \in U$. Existe $r > 0$ tal que $B(x_0, r) \subset U$. Pero

$$B(x_0, r) = W(x_0; e_1^*, \ldots, e_n^*; r)$$

y por tanto U es un abierto débil. (Nótese que estamos usando implícitamente que puesto que X es isomorfo a ℓ_∞^n, se tiene que $e_i^* \in X^*$ $(1 \leq i \leq n)$. □

De hecho podemos probar aún más, que la topología débil es metrizable si y sólo si X tiene dimensión finita. Para ello necesitamos previamente un lema algebraico que tendremos ocasión de utilizar en más de una ocasión.

Lema 6.3. *Sea X un espacio vectorial y sean $f, g_1, \ldots, g_n$ formas lineales sobre X tales que*

$$\cap_{i=1}^n \ker g_i \subset \ker f$$

Entonces f es una combinación lineal de los g_i's.

Demostración. Razonamos por inducción sobre n. Para $n = 1$ el lema es fácil y ya lo probamos en el Capítulo 4. Supongamos que ya lo hemos probado para $k \leq n-1$. Entonces dados $f, g_1, \ldots, g_n$ como en el enunciado del teorema aplicamos la hipótesis de inducción a

$$f_{|\ker g_n}, g_{1|\ker g_n}, \ldots, g_{n-1|\ker g_n}$$

y deducimos que, sobre el $\ker g_n$,

$$f = \sum_{i=1}^{n-1} a_i g_i$$

Por tanto $f - \sum_{i=1}^{n-1} a_i g_i$ se anula en $\ker g_n$, es decir

$$\ker g_n \subset \ker \left(f - \sum_{i=1}^{n-1} a_i g_i \right)$$

y por tanto del caso $n = 1$ obtenemos que

$$f - \sum_{i=1}^{n-1} a_i g_i = a_n g_n$$

de donde se obtiene lo pedido. □

Ahora podemos probar que la topología débil nunca es metrizable si X es de dimensión infinita.

PROPOSICIÓN 6.4. *Sea X un espacio normado. En ese caso la topología débil sobre X es metrizable si y sólo si X es de dimensión finita.*

DEMOSTRACIÓN. Ya vimos que si X es de dimensión finita entonces la topología débil coincide con la topología de la norma, y por tanto es metrizable.

Recíprocamente, supongamos que la topología débil es metrizable. Puesto que las topologías metrizables satisfacen el I Axioma de Numerabilidad, debe existir una sucesión $(x_n^*)_n \subset X^*$ tal que para todo entorno U de 0 existe un racional $\mathbb{Q} \ni \epsilon > 0$ y un natural $n(U) \in \mathbb{N}$ tales que U contiene a $W(0; x_1^*, \ldots, x_{n(U)}^*; \epsilon)$. Cada $x^* \in X^*$ genera el entorno débil $W(0; x^*; 1)$ que a su vez contiene uno de los entornos $W(0; x_1^*, \ldots, x_{n(W(0;x^*;1))}^*; \epsilon)$. Pero por el Lema 6.3, esto obliga a que x^* sea una combinación lineal de $x_1^*, \ldots, x_{n(W(0;x^*;1))}^*$. Esto implicaría que X^* tendría una base algebraica numerable, y puesto que X^* es un espacio de Banach, esto sólo puede ocurrir si X^*, y por tanto X tiene dimensión finita. □

De hecho se tiene que si $\dim X < \infty$ entonces existe una única topología vectorial separada sobre X (obviamente normable), aunque no probaremos ese resultado en este curso.

PROPOSICIÓN 6.5. *Sea X un espacio normado. Una red (o una sucesión) $(x_i)_{i\in I} \subset X$ tiende débilmente a $x_0 \in X$ si y sólo si para todo $x^* \in X^*$*

$$\lim_i x^*(x_i) = x^*(x_0)$$

DEMOSTRACIÓN. En efecto, supongamos que $\lim_i x^*(x_i) = x^*(x_0)$ para todo $x^* \in X^*$. Si W es un abierto de la topología débil que contiene a x_0, entonces existe un $W(x_0; x_1^*, \ldots, x_n^*; \epsilon) \subset W$. Puesto que, por hipótesis $\lim_i x_j^*(x_i) = x_j^*(x_0)$ para todo $1 \leq j \leq n$, existe $i_0 \in I$ tal que para todo $i \geq i_0$ (donde "$\geq$" denota el orden en I)

$$|x_j^*(x_i) - x_j^*(x_0)| < \epsilon$$

y por tanto para todo $i \geq i_0$ se tiene que $x_i \in W(x_0; x_1^*, \ldots, x_n^*; \epsilon) \subset W$, es decir $(x_i) \to x_0$ en la topología débil.

Recíprocamente, supongamos que $(x_i) \to x_0$ en la topología débil, y sea $x^* \in X^*$. Para todo $\epsilon > 0$, consideramos el entorno $W(x_0; x^*; \epsilon)$. Entonces existe $i_0 \in I$ tal que para todo $i \geq i_0$ se tiene que $x_i \in W(x_0; x^*; \epsilon)$, lo que implica que

$$|x^*(x_i) - x^*(x_0| < \epsilon$$

y de aquí se sigue que

$$\lim_i x^*(x_i) = x^*(x_0).$$

$\square$

Proposición 6.6. *Sea X un espacio normado y $x^* : X \longrightarrow \mathbb{K}$ una forma lineal. Entonces x^* es continua para la topología de la norma de X si y sólo si x^* es continua para la topología débil de X.*

Demostración. Las bases de entornos utilizadas para definir la topología débil son abiertos en la topología de la norma (porque son $\cap_{i=1}^n (x_i^*)^{-1}(A)$ con A abierto y x_i^* norma-continuos) y de ahí se sigue que los conjuntos abiertos en la topología débil son $\|\cdot\|-$abiertos, es decir, la topología de la norma es más fuerte (tiene más abiertos) que la topología débil (y de ahí el nombre). Por tanto si $x^* : X \longrightarrow \mathbb{K}$ es continuo para la topología débil, también lo es para la topología de la norma.

Recíprocamente, si $x^* \in X^*$, para todo $\epsilon > 0$ el conjunto $(x^*)^{-1}(\epsilon, \epsilon)$ es un abierto débil, y por tanto x^* es débilmente continua. $\square$

Una vez probado esto podemos probar el siguiente teorema

Teorema 6.7. *Sea $T : X \longrightarrow Y$ una aplicación entre dos espacios normados X e Y. Entonces T es norma-norma continua si y sólo si T es débil-débil continua.*

Demostración. Supongamos que T es norma-norma continua. Entonces para todo $y^* \in Y^*$ se tiene que $y^* \circ T : X \longrightarrow \mathbb{K}$ es norma continua; por tanto $y^* \circ T \in X^*$ y del resultado anterior se sigue que $y^* \circ T$ es débil continua, y de aquí se sigue que T es débil-débil continua.

Recíprocamente, si T es débil-débil continua, entonces para todo $y^* \in Y^*$ se tiene que $y^* \circ T$ es débil continua. Por tanto $y^* \circ T$ es norma continua. Ahora el Teorema 4.9, consecuencia del Principio de Acotación Uniforme, nos dice que T es norma-norma continua. □

Probamos a continuación el Teorema de Mazur.

TEOREMA 6.8 (Mazur). *Si K es un conjunto convexo del espacio normado X, entonces la clausura débil de K coincide con su clausura en norma.*

DEMOSTRACIÓN. Puesto que la topología de la norma es más fuerte, se tiene automáticamente que

$$\overline{K}^{\|\cdot\|} \subset \overline{K}^{w}$$

Para el otro contenido, supongamos que existe

$$x_0 \in \overline{K}^{w} \setminus \overline{K}^{\|\cdot\|}$$

Entonces, puesto que $\overline{K}^{\|\cdot\|}$ es convexo, por el teorema de Hahn-Banach existen $x_0^* \in X^*$ y $\alpha < \beta \in \mathbb{R}$ tales que

$$\sup_{x\in\overline{K}^{\|\cdot\|}} \Re(x_0^*(x)) \leq \alpha < \beta \leq \Re(x_0^*(x_0)) \tag{9}$$

Pero puesto que $x_0 \in \overline{K}^{w}$, se tiene que existe una red $(x_i)_{i\in I} \subset K$ tal que

$$x_0 = \lim_i x_i$$

ya vimos anteriormente que esto implicaba que

$$x^*(x_0) = \lim_i x^*(x_i)$$

pero esto es imposible ya que, según (9),

$$\lim_i \Re(x^*(x_i)) \leq \alpha < \beta \leq \Re(x_0^*(x_0))$$

□

El siguiente corolario, útil en ocasiones, se puede proponer como ejercicio

COROLARIO 6.9. *Sea $(x_n) \subset X$ una sucesión que tiende débilmente a 0. Entonces existe una sucesión (σ_n) de combinaciones convexas de los (x_n) tal que (σ_n) tiende en norma a 0.*

Aunque la topología débil no se puede caracterizar en general por sucesiones, muchas de las aplicaciones de la topología débil vendrán precisamente a través del estudio de la convergencia débil de sucesiones. Puesto que la topología de la norma es mas fuerte, es más fácil para una sucesión converger en la topología débil, y la mayoría de las aplicaciones de la topología débil vendrán precisamente de conseguir que una sucesión que no converge en norma sí lo haga débilmente, y buscar las consecuencias que se puedan extraer de eso.

Veamos esto con un poco de detalle

- Sea $(x_n) \subset X$ una sucesión tal que $x_n \stackrel{\|\cdot\|}{\to} x$. Entonces $x_n \stackrel{w}{\to} x$.
  Demostración: Para todo $x^* \in X^*$,
  $$|x^*(x_n) - x^*(x)| = |x^*(x_n - x)| \leq \|x_n - x\| \to 0$$
- Sea la sucesión $(e_n) \subset \ell_2$. Entonces $\|e_n\| = 1$ para todo $n \in \mathbb{N}$ pero $e_n \stackrel{w}{\to} 0$.
  Demostración: La primera afirmación es evidente. Para la segunda, sea $a \in \ell_2^*$. Ya vimos que entonces $a = (a_n) \in \ell_2$ y
  $$a(e_n) = \sum_m a_m \delta_{nm} = a_n$$
  y puesto que $(a_n) \in \ell_2$, se sigue que
  $$\lim_n a(e_n) = \lim_n a_n = 0$$

Puede ocurrir en cambio que en un espacio normado X la convergencia débil de sucesiones coincida con la convergencia en norma de sucesiones. Schur probó que este es el caso si $X = \ell_1$ por lo que a los espacios que verifican esta propiedad se les llama espacios de Schur.

TEOREMA 6.10 (Schur). *En ℓ_1 una sucesión $(x_n)_n$ tiende débilmente a x si y sólo si (x_n) tiende en norma a x.*

DEMOSTRACIÓN. Hacemos la demostración en el caso real. Se adapta al caso complejo de forma estandar y dejamos esto como ejercicio.

Veamos primero que si x_n tiende débilmente a 0 entonces también converge a 0 en norma. Si no fuera así, tomando subsucesiones podemos suponer la existencia de $\epsilon > 0$ tal que para todo $n \in \mathbb{N}$,

$$\|x_n\| = \sum_{i=1}^{\infty} |x_n^i| > \epsilon.$$

Sea N_1 tal que

$$\sum_{i=N_1+1}^{\infty} |x_1^i| < \frac{\epsilon}{5}.$$

Entonces

$$\sum_{i=1}^{N_1} |x_1^i| \geq \frac{4\epsilon}{5},$$

y por tanto

$$\sum_{i=1}^{N_1} \varepsilon_1^i x_1^i \geq \frac{4\epsilon}{5},$$

donde $\varepsilon_1^i = sgn(x_1^i)$. Obsérvese que para una elección arbitraria de signos $\{\varepsilon_i = \pm 1\}$ tales que $\varepsilon_i = \varepsilon_1^i$ si $1 \leq i \leq N_1$ se tiene que

$$\left|\sum_{i=1}^{\infty} \varepsilon_i x_1^i\right| = \left|\sum_{i=1}^{N_1} \varepsilon_1^i x_1^i + \sum_{i=N_1+1}^{\infty} \varepsilon_i x_1^i\right| \geq$$

$$\geq \left|\sum_{i=1}^{N_1} \varepsilon_1^i x_1^i\right| - \sum_{i=N_1+1}^{\infty} |x_1^i| \geq \frac{4\epsilon}{5} - \frac{\epsilon}{5} = \frac{3\epsilon}{5}.$$

A continuación, puesto que (x_n) tiende débilmente a 0, considerando todos los elementos $\varepsilon_{N_1} = (\pm 1, \overset{(N_1)}{\ldots}, \pm 1, 0, 0, \ldots) \in \ell_\infty = (\ell_1)^*$ elegimos n_2 tal que

$$\sum_{i=1}^{N_1} |x_{n_2}^i| < \frac{\epsilon}{5}.$$

Seguidamente consideramos $N_2 > N_1$ tal que

$$\sum_{i=N_2+1}^{\infty} |x_{n_2}^i| < \frac{\epsilon}{5}.$$

Entonces

$$\sum_{i=1}^{N_2} |x_{n_2}^i| \geq \frac{4\epsilon}{5},$$

y por tanto, dados los signos $\varepsilon_{N_1+1} = sgn(x_{n_2}^{N_1+1}), \cdots, \varepsilon_{N_2} = sgn(x_{n_2}^{N_2})$ tenemos que

$$\sum_{i=N_1+1}^{N_2} \varepsilon_i x_{n_2}^i = \sum_{i=N_1+1}^{N_2} |x_{n_2}^i| = \sum_{i=1}^{N_2} |x_{n_2}^i| - \sum_{i=1}^{N_1} |x_{n_2}^i| \geq \frac{4\epsilon}{5} - \frac{\epsilon}{5} = \frac{3\epsilon}{5}.$$

De nuevo nótese que para una elección arbitraria de signos $\{\varepsilon_i = \pm 1\}$ tales que $\varepsilon_i = \varepsilon_1^i$ si $1 \leq i \leq N_1$ y $\varepsilon_i = \varepsilon_{n_2}^i$ si $N_1 \leq i \leq N_2$ se tiene que

$$\left|\sum_{i=1}^{\infty} \varepsilon_i x_{n_2}^i\right| = \left|\sum_{i=N_1+1}^{N_2} \varepsilon_i x_{n_2}^i - \sum_{i=1}^{N_1} \varepsilon_i x_{n_2}^i + \sum_{i=N_2+1}^{\infty} \varepsilon_i x_1^i\right| \geq$$

$$\geq \left| \sum_{i=N_1+1}^{N_2} \varepsilon_i x_{n_2}^i \right| - \left| \sum_{i=1}^{N_1} \varepsilon_i x_{n_2}^i \right| - \left| \sum_{i=N_2+1}^{\infty} \varepsilon_i x_1^i \right| \geq$$

$$\geq \left| \sum_{i=N_1+1}^{N_2} \varepsilon_i x_{n_2}^i \right| - \sum_{i=1}^{N_1} |x_{n_2}^i| - \sum_{i=N_2+1}^{\infty} |x_1^i| \geq$$

$$\geq \frac{3\epsilon}{5} - \frac{\epsilon}{5} - \frac{\epsilon}{5} = \frac{\epsilon}{5}$$

Continuando la construcción por inducción, obtenemos dos sucesiones $1 < n_2 < n_3 < \cdots$ y $N_1 < N_2 < N_3 < \cdots$ y un vector $\varepsilon = (\varepsilon_i) \in \ell_\infty$ definido como

$$\varepsilon_i = sgn(x_{n_k}^i) \text{ si } N_{k-1} < i \leq N_k$$

de manera que, para todo $k \in \mathbb{N}$,

$$\varepsilon(x_{n_k}) \geq \frac{\epsilon}{5}$$

en contradicción con el hecho de que x_n tienda débilmente a 0.

Si ahora x_n tiende débilmente a x razonamos análogamente con la sucesión $(x_n - x)$. □

Hay varias demostraciones de este resultado fundamental. Hemos reproducido la de [**24**, Thm 99] que es elemental y utiliza el método de la "joroba deslizante" que tiene aplicaciones en otros contextos. En [**13**, Prop V.5.2] se puede encontrar una demostración más topológica, que utiliza el Teorema de Baire, el Teorema de Alaoglu y el hecho de que (B_{ℓ_∞}, w^*) es metrizable (por ser ℓ_1 separable). En [**15**, p. 85] aparece otra demostración que utiliza el Lema de Philips y una descripción del dual de ℓ_∞.

La topología débil*

Una vez estudiados algunos de los resultados básicos más importantes referidos a la topología débil, pasamos a estudiar la topología débil* sobre un espacio dual X^*.

Definición 6.11. *Sea X un espacio normado con dual X^*. La topología débil* en X^* es la topología generada por la siguiente base de entornos: dados $x_0^* \in X^*$, $n \in \mathbb{N}$, $x_1, \ldots, x_n \in X$, y $\epsilon > 0$ un entorno débil* de x_0^* viene definido por*

$$W(x_0^*; x_1, \ldots, x_n; \epsilon) = \{x^* \in X^* \text{ tales que}$$

$$|J(x_i)(x^* - x_0^*)| = |(x^* - x_0^*)(x_i)| < \epsilon \text{ para todo } 1 \leq i \leq n\}$$

Dado un espacio dual X^*, es obvio que la topología débil* sobre X^* es en general menos fuerte (tiene menos entornos) que la topología débil sobre X^*, puesto que la topología débil es la generada por todos las formas $x^{**} \in X^{**}$, mientras que la débil* es la generada tan sólo por las formas $J(x) \in X^{**}$, con $x \in X$.

Veamos que a pesar de ello la topología débil* tiene suficientes entornos para ser Hausdorff.

Dados $x^* \neq y^* \in X^*$, existe $x \in X$ tal que

$$(x^* - y^*)(x) > 0,$$

y ahora podemos razonar como lo hacíamos en el caso de la topología débil.

De nuevo es fácil ver que la topología débil* es una topología vectorial.

Siguiendo exactamente la misma demostración que para el caso de la topología débil, *mutatis mutandi* se demuestra que una red (o sucesión) $(x_i^*)_{i\in I} \subset X^*$ tiende débil* a $x_0^* \in X^*$ si y sólo si para todo $x \in X$

$$\lim_i J(x)(x_i^*) \lim_i x_i^*(x) = x_0^*(x)$$

Veamos a continuación que al igual que el dual de X con la topología débil es X^*, también el dual de X^* con la topología débil* es X.

TEOREMA 6.12. *Sea X un espacio de Banach y sea $x^{**} \in X^{**}$ una forma acotada tal que $x^{**} : X^* \longrightarrow \mathbb{K}$ es continuo para la topología débil*. Entonces existe $x \in X$ tal que $J(x) = x^{**}$ (o, dicho más brevemente, $x^{**} \in X$). Recíprocamente, para todo $x \in X$ se tiene que $x \in (X^*, w^*)^*$.*

DEMOSTRACIÓN. Si x^{**} es continuo para la topología débil*, entonces existe un entorno débil* del origen $W := W(0; x_1, \ldots, x_n; \epsilon)$ tal que $|x^{**}(x^*)| < 1$ para todo $x^* \in W$. Sea $x^* \in X^*$ tal que $J(x_i)(x^*) = x^*(x_i) = 0$ para todo $1 \leq i \leq n$. Entonces para todo $\lambda \in \mathbb{K}$, $\lambda x^* \in W$ y por tanto $|x^{**}(\lambda x^*)| < 1$ lo que implica que $x^{**}(x^*) = 0$. Por tanto

$$\cap_{i=1}^n \ker(J(x_i)) \subset \ker x^{**}$$

Ahora el Lema 6.3 nos garantiza que

$$x^{**} = \sum_i a_i J(x_i)$$

y puesto que $J(X)$ es un subespacio vectorial, se sigue que $x^{**} \in J(X)$.

La otra implicación es inmediata. $\square$

Enunciamos y probamos a continuación los Teoremas de Goldstine y Alaoglu.

TEOREMA 6.13 (Alaoglu). *Para todo espacio normado X, B_{X^*} es débil* compacto.*

DEMOSTRACIÓN. Por el Teorema de Tychonoff, el espacio $[-1,1]^{B_X}$ de las funciones $f : B_X \longrightarrow [-1,1]$ con la topología producto (la topología de la convergencia puntual) es un espacio compacto. Dado un $x^* \in B_{X^*}$, podemos identificar de forma natural x^* con un punto de $[-1,1]^{B_X}$, de forma que podemos identificar B_{X^*} con un subconjunto de $[-1,1]^{B_X}$. Puesto que la topología débil* es la topología de la convergencia puntual, esta identificación nos permite ver a (B_{X^*}, τ_{w^*}) como un subconjunto de $[-1,1]^{B_X}$ con la topología producto, de manera que para probar el teorema simplemente tenemos que probar que B_{X^*} es un cerrado de $[-1,1]^{B_X}$.

Sea $(x_i^*)_{i\in I} \subset B_{X^*}$ una red convergente puntualmente (débil*) a $f \in [-1,1]^{B_X}$. Veamos que f es "lineal" sobre B_X: Para todos $\alpha, \beta \in \mathbb{K}$, $x, y \in X$ tales que $\alpha x + \beta y \in B_X$, se tiene

$$f(\alpha x + \beta y) = \lim_i x_i^*(\alpha x + \beta y) =$$

$$= \lim_i \alpha x_i^*(x) + \lim_i \beta x_i^*(y) = \alpha f(x) + \beta f(y),$$

Además f está acotada en B_X y $\|f\| \leq 1$ puesto que $f(B_X) \subset [-1,1]$, por lo que $f \in B_{X^*}$. □

TEOREMA 6.14 (Goldstine). *Sea X un espacio normado. Entonces B_X es débil* densa en $B_{X^{**}}$ y por lo tanto X es débil*-denso en X^{**}.*

DEMOSTRACIÓN. Sea $x^{**} \in X^{**}$ y supongamos que $x^{**} \notin \overline{B_X}^{w^*}$. Aplicamos el Teorema de Hahn-Banach. Para ello observamos que

1. (X^{**}, w^*) es un espacio localmente convexo y $(X^{**}, w^*)^* = X^*$
2. $\overline{B_X}^{w^*} \subset (X^{**}, w^*)$ es un conjunto débil* cerrado y convexo.
3. x^{**} es débil* cerrado y convexo.

Por lo tanto podemos aplicar la versión geométrica del Teorema de Hahn-Banach para obtener un elemento $x^* \in X^*$ tal que

$$\sup_{y^{**} \in \overline{B_X}^{w^*}} x^*(y^{**}) < x^*(x^{**})$$

además podemos suponer sin perdida de generalidad que $\|x^*\| = 1$.

En ese caso

$$\sup_{y^{**} \in \overline{B_X}^{w^*}} x^*(y^{**}) \geq \|x^*\| = 1$$

por lo que

$$x^*(x^{**}) > 1,$$

luego $\|x^{**}\| > 1$ y por lo tanto

$$B_{X^{**}} \subset \overline{B_X}^{w^*}.$$

La última afirmación es ahora fácil. □

Corolario 6.15. *X es reflexivo si y sólo si B_X es débilmente compacta.*

Demostración. Supongamos que X es reflexivo. Entonces $X = X^{**}$ y $B_X = B_{X^{**}}$, y esta es débil* compacta. Puesto que la topología débil y la débil* coinciden en los reflexivos, tenemos que B_X es débil compacta.

Recíprocamente, si B_X es débil compacta, entonces es débil* cerrada en X^{**}. Puesto que la clausura débil* de B_X en $B_{X^{**}}$ es $B_{X^{**}}$, tenemos que $B_X = B_{X^{**}}$ y por tanto $X = X^{**}$. □

Si bien la topología débil* sobre X^* nunca es metrizable (si X es infinito-dimensional), sí ocurre a menudo que el conjunto compacto (B_{X^*}, w^*) es metrizable, y a menudo se puede sacar partido de ello. Veamoslo

Teorema 6.16. *Un espacio de Banach X es separable si y sólo si (B_{X^*}, w^*) es metrizable.*

Demostración. Sea $(x_n) \subset S_X$ una sucesión densa en S_X, y definamos una función $d : B_{X^*} \times B_{X^*} \longrightarrow [0, \infty)$ mediante

$$d(x^*, y^*) = \sum_{n=1}^{\infty} \frac{|x^*(x_n) - y^*(x_n)|}{2^n}.$$

Claramente la serie es siempre convergente y por tanto d está bien definida. Es fácil ver que d es una *distancia* en B_{X^*}. Veamos entonces que la aplicación identidad

$$Id : (B_{X^*}, w^*) \longrightarrow (B_{X^*}, d)$$

es un homeomorfismo, lo que automáticamente implicará que (B_{X^*}, w^*) es metrizable.

Sea $x_0^* \in B_{X^*}$ y sea

$$U = \{x^* \in B_{X^*} \text{ tales que } d(x^*, x_0^*) < \epsilon\}$$

un entorno abierto de $x_0^* \in (B_{X^*}, d)$. Sea $n_0 \in \mathbb{N}$ un natural tal que

$$\frac{1}{2^{n_0}} = \sum_{i=1}^{\infty} \frac{1}{2^{-n_0-i}} < \frac{\epsilon}{4}.$$

Sea entonces $W(x_0; x_1, \ldots, x_{n_0}; \epsilon)$, un entorno de $x_0^* \in (B_{X^*}, w^*)$. Para todo $x^* \in W(x_0; x_1, \ldots, x_{n_0}; \epsilon)$ se tiene que

$$d(x^*, x_0^*) = \sum_{n=1}^{\infty} \frac{|x_0^*(x_n) - x^*(x_n)|}{2^n} =$$

$$= \sum_{n=1}^{n_0} \frac{|x_0^*(x_n) - x^*(x_n)|}{2^n} + \sum_{n=n_0+1}^{\infty} \frac{|x_0^*(x_n) - x^*(x_n)|}{2^n} \leq$$

$$\leq \frac{\epsilon}{2} + \sum_{n=1}^{\infty} \frac{1}{2^{-n_0-n+1}} < \epsilon,$$

y por tanto $W(x_0; x_1, \ldots, x_{n_0}; \epsilon) \subset U$ y se tiene que

$$Id : (B_{X^*}, w^*) \longrightarrow (B_{X^*}, d)$$

es continua. Al ser (B_{X^*}, w^*) compacto, (B_{X^*}, d) Hausdorff e Id biyectiva se tiene automáticamente que $Id^{-1} : (B_{X^*}, d) \longrightarrow (B_{X^*}, w^*)$ es continua, es decir Id es un homeomorfismo y por tanto (B_{X^*}, w^*) es metrizable.

Recíprocamente, si (B_{X^*}, w^*) es metrizable entonces verifica el I Axioma de Numerabilidad y por tanto sabemos que existe una sucesión $(U_n)_n$ de entornos del origen de (B_{X^*}, w^*) tales que $\cap_{n=1}^{\infty} U_n = \{0\}$. Por la definición de la topología débil*, para cada $n \in \mathbb{N}$ existe ϵ_n y un conjunto finito $F_n = \{x_{n_1}, \ldots, x_{n_m}\} \subset X$ tal que

$$W(0; x_{n_1}, \ldots, x_{n_m}; \epsilon_n) \subset U_n$$

Sea $F = \cup_{n=1}^{\infty} F_n$. Así definido F es contable. Además $(F^{\perp})^{\perp}$ es la envoltura lineal y cerrada de F y por tanto es separable. Pero si $x \in F^{\perp}$ quiere decir que para todo $n \geq 1$ y para todo $x \in F_n$

$$\left| \frac{x^*}{\|x^*\|}(x) \right| < \epsilon_n.$$

Por tanto $\frac{x^*}{\|x^*\|} \in U_n$ para todo $n \geq 1$, de donde $x^* = 0$. Es decir $F^{\perp} = \{0\}$ y por tanto $(F^{\perp})^{\perp} = X$, por lo que X es separable. □

Corolario 6.17. *[Banach 1932] Sea X un espacio de Banach separable. Entonces toda sucesión acotada de X^* tiene una subsucesión débil* convergente.*

Demostración. Si X es separable, (B_{X^*}, w^*) es metrizable, por lo que la compacidad se caracteriza por sucesiones. □

Veamos dos aplicaciones de este resultado. La primera es al análisis de Fourier.

TEOREMA 6.18. *Sea $(\alpha_n)_{n\in\mathbb{Z}}$ una sucesión indexada por $\mathbb{Z}$. Para cada $m \in \mathbb{N}$ definamos la función $s_m : [-\pi,\pi] \longrightarrow \mathbb{C}$ como*

$$s_m(t) = \sum_{n=-m}^{m} \alpha_n e^{\imath nt}$$

y sea

$$a_m = \frac{\sum_{j=0}^{m-1} s_j}{m}.$$

Sea $1 < q \leq \infty$ tal que la sucesión $(a_m) \subset L_q[-\pi,\pi]$ es acotada. Entonces existe $y \in L_q[-\pi,\pi]$ tal que para todo $n \in \mathbb{Z}$

$$\alpha_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} y(t)e^{-\imath nt}dt,$$

es decir

$$\sum_{n=-\infty}^{\infty} \alpha_n e^{\imath nt}$$

es la serie de Fourier de y.

DEMOSTRACIÓN. Recordemos la notación: Para todo $x \in L_1[-\pi,\pi]$ y $n \in \mathbb{Z}$, definimos el n-simo coeficiente de Fourier de x como

$$\hat{x}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)e^{-\imath nt}dt.$$

Resulta claro de las definiciones que para todo $n \in \mathbb{Z}$ y para todo $m > |n|$ se tiene

$$\hat{a}_m(n) = \frac{m-|n|}{m}\alpha_n.$$

Sea q como en la hipótesis y sea $1 \leq p < \infty$ de manera que $\frac{1}{p}+\frac{1}{q}=1$. Puesto que por hipótesis $(a_m)_m \in L_q = (L_p)^*$ es una sucesión acotada, podemos aplicar el Corolario 6.17 y obtenemos una subsucesión (a_{m_j}) débil* convergente a $y \in L_q$. Para todo $n \in \mathbb{Z}$ sea $x_n(t) = \frac{e^{-\imath nt}}{2\pi}$. Entonces para todos m, n

$$a_m(x_n) = \hat{a}_m(n), \quad \text{y} \quad y(x_n) = \hat{y}(n)$$

y por tanto

$$\hat{y}(n) = \lim_j \hat{a}_{m_j}(n) = \lim_j \frac{m_j - |n|}{m_j}\alpha_n = \alpha_n.$$

□

Otra aplicación que puede ser interesante estudiar es el llamado *Principio de seleccion de Helly*. Se puede ver una exposición sencilla por ejemplo en ([**31**, p. 273 y ss]).

Finalmente, de manera opcional se puede probar el Teorema de Eberlein, que nos dice que, aunque la topología débil no se pueda caracterizar por sucesiones, la compacidad débil *sí* se puede caracterizar por sucesiones.

TEOREMA 6.19 (Eberlein). *Sea X un espacio de Banach y $A \subset X$. Entonces A es (relativamente) débilmente compacto si y sólo si A es secuencialmente (relativamente) compacto.*

DEMOSTRACIÓN. En [**15**, p. 18] se puede ver una demostración que requiere una cierta dosis de topología. Una demostración totalmente elemental usando tan sólo la convergencia débil de sucesiones de un enunciado algo más sencillo "X es reflexivo si y sólo si toda sucesión acotada tiene una subsucesión débilmente convergente" se puede ver en [**31**, p. 288]. □

Prácticas sugeridas

EJERCICIO 6.1. *Probar que una sucesión $(f_n) \subset C[0,1]$ converge débilmente a f si y sólo si (f_n) está acotada y para todo $t \in [0,1]$ $f_n(t)$ tiende a $f(t)$.*

EJERCICIO 6.2. *Dado un espacio de Banach separable X, encontrar $T \in \mathcal{L}(\ell_2; X)$ tal que $T(\ell_2)$ sea denso en X.* **Sugerencia:** *Sea $(y_n) \subset X$ una sucesión densa y sea $T : \ell_2 \longrightarrow X$ dado por*

$$T(a) = \sum_n \frac{a_n y_n}{2^n}$$

EJERCICIO 6.3. *X es reflexivo si y sólo si X^* es reflexivo*

EJERCICIO 6.4. *Si X es reflexivo e Y es un subespacio vectorial cerrado de X entonces Y es reflexivo.*

Capítulo 7

Operadores compactos

En el estudio de las aplicaciones lineales entre espacios de Banach (o simplemente normados) pronto se ve que muchos problemas que en el caso de espacios finito-dimensionales tienen solución resultan muy difíciles, o imposibles, de solucionar. Con frecuencia el problema es esencialmente la no compacidad de la bola unidad. Buena parte de las técnicas del Análisis Funcional van dirigidas a solucionar este problema. Una de las formas de abordarlo es transformar la bola unidad mediante un operador en un conjunto *relativamente compacto*. A los operadores que cumplen esa condición les llamamos *operadores compactos*.

Es claro que los operadores de *rango finito* son compactos. Además a menudo los operadores compactos son precisamente la clausura en el espacio normado de los operadores de los operadores de rango finito.

Vemos también la *propiedad de ideal* de los operadores compactos, así como el fundamental *Teorema de Schauder*, cuya demostración constituye una bonita aplicación del Teorema de Ascoli-Arzela.

Para terminar el capítulo, estudiamos la relación de los operadores compactos con los operadores *completamente continuos*.

Hemos basado la preparación de este capítulo en [**31**], [**24**] y [**13**].

Operadores compactos

Empezamos definiendo los operadores compactos.

Definición 7.1. *Sean X e Y espacios normados. Un operador $T : X \longrightarrow Y$ se dice compacto si $\overline{T(B_X)}$ es un conjunto compacto. Denotaremos por $\mathcal{K}(X;Y)$ al espacios de los operadores compactos de X en Y con la norma heredada de $\mathcal{L}(X;Y)$. Si $X = Y$ escribiremos $\mathcal{K}(X)$ en lugar de $\mathcal{K}(X;X)$.*

Definición 7.2. *Sean X e Y espacios normados. Un operador $T : X \longrightarrow Y$ se dice de rango finito si* $\dim(Im(T)) < \infty$. *Denotaremos por $\mathcal{F}(X;Y)$ al espacios de los operadores de rango finito de X en Y con la norma heredada de $\mathcal{L}(X;Y)$. Si $X = Y$ escribiremos $\mathcal{F}(X)$ en lugar de $\mathcal{F}(X;X)$.*

Nótese primeramente que, como ya vimos, se sigue del Lema de Riesz que la bola unidad de los espacios de dimensión infinita no es compacta, por lo que si $\dim X = \infty$, el operador identidad $Id : X \longrightarrow X$ nunca es compacto. También el Teorema de la Aplicación Abierta nos garantiza que los operadores sobreyectivos (sobre espacios de dimensión infinita) no son compactos.

Observación 7.3. *Si $T \in S_{X^*}$ es tal que T no alcanza su norma en B_X, entonces $T(B_X) = (-1, 1)$, y tenemos que $T \in \mathcal{L}(X; \mathbb{R})$ pero $\overline{T(B_X)}$ no es compacto. Por tanto no se puede sustituir $T(B_X)$ por $\overline{T(B_X)}$ en la definición de operador compacto.*

Si se quiere un ejemplo explícito de esta situación, considérese por ejemplo la forma $T : c_0 \longrightarrow \mathbb{K}$ definida por $T(e_n) = \frac{1}{2^n}$. Entonces

$$\|T\| = 1$$

pero la norma no se alcanza en la bola de c_0 (se alcanza en la bola de ℓ_∞, por ejemplo en $\mathbb{1}$).

Proposición 7.4. *Sean X, Y espacios normados, $T : X \longrightarrow Y$ un operador. Entonces*

(i) *T es compacto si y sólo si dada cualquier sucesión acotada $(x_n)_n \subset X$, la sucesión $T(x_n)_n$ tiene una subsucesión convergente.*
(ii) *Si T es compacto, entonces $T(B_X)$ es un conjunto precompacto. Recíprocamente, si Y es Banach y $T(B_X)$ es precompacto entonces T es compacto.*
(iii) *Si T es continuo y de rango finito entonces T es compacto y $Im(T)$ es cerrado en Y. Recíprocamente, si X, Y son espacios de Banach, T es compacto e $Im(T)$ es cerrado automáticamente T es de rango finito.*

Demostración. (i) Sea T compacto y $(x_n) \subset X$ una sucesión tal que $\|x_n\| \leq C$ para todo $n \in \mathbb{N}$. Entonces $T(\frac{x_n}{C}) \in T(B_X)$ para todo $n \in \mathbb{N}$. Puesto que $\overline{T(B_X)}$ es compacto, sabemos que existe una subsucesión de $T(\frac{x_n}{C})$ que converge en $\overline{T(B_X)} \subset Y$. Por tanto una subsucesión de $T(x_n)$ converge en Y.

Recíprocamente, para ver que $\overline{T(B_X)}$ es compacto usamos la caracterización por sucesiones (puesto que estamos en un espacio metrizable). Sea $(y_n) \subset \overline{T(B_X)}$. Entonces existe $(x_n) \subset B_X$ tal que

$$\|y_n - T(x_n)\| < \frac{1}{n}$$

Por hipótesis existe una subsucesión $(x_{n_j})_j \subset (x_n)_n$ tal que $T(x_{n_j})_j$ converge a $y \in Y$. De ahí se sigue con facilidad que la subsucesión $(y_{n_j})_j \subset (y_n)_n$ también converge a y, y por tanto $y \in \overline{T(B_X)}$. Lo que prueba que $\overline{T(B_X)}$ es compacto.

(ii) Si T es compacto, entonces $\overline{T(B_X)}$ es precompacto, y por tanto $T(B_X)$ también lo es.

Recíprocamente, si Y es Banach y $T(B_X)$ es precompacto, entonces $\overline{T(B_X)}$ es precompacto y completo, y por tanto compacto.

Probemos (iii): Si T es de rango finito entonces $Im(T)$ es cerrado en Y por ser un espacio de dimensión finita. Entonces $\overline{T(B_X)} \subset Im(T)$ y $\overline{T(B_X)}$ es un conjunto cerrado y acotado en un espacio de dimensión finita, luego es compacto.

Recíprocamente, sean X e Y espacios de Banach, $T : X \longrightarrow Y$ compacto tal que $T(X)$ es cerrado. Por ser T compacto, es continuo. Además $T(X)$ es un espacio de Banach y $T : X \longrightarrow T(X)$ es un operador sobre. Por el Teorema de la Aplicación Abierta, $T(U_X)$ es un abierto en $T(X)$. Por tanto existe $\delta > 0$ tal que

$$E := \{y \in T(X) \text{ tales que } \|y\| < \delta\} \subset T(B_X)$$

y puesto que $T(X)$ es cerrado, tenemos que

$$\overline{E} = \{y \in T(X) \text{ tales que } \|y\| \leq \delta\} \subset \overline{T(B_X)} \subset T(X)$$

Puesto que $\overline{T(B_X)}$ es compacto, tenemos que la bola cerrada de centro 0 y radio δ de $T(X)$ es compacta. Y esto implica que $\dim(T(X)) < \infty$. □

Veamos a continuación que los operadores compactos forman un ideal de operadores cerrado en la norma usual de operadores

Teorema 7.5. *Sean X, Y espacios normados. Entonces*

1. *Para todos $T, S \in \mathcal{K}(X;Y)$, para todos $\alpha, \beta \in \mathbb{K}$, el operador*
$$\alpha T + \beta S : X \longrightarrow Y$$
es compacto.
2. *(Propiedad de ideal de operadores) Si E, F son espacios normados $R : E \longrightarrow X$, $S : Y \longrightarrow F$ son operadores y $T \in \mathcal{K}(X;Y)$ entonces*
$$S \circ T \circ R : E \longrightarrow F$$
es compacto.

3. *Si Y es un espacio de Banach, $\mathcal{K}(X;Y)$ es un subespacio cerrado de $\mathcal{L}(X;Y)$*

Demostración. 1. Es muy fácil ver que si T es compacto también lo es αT, por ejemplo usando la caracterización por sucesiones. De esta misma forma demostramos que la suma de operadores compactos es compacto: Sean T, S como en la hipótesis, y sea una sucesión $(x_n) \in B_X$. Entonces existe una subsucesión $(x_{n_j})_j \subset (x_n)$ tal que $T(x_{n_j})$ converge. A su vez existe una subsubsucesión $(x_{n_{j_k}})_k \subset (x_{n_j})$ tal que $S(x_{n_{j_k}})$ converge. Entonces

$$(T+S)(x_{n_{j_k}})$$

converge, lo que prueba que $S+T$ es compacto.

2. Se puede demostrar también por sucesiones, pero lo haremos directamente con la definición. Notemos que

$$\overline{S \circ T \circ R(B_E)} = \overline{S \circ T(R(B_E))} \subset \overline{S \circ T(\|R\| B_X)} =$$
$$= \overline{S(T(\|R\| B_X))} \subset \overline{S(\overline{T(\|R\| B_X)})}$$

Puesto que $\overline{T(B_X)}$ es compacto, también lo es $\overline{T(\|R\| B_X)}$ y puesto que S es continua, transforma compactos en compactos. Por tanto $\overline{S \circ T \circ R(B_E)}$ es un cerrado contenido en un compacto, y por tanto es compacto.

3. Sea $(T_n)_n \subset \mathcal{K}(X;Y)$ una sucesión que converge a $T \in \mathcal{L}(X;Y)$ en la norma de operadores. Hemos de comprobar que T es compacto. Por la Proposición 7.4 basta comprobar que $T(B_X)$ es precompacto. Sea $\epsilon > 0$. Sea $n \in \mathbb{N}$ tal que

$$\|T_n - T\| \leq \frac{\epsilon}{3}.$$

Puesto que T_n es precompacto, existen $x_1, \ldots, x_m \in B_X$ tales que

$$T_n(B_X) \subset \cup_{i=1}^m B(T_n(x_i), \frac{\epsilon}{3}).$$

Entonces, dado $x \in B_X$, sea x_j tal que $T_n(x) \in B(T_n(x_j), \frac{\epsilon}{3})$. Así, se tiene que

$$\|T(x) - T(x_j)\| \leq \|(T - T_n)(x)\| + \|T_n(x) - T_n(x_j)\| + \|(T_n - T)(x_j)\| \leq$$
$$\leq \|T - T_n\| \|x\| + \frac{\epsilon}{3} + \|T - T_n\| \|x_j\| \leq 3\frac{\epsilon}{3} = \epsilon$$

y por tanto

$$T(B_X) \subset \cup_{i=1}^m B(T_n(x_i), \epsilon),$$

lo que implica que T es compacto. □

TEOREMA 7.6 (Schauder). *Sean X, Y espacios de Banach, $T \in \mathcal{L}(X;Y)$. Entonces T es compacto si y sólo si T^* es compacto.*

DEMOSTRACIÓN. Sea $T : X \longrightarrow Y$ un operador compacto. Sea $(y_n^*)_n \subset B_{Y^*}$. Para todos $y, z \in Y$

$$|y_n^*(y) - y_n^*(z)| \leq \|y_n^*\| \|y - z\| \leq \|y - z\|$$

Sea $K = \overline{T(B_X)}$. K es un espacio compacto metrizable y por lo anterior el conjunto

$$\{y_n^*; n \in \mathbb{N}\}$$

es un conjunto de funciones equicontinuas uniformemente acotadas. Entonces el Teorema de Ascoli-Arzela nos dice que existe una subsucesión $(y_{n_j}^*)_{j\in\mathbb{N}}$ que converge uniformemente en K. Entonces, para todo $i, j \in \mathbb{N}$,

$$\|T^*(y_{n_i}^*) - T^*(y_{n_j}^*)\| = \sup_{x \in B_X} |T^*(y_{n_i}^* - y_{n_j}^*)(x)| =$$
$$= \sup_{x \in B_X} |(y_{n_i}^* - y_{n_j}^*)T(x)| \leq \sup_{y \in K} |y_{n_i}^*(y) - y_{n_j}^*(y)|$$

Puesto que la sucesión $(y_{n_j})_{j\in\mathbb{N}}$ es uniformemente Cauchy en K, se sigue que $(T^*(y_{n_j}^*))_j$ es una sucesión de Cauchy en X^*. Por tanto, puesto que X^* es Banach, $(T^*(y_{n_j}^*))_j$ debe converger y de ahí se sigue que T^* es compacto.

La otra implicación es esencialmente la demostración de que los operadores compactos forman un ideal inyectivo de operadores, aunque no presentaremos esa noción a nuestros alumnos:

Supongamos que T^* es compacto. Entonces $T^{**} = (T^*)^*$ también es compacto por lo anterior. Puesto que $T^{**}J_X = J_Y T$, la propiedad de ideal de los operadores compactos nos garantiza que $J_Y T$ es compacto. Esto quiere decir $J_Y(T(B_X))$ es precompacto. Puesto que J_Y es una isometría, se sigue que $T(B_X)$ es precompacto, y de aquí se sigue que T es compacto. □

EJEMPLO 7.7. *Ya hemos visto que todo operador de rango finito es compacto, y puesto que los operadores compactos $\mathcal{K}(X;Y)$ forman un subespacio cerrado en el espacio de los operadores, se sigue que el límite (en la norma de operadores) de una sucesión de operadores de rango finito es compacto. De hecho, en muchos casos (por ejemplo si Y es un espacio con base de Schauder), todo operador compacto es el límite de una sucesión de operadores de rango finito. En 1932 Banach preguntó si en un espacio separable todo operador compacto se podría escribir como el límite de una sucesión de operadores de rango finito, es* *decir si* $\mathcal{K}(X) = \overline{\mathcal{F}(X)}$. *Más tarde, en 1953 Grothendieck* [**23**] *definió la Propiedad de Aproximación de la siguiente forma*

X tiene la Propiedad de Aproximación si y sólo si para todo espacio de Banach Y, $\mathcal{K}(X;Y) = \overline{\mathcal{F}(X;Y)}$

Hubo que esperar a 1973 para que Enflo [**19**] *probara que no todo espacio de Banach tiene la Propiedad de Aproximación, y diera simultáneamente una respuesta negativa a la pregunta de Banach.*

Es fácil ver directamente que para los espacios de sucesiones separables sí se tiene que $\mathcal{K}(X) = \overline{\mathcal{F}(X)}$.

PROPOSICIÓN 7.8. *Sea* X *uno de los espacios* c_0, ℓ_p $(1 \leq p < \infty)$. *Entonces* $\mathcal{K}(X) = \overline{\mathcal{F}(X)}$.

DEMOSTRACIÓN. Sea $(e_n)_{n\in\mathbb{N}}$ la base canónica habitual de estos espacios. Dado $n \in \mathbb{N}$, sea $\pi_n : X \longrightarrow X$ el operador que a cada $x = (x_i)_{i\in\mathbb{N}}$ le asocia

$$\pi_n(x) = \sum_{i=1}^{n} x_i e_i = (x_1, \ldots, x_n, 0, 0, \ldots).$$

Es inmediato ver que para cualquiera de los espacios mencionados π_n es continuo y que

$$\lim_n \pi_n(x) = Id(x) = x.$$

Sea ahora $T \in \mathcal{K}(X)$. Claramente $\pi_n \circ T \in \mathcal{F}(X)$ para todo $n \in \mathbb{N}$. Sólo nos falta ver que $\pi_n \circ T$ converge a T en la norma de $\mathcal{L}(X)$. Esto es lo mismo que ver que $(\pi_n - Id)(T(x))$ converge a 0 uniformemente en B_X. Para probar eso es suficiente probar que $\pi_n - Id$ tiende a 0 uniformemente en el compacto $\overline{T(B_X)}$ y esto se sigue del Teorema de Banach-Steinhaus. $\square$

EJEMPLO 7.9. *Sea* $T : \ell_2 \longrightarrow \ell_2$ *el operador dado por*

$$T(x) = \left(\frac{x_n}{2^n}\right)_n.$$

Veamos que T *es compacto, mientras que claramente no tiene rango finito.*

Para ver que es compacto, veamos que $T(B_{\ell_2}) \subset \ell_2$ *es precompacto. Dado* $\epsilon > 0$ *vamos a encontrar un* 2ϵ*-cubrimiento finito de* $T(B_{\ell_2})$. *Sea* $n_0 \in$ *tal que* $2^{-n_0} \leq \epsilon$. *Obsérvese que el conjunto*

$$A = \{(\frac{x_1}{2^1}, \frac{x_2}{2^2}, \ldots, \frac{x_{n_0}}{2^{n_0}}, 0, 0, \ldots); |x_n| \leq 1\}$$

es compacto en ℓ_2 (es acotado y cerrado en $\mathbb{R}^{n_0}$ que a su vez es cerrado en ℓ_2) y por tanto admite un ϵ-cubrimiento centrado en una ϵ-red F. Veamos que F es una $\sqrt{2}\epsilon$-red de $T(B_{\ell_2})$. Sea $x = (x_n) \in B_{\ell_2}$. Entonces $x_n \leq 1$ para todo $n \in \mathbb{N}$, y por tanto existe $f \in F$ tal que

$$\sum_{n=1}^{n_0} |2^{-n}x_n - f|^2 < \epsilon^2$$

Entonces

$$\| \left(\frac{x_n}{2^n}\right)_n - f\|_2^2 = \sum_{n=1}^{n_0} \left|\frac{x_n}{2^n} - f_n\right|^2 + \sum_{n=n_0+1}^{\infty} \left|\frac{x_n}{2^n}\right|^2 < \epsilon^2 + \epsilon^2 = 2\epsilon^2$$

EJEMPLO 7.10. *Sea $A = (a_{i,j})_{i,j\in\mathbb{N}}$ una matriz infinita y sea $1 \leq p \leq \infty$.*

Si $p = 1$, supongamos que $\gamma(j) = \sum_{i=1}^{\infty} |a_{i,j}| < \infty$ para todo $j \in \mathbb{N}$ y que $\gamma(j) \to 0$ cuando $j \to \infty$.

Si $p = \infty$ supongamos que $\delta(i) = \sum_{j=1}^{\infty} |a_{i,j}| < \infty$ para todo $i \in \mathbb{N}$ y que $\delta(i) \to 0$ cuando $i \to \infty$.

Si $1 < p < \infty$ y $\frac{1}{p} + \frac{1}{q} = 1$, supongamos que

$$\beta_p = \left(\sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty} |a_{i,j}|^q\right)^{\frac{p}{q}}\right)^{\frac{1}{p}} < \infty.$$

Entonces el operador

$$T_A : \ell_p \longrightarrow \ell_p$$

dado por

$$T_A(x) = \left(\sum_{j=1}^{\infty} a_{i,j}x_j\right)_i$$

(está bien definido y) es compacto.

Comencemos viendo el caso $p = 1$. Veamos en primer lugar que T_A está bien definido y es continuo. Notemos que si $\gamma(j) \to 0$ en particular se tiene que $\sup_j \gamma(j) = C < \infty$. Entonces

$$\|T_A(x)\|_1 = \left\|\left(\sum_{j=1}^{\infty} a_{i,j}x_j\right)_i\right\|_1 = \sum_{i=1}^{\infty}\left|\sum_{j=1}^{\infty} a_{i,j}x_j\right| \leq \sum_i\sum_j |a_{i,j}x_j| =$$

$$= \sum_j\sum_i |a_{i,j}||x_j| \leq \sum_j |x_j| \sum_i |a_{i,j}| = \sum_j |x_j|\gamma(j) \leq M\|x\|_1$$

Veamos ahora que T es compacto. Sea A_n la matriz doblemente infinita definida que se obtiene al mantener las primeras n columnas de A sin cambios y colocar ceros en las restantes columnas. El operador T_n asociado a la matriz A_n, dado por

$$T_n(x) = \left(\sum_{j=1}^{n} a_{i,j} x_j \right)_i$$

actúa igual que T, salvo que previamente "trunca" x y se queda sólo con sus n primeras coordenadas. Es claro que $T_n = T \circ \pi_n$ donde π_n vienen definidas como en el Ejemplo 7.8. Es claro entonces que para cada $n \in \mathbb{N}$ el operador T_n es continuo y tiene rango finito . Lo único que tenemos que probar es que T_n tiende a T. Pero

$$\|T_n - T\| = \sup_{x \in B_{\ell_1}} \|(T_n - T)(x)\| = \sup_{x \in B_{\ell_1}} \left\| \left(\sum_{j=n+1}^{\infty} a_{i,j} x_j \right)_i \right\|_1 =$$

$$= \sup_{x \in B_{\ell_1}} \sum_{i=1}^{\infty} \left| \sum_{j=n+1}^{\infty} a_{i,j} x_j \right| \leq \sup_{x \in B_{\ell_1}} \sum_{i=1}^{\infty} \sum_{j=n+1}^{\infty} |a_{i,j} x_j| =$$

$$= \sup_{x \in B_{\ell_1}} \sum_{j=n+1}^{\infty} \sum_{i} |a_{i,j}||x_j| \leq \sup_{x \in B_{\ell_1}} \sum_{j=n+1}^{\infty} |x_j| \sum_{i} |a_{i,j}| =$$

$$= \sup_{x \in B_{\ell_1}} \sum_{j=n+1}^{\infty} |x_j| \gamma(j) \leq \sup_{x \in B_{\ell_1}} \|x\| \sup_{j>n} \gamma(j) = \sup_{j>n} \gamma(j) \to 0.$$

Veamos ahora el caso $p = \infty$. Sea A^t la matriz traspuesta de A. Por el caso $p = 1$, A^t define un operador compacto $T_{A^t} : \ell_1 \longrightarrow \ell_1$. Extendiendo los razonamientos hechos para espacios finito dimensionales en el Ejercicio 5.1 es fácil ver que $A = (A^t)^t$ es la matriz asociada al operador traspuesto de T_{A^t}, es decir $T = (T_{A^t})^*$. Entonces el Teorema de Schauder nos garantiza que T es compacto.

Veamos ahora el caso $1 < p < \infty$. Veamos en primer lugar que $T_A : \ell_p \longrightarrow \ell_p$ (está bien definido y) es continuo. Nótese el uso de la desigualdad de Hölder.

$$\|T_A(x)\|_p = \left\| \left(\sum_{j=1}^{\infty} a_{i,j} x_j \right)_i \right\|_p \leq \left\| \left(\sum_{j=1}^{\infty} |a_{i,j} x_j| \right)_i \right\|_p \leq$$

$$\leq \left\| \left(\left(\sum_{j=1}^{\infty} |a_{i,j}|^q \right)^{\frac{1}{q}} \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{\frac{1}{p}} \right)_i \right\|_p =$$

$$= \|x\|_p \left\| \left(\left(\sum_{j=1}^{\infty} |a_{i,j}|^q \right)^{\frac{1}{q}} \right)_i \right\|_p = \|x\|_p \beta_p.$$

Ya hemos visto que T_A es continuo. Veamos que es incluso compacto. Podemos de nuevo considerar las matrices A_n y sus operadores de rango finito asociados $T_n = T \circ \pi_n$. De nuevo sólo necesitamos comprobar que T_n tiende a T en la norma de operadores.

$$\|T_n - T_A\| = \sup_{x \in B_{\ell_p}} \|(T_n - T_A)(x)\|_p = \sup_{x \in B_{\ell_p}} \left\| \left(\sum_{j=n+1}^{\infty} a_{i,j} x_j \right)_i \right\|_p \leq$$

$$\leq \sup_{x \in B_{\ell_p}} \left\| \left(\sum_{j=n+1}^{\infty} |a_{i,j} x_j| \right)_i \right\|_p \leq$$

$$\leq \sup_{x \in B_{\ell_p}} \left\| \left(\left(\sum_{j=n+1}^{\infty} |a_{i,j}|^q \right)^{\frac{1}{q}} \left(\sum_{j=n+1}^{\infty} |x_j|^p \right)^{\frac{1}{p}} \right)_i \right\|_p =$$

$$= \sup_{x \in B_{\ell_p}} \|x\|_p \left\| \left(\left(\sum_{j=n+1}^{\infty} |a_{i,j}|^q \right)^{\frac{1}{q}} \right)_i \right\|_p \to 0.$$

También proponemos como ejercicio demostrar que si una de las sucesiones $\gamma(j)$ y $\delta(i)$ definidas más arriba es acotada y la otra tiende a 0 entonces $T : \ell_p \longrightarrow \ell_p$ es compacto.

Veamos el análogo continuo del ejemplo de arriba, lo que se conoce como operadores integrales de Fredholm.

Ejemplo 7.11. *Sea $k(\cdot,\cdot) : [0,1] \times [0,1] \longrightarrow \mathbb{K}$ una función a la que de ahora en adelante nos referiremos como núcleo de Fredholm. El núcleo k nos permite definir una aplicación entre espacios funcionales (de momento por precisar)*

$$x \mapsto T_k(x)(s) = \int_0^1 k(s,t) x(t) dt$$

Estudiemos las características de este operador en función del núcleo k *y del espacio en que lo definamos.*

Supongamos inicialmente que $k : [0,1] \times [0,1] \longrightarrow \mathbb{K}$ *es continua. Sea* $x : [0,1] \longrightarrow \mathbb{K}$ *una función integrable. Si* $s_n \to s \in [0,1]$ *entonces el Teorema de la Convergencia Dominada nos garantiza que*

$$T_k(x)(s_n) \to T_k(x)(s),$$

es decir, $T(x) : [0,1] \longrightarrow \mathbb{K}$ *es continua.*

Supongamos ahora que $(x_n)_n$ *es una sucesión de funciones integrables uniformemente acotadas en norma 1, es decir, existe* $\alpha > 0$ *tal que* $\|x_n\|_1 \leq \alpha$ *para todo* $n \in \mathbb{N}$. *Veamos que la sucesión* $(T_k(x_n))_n$ *tiene una subsucesión que converge uniformemente (es decir, en* $\|\cdot\|_\infty$*) en* $[0,1]$.

Por ser k *continuo, existe* $\beta > 0$ *tal que* $\|k\|_\infty \leq \beta$. *Entonces es fácil ver que*

$$\|T_k(x_n)\|_\infty \leq \alpha\beta$$

para todo $n \in \mathbb{N}$. *Es decir,* $(T_k(x_n))_n$ *está acotada en norma infinito. Además la sucesión es equicontinua en* $[0,1]$. *Veámoslo.* k *es uniformemente continua en* $[0,1] \times [0,1]$ *(por la compacidad de* $[0,1] \times [0,1]$*). Entonces para todo* $\epsilon > 0$ *existe* $\delta > 0$ *tal que si* $|s-u| < \delta$ *y* $|t-v| < \delta$ *entonces*

$$|k(s,t) - k(u,v)| < \epsilon.$$

Por lo tanto, para todo $n \in \mathbb{N}$ *y para todos* $s, u \in [0,1]$ *con* $|s-u| < \delta$ *se tiene*

$$|T_k(x_n)(s) - T_k(x_n)(u)| = \int_0^1 (k(s,t) - k(u,t))x_n(t)dt \leq$$

$$\leq \int_0^1 |k(s,t) - k(u,t)||x_n(t)|dt < \epsilon\alpha.$$

Ahora, el Teorema de Ascoli-Arzela nos garantiza que la sucesión $(T_k(x_n))_n$ *tiene una subsucesión uniformemente convergente (es decir, convergente en* $\|\cdot\|_\infty$.

Sean entonces X *e* Y *dos cualesquiera de los espacios* $C[0,1]$ *o* $L_p[0,1]$ *(*$1 \leq p \leq \infty$*). Recordemos que para todo* $1 \leq p \leq \infty$ *se tiene que*

$$\|\cdot\|_1 \leq \|\cdot\|_p \leq \|\cdot\|_\infty$$

y

$$C[0,1] \subset L_\infty[0,1] \subset L_p[0,1] \subset L_1[0,1]$$

Por tanto, si $x \in X$ entonces x es integrable y (es decir, $x \in L_1[0,1]$) y si $(x_n) \subset X$ es una sucesión acotada, entonces (x_n) está acotada en $\|\cdot\|_1$.

Además, si $y \in C[0,1]$ entonces $y \in Y$, y si (y_n) es una sucesión uniformemente convergente entonces (y_n) converge en la norma de Y.

Por tanto, de los razonamientos anteriores se sigue que el operador

$$T_k : X \longrightarrow Y$$

(está bien definido y) es compacto.

Veamos que otras situaciones en las que podemos probar con facilidad la compacidad de $T_k : X \longrightarrow Y$. Sea $1 < p \leq \infty$, sea $X = L_p[0,1]$, sea $Y = L_q[0,1]$ con $\frac{1}{p} + \frac{1}{q} = 1$ y supongamos que $k \in L_q([0,1] \times [0,1])$. Entonces, para todo $x \in X$ y $s \in [0,1]$ usando la desigualdad de Hölder análogamente a como lo hicimos en el Ejemplo 7.10 tenemos

$$|T_k(x)(s)| \leq \int_0^1 |k(s,t)||x(t)|dt \leq \|x\|_p \left(\int_0^1 |k(s,t)|^q dt \right)^{\frac{1}{q}}$$

y por lo tanto

$$\|T_k(x)\|_q = \left(\int_0^1 |T_k(x)(s)|^q ds \right)^{\frac{1}{q}} \leq$$

$$\leq \|x\|_p \left(\int_0^1 \int_0^1 |k(s,t)|^q dt ds \right)^{\frac{1}{q}} = \|x\|_p \|k\|_q$$

de donde se sigue la continuidad de T_k.

Para ver la compacidad de T_k, puesto que $C([0,1] \times [0,1])$ es denso en $L_q([0,1] \times [0,1])$ (esto se prueba análogamente a la densidad de $C[0,1]$ en $L_q[0,1]$), sea $(k_n) \subset C([0,1] \times [0,1])$ una sucesión tal que $\|k_n - k\|_q \to 0$ y sea $T_n : L_p[0,1] \longrightarrow L_q[0,1]$ el operador de Fredholm de núcleo k_n. Entonces

$$\|T_k - T_n\|_{\mathcal{L}(L_p;L_q)} = \sup_{x \in B_{L_p[0,1]}} \|(T_k - T_n)(x)\|_q =$$

$$= \sup_{x \in B_{L_p[0,1]}} \left(\int_0^1 |(T_k - T_n)(x)(s)|^q ds \right)^{\frac{1}{q}} \leq$$

$$\leq \sup_{x \in B_{L_p[0,1]}} \|x\|_p \left(\int_0^1 \int_0^1 |k(s,t) - k_n(s,t)|^q dt ds \right)^{\frac{1}{q}} = \|k - k_n\|_q \to 0.$$

A menudo el interés de los operadores compactos radica en su capacidad de "mejorar la convergencia". Para estudiar esto necesitamos previamente una definición.

DEFINICIÓN 7.12. *Sean X, Y espacios de Banach. Un operador $T : X \longrightarrow Y$ se dice* completamente continuo *si transforma sucesiones débilmente convergentes en sucesiones convergentes en norma, es decir, si para todo sucesión $(x_n)_n \subset X$ tal que x_n tiende débilmente a $x \in X$ se tiene que $T(x_n)$ converge en norma a $T(x)$.*

TEOREMA 7.13. *Sean X, Y espacios de Banach. Entonces si T es compacto T es completamente continuo. Recíprocamente, si X es reflexivo y $T : X \longrightarrow Y$ es completamente continuo entonces T es compacto (de hecho basta pedir $X \not\supset \ell_1$, pero esto se sigue de un resultado muy profundo de Rosenthal).*

DEMOSTRACIÓN. Supongamos que T es compacto y x_n tiende débilmente a x. $(x_n) \subset X$ es débilmente acotada, y por tanto acotada por el Teorema 4.9. Si $T(x_n) \not\to T(x)$ podemos suponer, pasando a una subsucesión, que existe $\epsilon > 0$ tal que

$$\|T(x_n) - T(x)\| \geq \epsilon \tag{10}$$

para todo $n \in \mathbb{N}$. Por ser T compacto y $(x_n)_n$ acotada, existe una subsucesión (x_{n_j}) tal que $T(x_{n_j})$ converge a $y \in Y$. Entonces, de (10) se sigue que

$$\|y - T(x)\| \geq \epsilon$$

de forma que $y \neq T(x)$. Sin embargo, para todo $y^* \in Y^*$ se tiene que $y^* \circ T \in X^*$ y por tanto de la hipótesis sobre (x_n) se sigue que

$$(y^* \circ T)(x) = \lim_j (y^* \circ T)(x_{n_j}) = y^* \left(\lim_j T(x_{n_j}) \right) = y^*(y)$$

y por tanto $x = y$ y hemos alcanzado una contradicción.

Para probar la otra implicación necesitamos el Teorema de Eberlein, que no es seguro que incluyamos en el programa. Si lo hemos dado, la demostración es sencilla. Supongamos que X es reflexivo y $T : X \longrightarrow Y$ es completamente continuo. Si $(x_n) \subset B_X$, por el Teorema de Eberlein (x_n) admite una subsucesión (x_{n_k}) débilmente convergente a un cierto $x \in B_X$ (por el Teorema de Mazur, las clausuras en norma y débil de un convexo coinciden). Al ser T completamente continuo, la sucesión $(T(x_{n_k}))_k$ converge en norma, lo que termina la demostración. □

Prácticas sugeridas

EJERCICIO 7.1. *Sea X un espacio de Banach y sea $T \in \mathcal{K}(X) \setminus F(X)$. Demostrar que $0 \in \overline{T(S_X)}$.* **Sugerencia:** *Utilizar el Teorema de la Aplicación Abierta.*

EJERCICIO 7.2. *Sea $C^{(1)}[0,1]$ el espacio de las funciones con derivada primera continua con la norma $\|f\| = \|f\|_\infty + \|f'\|_\infty$. Demostrar que la inclusión formal*

$$i : C^{(1)}[0,1] \longrightarrow C[0,1]$$

es un operador compacto. **Sugerencia:** *Utilizar el Teorema de Ascoli-Arzela.*

EJERCICIO 7.3. *Sea $1 < p < \infty$ y sea $T : c_0 \longrightarrow \ell_p$ un operador continuo. Demostrar que T es compacto.*

Capítulo 8

Teoría espectral de operadores compactos

Comenzamos en este capítulo el estudio de uno de los grandes temas del Análisis Funcional, la Teoría Espectral. Un enfoque posible es comenzar estudiando álgebras de Banach, demostrar en ese marco general los resultados que se van a necesitar y posteriormente observar que los operadores de un espacio en sí mismo es un álgebra de Banach y obtener los resultados de la Teoría Espectral prácticamente como corolario. Sin embargo, pensamos que ese enfoque es algo arduo, puesto que le exige al alumno el estudio abstracto inicial de las álgebras de Banach sin haber trabajado previamente en un modelo intuitivo en el que apoyarse. Preferimos empezar enunciando y demostrando los resultados para el caso de operadores en este capítulo y dejar para más adelante (y no en esta asignatura) la generalización de estos resultados al contexto de álgebras de Banach.

Empezamos estudiando operadores *inversibles* y perturbaciones inversibles de la identidad, para a continuación definir el *espectro* de un operador y sus autovalores. Probamos que el espectro es compacto y probamos el *Teorema de Gelfand*, utilizando variable compleja. Tras ello se puede probar la *Fórmula del Radio Espectral*, aunque no la necesitamos. A continuación desarrollamos la *Teoría Espectral de Operadores Compactos en Espacios de Banach*, conocida como *Teoría de Riesz-Schauder*. Como corolario de esta teoría se obtiene la *Alternativa de Fredholm*, de gran utilidad en las aplicaciones.

Hemos seguido principalmente **[31]** y **[24]** en la preparación de la primera parte del capítulo y **[33]** y **[31]** en la presentación de la Teoría de Riesz-Schauder.

Teoría espectral de operadores compactos

Empezamos definiendo operadores inversibles, y estudiando algunas de sus propiedades básicas.

DEFINICIÓN 8.1. *Dados dos espacios de Banach X, Y, un operador $T : X \longrightarrow Y$ es inversible si es un isomorfismo biyectivo.*

OBSERVACIÓN 8.2. *$T \in \mathcal{L}(X;Y)$ es inversible si y sólo si existe $T^{-1} \in \mathcal{L}(Y;X)$ tal que $T^{-1}T = I_X$ y $TT^{-1} = I_Y$*

Se sigue del Teorema de la Aplicación Abierta que T es inversible si y sólo si T es biyectivo.

Es fácil ver que si $T \in \mathcal{L}(X;Y)$ y $S \in \mathcal{L}(Y;Z)$ son inversibles entonces ST es inversible y

$$(ST)^{-1} = T^{-1}S^{-1}$$

De forma análoga si $T, S \in \mathcal{L}(X)$, es fácil ver que S y T son inversibles si y sólo si ST y TS lo son.

PROPOSICIÓN 8.3. *$T : X \longrightarrow Y$ es inversible si y sólo si $T^* : Y^* \longrightarrow X^*$ es inversible, y en ese caso $(T^*)^{-1} = (T^{-1})^*$.*

DEMOSTRACIÓN. Si T es inversible tenemos que

$$T^*(T^{-1})^* = (T^{-1}T)^* = (I_X)^* = I_{X^*}$$

y

$$(T^{-1})^*T^* = (TT^{-1})^* = (I_Y)^* = I_{Y^*}.$$

Recíprocamente, si T^* es inversible, también lo es T^{**}. Si definimos $S = (T^{**})^{-1}_{|Y}$ es fácil ver que S es un operador inyectivo de Y en X (para ver que S toma valores en Y, notemos que para todo $y \in Y$ existe $x \in X$ tal que $T(x) = y$. Por lo tanto $T^{**}(x) = y$ y $S(y) = x \in X$). Además claramente $ST = I_X$ y $TS = I_Y$, de forma que T es inversible. $\square$

LEMA 8.4. *Sea X un espacio de Banach y $T \in \mathcal{L}(X)$ un operador tal que*

$$\|T\| < 1.$$

Entonces el operador $(I - T) \in \mathcal{L}(X)$ es inversible y además

$$(I - T)^{-1} = \sum_{n=0}^{\infty} T^n,$$

donde la convergencia es en la norma de operadores, y

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

DEMOSTRACIÓN. Nótese en primer lugar que

$$\sum_{n=0}^{\infty} \|T^n\| \leq \sum_{n=0}^{\infty} \|T\|^n = \frac{1}{1 - \|T\|}.$$

Es decir, la serie $\sum_{n=0}^{\infty} T^n$ es absolutamente convergente en el espacio de Banach $\mathcal{L}(X)$ y por tanto la serie $\sum_{n=0}^{\infty} T^n$ converge en $\mathcal{L}(X)$. Por tanto

$$(I-T)\sum_{n=0}^{\infty} T^n = (I-T)+(T-T^2)+(T^2-T^3)+\cdots = I$$

y

$$\left(\sum_{n=0}^{\infty} T^n\right)(I-T) = (I-T)+(T-T^2)+(T^2-T^3)+\cdots = I$$

□

Lema 8.5. *Sea X un espacio de Banach y $T \in \mathcal{L}(X)$ un operador tal que existe $p \in \mathbb{N}$ de manera que*

$$\|T^p\| < 1.$$

Entonces el operador $(I-T) \in \mathcal{L}(X)$ es inversible y además

$$(I-T)^{-1} = \sum_{n=0}^{\infty} T^n,$$

donde la convergencia es en la norma de operadores, y

$$\|(I-T)^{-1}\| \leq \frac{1+\|T\|+\cdots+\|T^{p-1}\|}{1-\|T^p\|}.$$

Demostración. Puesto que

$$\|T^{pn+j}\| \leq \|T^p\|^n \|T^j\|$$

y que

$$\|T^p\| < 1$$

se tiene que

$$\sum_{n=0}^{\infty} \|T^n\| = \sum_{n=0}^{\infty} \|T^{pn}\| + \sum_{n=0}^{\infty} \|T^{pn+1}\| + \cdots + \sum_{n=0}^{\infty} \|T^{pn+p-1}\| \leq$$

$$\leq \sum_{n=0}^{\infty} \|T^p\|^n (1+\|T\|+\cdots\|T^{p-1}\|) = \frac{1+\|T\|+\cdots+\|T^{p-1}\|}{1-\|T^p\|}$$

Ahora la demostración sigue muy similar a la del lema anterior. □

Lema 8.6. *Sea X un espacio de Banach y $S, T \in \mathcal{L}(X)$. Si T es inversible y*

$$\|T^{-1}(S-T)\| < 1$$

entonces S es inversible y

$$S^{-1} = T^{-1}\sum_{n=0}^{\infty}\left((S-T)T^{-1}\right)^n,$$

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

y

$$\|S^{-1}-T^{-1}\| \leq \frac{\|T^{-1}\|^2\|S-T\|}{1-\|(S-T)T^{-1}\|}$$

Demostración. Si $\|(T-S)T^{-1}\| < 1$, el Lema 8.4 nos dice que

$$I-(T-S)T^{-1} = I - I + ST^{-1} = ST^{-1}$$

es inversible y

$$(ST^{-1})^{-1} = \sum_{n=0}^{\infty}\left((T-S)T^{-1}\right)^n.$$

Como T es inversible, se sigue que

$$(ST^{-1})T = S$$

es inversible y

$$S^{-1} = T^{-1}(ST^{-1})^{-1} = T^{-1}\sum_{n=0}^{\infty}(T-S)T^{-1})^n.$$

Por hipótesis, $\|(T-S)T^{-1}\| < 1$ y por tanto

$$\|S^{-1} \leq \|T^{-1}\|\sum_{n=0}^{\infty}\|(T-S)T^{-1}\|^n = \frac{\|T^{-1}\|}{1-\|(T-S)T^{-1}\|}.$$

Además, como

$$(S^{-1}-T^{-1}) = S^{-1}TT^{-1} - S^{-1}ST^{-1} = S^{-1}(T-S)T^{-1}$$

tenemos que

$$\|S^{-1}-T^{-1}\| \leq \|S^{-1}\|\|(T-S)T^{-1}\| \leq \frac{\|T^{-1}\|^2\|S-T\|}{1-\|(S-T)T^{-1}\|}.$$

□

Corolario 8.7. *Sea X un espacio de Banach. Entonces el conjunto $\mathcal{C}$ de operadores inversibles en X es un conjunto abierto de $\mathcal{L}(X)$ y la aplicación $T \mapsto T^{-1}$ es un homeomorfismo de $\mathcal{C}$ en sí mismo.*

Demostración. Si T es inversible, consideremos la bola de centro T y radio $\frac{1}{\|T^{-1}\|}$. Entonces, para todo S en esa bola, se tiene que

$$\|S-T\| < \frac{1}{\|T^{-1}\|}$$

y por tanto

$$\|T^{-1}(S-T)\| \leq \|T^{-1}\|\|S-T\| < 1$$

y el lema anterior nos dice que S es inversible. Por tanto $\mathcal{C}$ es abierto.

Sea ahora $(T_n)_n \subset \mathcal{C}$ una sucesión de operadores inversibles tales que $T_n \to T \in \mathcal{C}$. Entonces a partir de cierto n_0 se tiene que

$$\|T_n - T\| < \frac{1}{\|T^{-1}\|}$$

y de nuevo por el lema anterior tenemos que

$$\|T_n^{-1} - T^{-1}\| \leq \frac{\|T^{-1}\|^2\|T_n - T\|}{1 - \|(T_n - T)T^{-1}\|}$$

lo que implica que $\|T_n^{-1} - T^{-1}\| \to 0$, es decir, la inversión es continua, y esto es todo lo que hace falta probar. □

La idea básica de la teoría espectral es estudiar los valores de $k \in \mathbb{K}$ para los que $T - kI$ es, o no, inversible. Recordemos que esa es la pregunta básica a la que nos lleva la teoría de diagonalización.

Tenemos entonces la siguiente definición:

Definición 8.8. *Sea $T : X \longrightarrow X$. El espectro de T es el conjunto*

$$\sigma(T) = \{k \in \mathbb{K} \text{ tales que } T - kI \text{ no es inversible } \}.$$

El conjunto $\rho(T) := \mathbb{K} \setminus \sigma(T)$ se denomina resolvente de T.

Y la pregunta será cómo calcular los *valores espectrales* $k \in \sigma(T)$.

Veremos que nos interesará destacar dos conjuntos dentro del espectro.

Definición 8.9. *Sea $T : X \longrightarrow X$. El conjunto de autovalores de T es el conjunto*

$$\sigma_e(T) = \{k \in \mathbb{K} \text{ tales que } T - kI \text{ no es inyectivo } \}.$$

El conjunto de autovalores aproximados de T es el conjunto

$$\sigma_a(T) = \{k \in \mathbb{K} \text{ tales que } T - kI \text{ no está acotado inferiormente } \}.$$

Notemos que $k \in \sigma_e(T)$ si y sólo si existe $0 \neq x \in X$ tal que

$$T(x) = kx$$

En ese caso k es un *autovalor* y x es un *autovector* asociado a k. El subespacio

$$\ker(T - kI)$$

es el *autoespacio* asociado a k.

Notemos también que $k \in \sigma_a(T)$ si y sólo si existe una sucesión $(x_n)_n \subset S_X$ tal que

$$\|T(x_n) - kx_n\| \to 0$$

En ese caso k es un *autovalor aproximado* de T. Si $k \in \sigma_e(T)$ es un autovalor y x es uno de sus autovectores haciendo $x_n = \frac{x}{\|x\|}$ para todo n se ve que k es un autovalor aproximado. Por tanto se tiene

$$\sigma_e(T) \subset \sigma_a(T) \subset \sigma(T)$$

Veamos que de hecho para los operadores de rango finito los tres conjuntos son el mismo.

TEOREMA 8.10. *Sea X un espacio normado y $T \in \mathcal{F}(X)$. Entonces*

$$\sigma_e(T) = \sigma_a(T) = \sigma(T)$$

DEMOSTRACIÓN. Puesto que $\sigma_e(T) \subset \sigma_a(T) \subset \sigma(T)$, basta ver que $\sigma(T) \subset \sigma_e(T)$, equivalentemente que $(\sigma_e(T))^c \subset (\sigma(T))^c$. Sea entonces $k \notin \sigma_e(T)$, de forma que $T - kI$ es inyectivo. Veamos que entonces $T - kI$ es inversible, es decir $k \notin \sigma(T)$.

Consideramos primero el caso en que $\dim X = n < \infty$. En ese caso el resultado es elemental y se estudia en Álgebra Lineal: Puesto que

$$\dim \ker(T - kI) + \dim Im(T - kI) = n,$$

si $T-kI$ es inyectivo entonces $\ker(T-kI) = \{0\}$ y por tanto $\dim Im(T-kI) = n$, es decir $T - kI$ es sobreyectiva y por tanto inversible.

Consideramos ahora el caso en que X tiene dimensión infinita. Entonces $k \neq 0$, ya que si T tiene rango finito no puede ser inyectivo (por ejemplo, si $\{x_n; n \in \mathbb{N}\}$ es un sistema linealmente independiente, entonces $\{T(x_n); n \in \mathbb{N}\}$ sería un sistema infinito linealmente independiente contenido en un espacio de dimensión finita).

Sea

$$S = (T - kI)_{|_{Im(T)}} : Im(T) \longrightarrow Im(T)$$

(para ver que S toma efectivamente valores en $Im(T)$, nótese que

$$(T - kI)(T(x)) = T(T(x)) - kI(T(x)) = T(T - kI)(x) \in Im(T).)$$

Como $T - kI$ es inyectivo, también lo es S. Por el caso anterior, S es sobreyectiva. Sea ahora $y \in X$. Entonces $T(y) \in Im(T)$ y por tanto existe $u \in Im(T)$ tal que $S(u) = T(y)$. Es decir,

$$(T - kI)(u) = T(y), \text{ equivalentemente } T(u - y) = ku.$$

Sea $x = \frac{u-y}{k}$. Entonces

$$T(x) = T\left(\frac{u-y}{k}\right) = \frac{ku}{k} = u = kx + y,$$

es decir

$$(T - kI)(x) = y$$

y por tanto $T - kI$ es sobreyectiva, y ahora terminamos con el Teorema de la Aplicación Abierta. □

En cambio no es difícil dar ejemplos de operadores para los que $\sigma_e(T) \neq \sigma_a(T)$ y $\sigma_a(T) \neq \sigma(T)$ ([**31**, 12.7]).

Empecemos a estudiar el espectro de un operador

TEOREMA 8.11. *Sea X un espacio de Banach y $T \in \mathcal{L}(X)$. Entonces*

1. *Sea $k \in \mathbb{K}$ tal que $|k|^m > \|T^m\|$ para algún $m \in \mathbb{N}$. Entonces $k \notin \sigma(T)$ y*
$$(T - kI)^{-1} = -\sum_{n=0}^{\infty} \frac{T^n}{k^{n+1}}.$$
Por tanto, para todo $k \in \sigma(T)$ se tiene que
$$|k| \leq \inf_n \|T^n\|^{\frac{1}{n}} \leq \|T\|.$$
2. *$\sigma(T) \subset \mathbb{K}$ es un conjunto compacto*
3. *Si X tiene dimensión infinita y $T \in \mathcal{K}(X)$ entonces $0 \in \sigma(T)$.*

DEMOSTRACIÓN. 1. Notemos que $k \neq 0$ y que

$$T - kI = -k\left(I - \frac{T}{k}\right).$$

El Lema 8.5 nos dice que $T - kI$ es inversible y que

$$(T - kI)^{-1} = -\frac{1}{k}\left(I - \frac{T}{k}\right)^{-1} = -\frac{1}{k}\sum_{n=0}^{\infty}\frac{T^n}{k^n} = -\sum_{n=0}^{\infty}\frac{T^n}{k^{n+1}}$$

El resto es fácil

2. De lo anterior se sigue que $\sigma(T) \subset \{k \in \mathbb{K}; |k| \leq \|T\|\}$, y por tanto está acotado. Sólo tenemos que ver que es cerrado. Sea $(k_n)_n \subset \sigma(T)$ una sucesión tal que $k_n \to k$. Entonces es trivial que

$$T - k_n I \to T - kI$$

Puesto que los operadores inversibles formaban un conjunto abierto $\mathcal{C}$, se sigue que su complementario, los no inversibles, forman un conjunto cerrado. Puesto que, para todo $n \in \mathbb{N}$ $T - k_n I$ no es inversible, se sigue que $T - kI$ tampoco lo es, es decir $k \in \sigma(T)$.

3. Un operador inversible T define una norma $|||\cdot|||$ en X equivalente a la norma original de X dada por

$$|||x||| = \|T(x)\|.$$

Por tanto si T fuera inversible definiría en X una norma equivalente cuya bola unidad $T(B_X)$ es un compacto, lo que implicaría que X es de dimensión finita. $\square$

Ejemplo 8.12. *Sea $k : [0,1] \times [0,1] \longrightarrow \mathbb{K}$ una función continua y sea $T : C[0,1] \longrightarrow C[0,1]$ el operador definido como*

$$T(f)(s) = \int_0^1 k(s,t)f(t)dt \text{ para todo } s \in [0,1].$$

En el Ejemplo 7.11 ya vimos que T está bien definido y es compacto. Además

$$\|T(f)\| = \sup_s |T(f)(s)| \leq \sup_{s,t} |k(s,t)| \|f\| = \|k\|_\infty \|f\|_\infty,$$

luego

$$\|T\| \leq \|k\|_\infty.$$

En ese caso decimos que T es un operador integral de Fredholm con núcleo continuo $k(\cdot,\cdot)$.

Si S es otro operador integral de Fredholm con núcleo continuo $h(\cdot,\cdot)$, es fácil ver que TS es también un operador integral de Fredholm con núcleo continuo $k * h$, donde para todos $0 \leq s,t \leq 1$,

$$k * h(s,t) = \int_0^1 k(s,u)h(u,t)du$$

y

$$\|TS\| \leq \|k\|_\infty \|h\|_\infty.$$

Haciendo $T = S$ (es decir, $h = k$) y aplicando inducción tenemos que para todo $n \in \mathbb{N}$ T^n es un operador integral de Fredholm con núcleo *continuo*

$$k^{(n)}(s,t) = \int_0^1 \cdots \int_0^1 k(s,u_1)k(u_1,u_2)\cdots k(u_{n-1},t)du_1\cdots du_{n-1}.$$

Sea k un núcleo continuo tal que $\|k\|_\infty \leq 1$. Entonces $\|T\| < 1$ y por el Lema 8.4 se tiene que $I-T$ es inversible e

$$(11)\quad (I-T)^{-1}(f)(s) = \sum_{n=0}^{\infty} T^n(f)(s) = f(s) + \sum_{n=1}^{\infty}\int_0^1 k^{(n)}(s,t)f(t)dt$$

Observemos que la serie $\sum_{n=1}^{\infty} k^{(n)}(s,t)$ converge uniforme y absolutamente en $(s,t)\in[0,1]\times[0,1]$ (ya que para todo $n\in\mathbb{N}$ $\|k^{(n)}\| \leq \|k\|^n$) y por tanto

$$h(s,t) := \sum_{n=0}^{\infty} k^{(n)}(s,t)$$

es una función continua en $[0,1]\times[0,1]$. Además, por el Teorema de la Convergencia Acotada podemos intercambiar el sumatorio y la integral en (11) y tenemos que

$$(I-T)^{-1}(f)(s) = f(s) + \int_0^1 h(s,t)f(t)dt,$$

es decir

$$(I-T)^{-1} = I + B,$$

donde B es otro operador integral de Fredholm con núcleo continuo.

Si x es un autovector del autovalor k de T entonces se tiene que $T(span[x]) \subset span[x]$. Decimos entonces que $span[x]$ es *invariante* por T. En general un subespacio $V \subset X$ se dice invariante si por T si $T(V)\subset V$. Claramente $\{0\}$ y X son subespacios invariantes para todo $T\in\mathcal{L}(X)$. Durante un largo tiempo no se supo si existía algún espacio de Banach X y algún operador definido en él T que no tuviera ningún subespacio invariante no trivial (esto es, distinto de $\{0\}$ y X); recientemente P. Enflo ([**18**]) encontró tales X y T. Aún no se sabe si dado un espacio de *Hilbert* H existe $T\in\mathcal{L}(H)$ sin subespacios invariantes no triviales. En cambio, recientemente Lomonosov ha probado que todo operador *compacto* entre espacios de Hilbert admite subespacios invariantes no triviales. Una exposición de estos problemas se puede ver en [**20**] y la bibliografía allí citada.

En relación a esto, veamos a continuación un ejemplo de un operador entre espacios de Hilbert sin autovalores.

Ejemplo 8.13. *Sea* $T : L_2[0,1] \longrightarrow L_2[0,1]$ *(sobre los números complejos) dado por*

$$T(f(t)) = tf(t)$$

para todo $t \in [0,1]$.

Para ver que no tiene autovalores, si $k \in \mathbb{C}$ *fuera un autovalor, existiría* $f(t) \in L_2[0,1]$ *tal que, para casi todo* $t \in [0,1]$

$$tf(t) = kf(t), \text{ es decir } (k-t)f(t) = 0,$$

de donde, tomando $t \neq k$, *tenemos que* $f(t) = 0$ *para casi todo* $t \in [0,1]$, *es decir* $f(t) = 0$.

En cambio, veamos que todo $k \in [0,1]$ *pertenece al espectro de* T. *Sea* $\lambda \in [0,1]$ *y sea* $\epsilon > 0$ *tal que* $[\lambda, \lambda+\epsilon] \subset [0,1]$ *o* $[\lambda-\epsilon, \lambda] \subset [0,1]$. *Supongamos por ejemplo que* $[\lambda, \lambda+\epsilon] \subset [0,1]$. *Sea*

$$f_\epsilon = \begin{cases} \dfrac{1}{\sqrt{\epsilon}} \quad , & \text{si } t \in [\lambda, \lambda+\epsilon] \\ \\ 0 \quad , & \text{si } t \notin [\lambda, \lambda+\epsilon] \end{cases}$$

Claramente $\|f_\epsilon\|_2 = 1$. *Además*

$$(T-\lambda I)(f_\epsilon)(t) = T(f_\epsilon)(t) - \lambda f_\epsilon(t) = tf_\epsilon(t) - \lambda f_\epsilon(t) = f_\epsilon(t-\lambda)$$

y por tanto

$$\|(T-\lambda I)(f_\epsilon)\|_2^2 = \int_0^1 \frac{1}{\epsilon}(\lambda - t)^2 dt = \cdots = \frac{\epsilon^2}{3}$$

y por tanto $(T-\lambda I)(f_\epsilon) \to 0$ *cuando* ϵ *tiende a* 0. *En consecuencia* $T - \lambda I$ *no está acotado inferiormente y por ello no puede ser inversible.*

Observemos finalmente que T *tiene una gran abundancia de subespacios invariantes no triviales. En particular, es fácil ver que para todo* $r \in (0,1)$, $L_2[0,r]$ *es invariante por* T.

Uno de los resultados fundamentales de la Teoría Espectral es el Teorema de Gelfand-Mazur que enunciamos a continuación. Su demostración es una bonita aplicación de la Teoría de Variable Compleja.

Teorema 8.14 (Gelfand). *Si* X *es un espacio de Banach complejo y* $T \in \mathcal{L}(X)$ *entonces* $\sigma(T) \neq \emptyset$.

Demostración. Sea $f \in \mathcal{L}(X)^*$. Recordemos que $\rho(T) = \sigma(T)^c$ es la *resolvente* de T. Definimos la función

$$w_f : \rho(T) \longrightarrow \mathbb{C}$$

como

$$w_f(z) = f((T - zI)^{-1}).$$

Sabemos que $\rho(T) \subset \mathbb{C}$ es un abierto (porque $\sigma(T)$ es cerrado). Veamos que w_f es analítica en $\rho(T)$: sea $z_0 \in \rho(T)$. Entonces

$$(T - z_0 I) - (T - zI) = (z - z_0)I$$

y por tanto, componiendo por la derecha con $(T - zI)^{-1}$ y por la izquierda con $(T - z_0 I)^{-1}$ tenemos

$$(T - zI)^{-1} - (T - z_0 I)^{-1} = (z - z_0)(T - z_0 I)^{-1}(T - zI)^{-1}$$

y por lo tanto

$$\lim_{z \to z_0} \frac{w_f(z) - w_f(z_0)}{z - z_0} = \lim_{z \to z_0} f\left(\frac{(T - zI)^{-1} - (T - z_0 I)^{-1}}{z - z_0}\right) =$$
$$= \lim_{z \to z_0} f\left((T - z_0 I)^{-1}(T - zI)^{-1}\right) = f((T - z_0 I)^{-1})$$

por la continuidad de f y por la continuidad de la inversión probada en el Corolario 8.7.

Supongamos ahora que $\sigma(T) = \emptyset$. Entonces $\rho(T) = \mathbb{C}$ por lo que w_f es analítica en $\mathbb{C}$, es decir, es una función entera. Además, si $|z| > \|T\|$, usando el Lema 8.4 se tiene que

$$\|(T - zI)^{-1}\| = \left\| \frac{1}{z}\left(\frac{T}{z} - I\right)^{-1} \right\| \leq \tag{12}$$
$$\leq \frac{1}{|z| - \|T\|} \to 0 \text{ cuando } |z| \to \infty.$$

Como f es lineal y continua,

$$w_f(z) = f((T - zI)^{-1}) \to 0 \text{ cuando } |z| \to \infty$$

y por tanto w_f está acotada en $\mathbb{C}$.

Por tanto el Teorema de Liouville nos dice que w_f, siendo entera y acotada debe de ser constante. De (12) se sigue que w_f debe de ser siempre 0. En particular

$$0 = w_f(0) = f(T^{-1}).$$

Como esto ocurre para todo $f \in \mathcal{L}(X)^*$ se sigue que $T^{-1} = 0$, lo cual es imposible. Por tanto $\sigma(T) \neq 0$. $\square$

Si se desea, y en función del tiempo disponible, en relación con esto se puede definir el *radio espectral* de un operador $T : X \longrightarrow X$ como

$$r_\sigma(T) = \sup_{k \in \sigma(T)} |k|$$

y a continuación demostrar la

PROPOSICIÓN 8.15 (Fórmula del Radio Espectral). *Si X es un espacio de Banach complejo y $T \in \mathcal{L}(X)$ entonces*

$$r_\sigma(T) = \inf_{n\in\mathbb{N}} \|T^n\|^{\frac{1}{n}} = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}}.$$

DEMOSTRACIÓN. Una demostración adecuada para este curso se puede ver en [**31**, 12.8 b]. □

Pasamos a estudiar la llamada Teoría de Riesz-Schauder. Necesitamos unos cuantos resultados previos, interesantes en sí mismos, antes de llegar a los resultados principales de la teoría. Hemos seguido para esta sección [**33**] principalmente.

LEMA 8.16. *Sea X un espacio de Banach, $T \in \mathcal{K}(X)$. Entonces $\ker(I-T)$ tiene dimensión finita.*

DEMOSTRACIÓN. Sea U la bola unidad cerrada de $\ker(I-T)$. Entonces $T(U) = U$. Puesto que T es compacto y $U \subset B_X$, se tiene que $U = T(U)$ es relativamente compacto, y por tanto compacto (puesto que es cerrado). Por tanto, si la bola unidad es compacta el espacio debe de ser de dimensión finita. □

LEMA 8.17. *Sea X un espacio de Banach, $T \in \mathcal{K}(X)$. Si $(x_n)_n \subset X$ es una sucesión acotada tal que $x_n - T(x_n) \to y \in X$ entonces existen $x \in X$ y una subsucesión $(x_{n_j})_j \subset (x_n)_n$ tales que $x_{n_j} \to x \in X$ y*

$$x - T(x) = y.$$

DEMOSTRACIÓN. Por la compacidad de T, existe una subsucesión $(x_{n_j})_j \subset (x_n)_n$ tal que $T(x_{n_j})$ converge a $z \in X$. Entonces

$$x_{n_j} = x_{n_j} - T(x_{n_j}) + T(x_{n_j}) \to y + z$$

Llamamos $x = y + z$ y, puesto que T es continuo, se tiene

$$(I-T)(x) = \lim_j x_{n_j} - T(x_{n_j}) = y + z - z = y.$$

□

LEMA 8.18. *Sea X un espacio de Banach, $T \in \mathcal{K}(X)$. Entonces $Im(I-T)$ es cerrado y tiene codimensión finita.*

DEMOSTRACIÓN. Veamos en primer lugar que $Im(I-T)$ es cerrado. Sea $q : X \longrightarrow X/\ker(I-T)$ la aplicación cociente. Entonces sabemos que $I-T$ factoriza como

$$I - T = (\widetilde{I-T}) \circ q$$

donde

$$\widetilde{I-T} : X/\ker(I-T) \longrightarrow X$$

es la aplicación canónica dada por

$$\widetilde{I-T}([x]) = (I-T)(x).$$

Es un ejercicio comprobar que $\widetilde{I-T}$ está bien definida, es lineal y continua, con $\|\widetilde{I-T}\| = \|I-T\|$.

Veamos que existe $c > 0$ tal que, para todo $x \in X$

$$\|x - T(x)\| = \|(\widetilde{I-T})([x])\| \geq c\|[x]\| \tag{13}$$

Si no fuera así, existiría una sucesión $(x_n) \subset X$ tal que

$$\lim_n \|x_n - T(x_n)\| = 0 \text{ y } \|[x_n]\| = 1 \text{ para todo } n \in \mathbb{N}. \tag{14}$$

Por la definición de la norma cociente podemos suponer que, para todo $n \in \mathbb{N}$,

$$1 \leq \|x_n\| \leq 2.$$

Puesto que T es compacto, existen $y \in X$ y una subsucesión $(x_{n_j})_j \subset (x_n)_n$ tal que $T(x_{n_j}) \to y \in X$. Por (14) sabemos que

$$\lim_j x_{n_j} = y$$

y por tanto

$$y - T(y) = \lim_j x_{n_j} - T(x_{n_j}) = y - y = 0,$$

es decir $y \in \ker(I-T)$. Pero ahora usando la otra mitad de (14) tenemos una contradicción puesto que

$$1 = \|[x_{n_j}]\| \leq \|x_{n_j} - y\| \to 0$$

Por tanto sabemos que (13) es cierto. De (13) se sigue que $(\widetilde{I-T})$ es un isomorfismo sobre su imagen y en particular $Im(\widetilde{I-T}) \subset X$ es cerrado. Puesto que

$$Im(I-T) = Im(\widetilde{I-T}),$$

se sigue que $Im(I-T)$ es cerrado.

Veamos ahora que $Im(I-T)$ tiene codimensión finita. Puesto que $Im(I-T) \subset X$ es un subespacio cerrado, podemos considerar el espacio cociente $X/Im(I-T)$ y sabemos que

$$(X/Im(I-T))^* = (Im(I-T))^{\perp}$$

y por otro lado es un ejercicio ver que $(Im(I-T))^{\perp} = \ker((I-T)^*) = \ker(I^* - T^*) = \ker(I_{X^*} - T^*)$. Puesto que T es compacto, el Teorema de Schauder nos dice que T^* es compacto. Ahora el Lema 8.16 nos dice que $\ker(I_{X^*} - T^*)$ tiene dimensión finita, y de

$$\ker(I_{X^*} - T^*) = (X/Im(I-T))^*$$

se sigue que $X/Im(I-T)$ tiene dimensión finita, es decir $Im(I-T)$ tiene codimensión finita. □

Observemos que si X es un espacio de Banach y $T \in \mathcal{L}(X)$, por la fórmula del binomio de Newton, para todo $n \in \mathbb{N}$ se tiene

$$(I-T)^n = I - \sum_{j=1}^{n} \binom{n}{j} (-1)^{j-1} T^j =: I - T_n$$

Si además T es compacto, puesto que los operadores compactos forman un subespacio vectorial, se sigue que cada uno de los T_n así definidos es compacto. Por tanto, para todo $n \in \mathbb{N}$:

- $\ker(I-T)^n$ tiene dimensión finita, y claramente $\ker(I-T)^n \subset \ker(I-T)^{n+1}$.
- $Im(I-T)^n$ tiene codimensión finita, es cerrado y $Im(I-T)^n \supset Im(I-T)^{n+1}$.

Veamos que los núcleos $\ker(I-T)^n$ no pueden crecer indefinidamente, sino que a partir de cierto n_0 se estabilizan.

PROPOSICIÓN 8.19. *Sea X un espacio de Banach y $T \in \mathcal{K}(X)$. Entonces existe $n_0 \in \mathbb{N}$ tal que, para todo $n \geq n_0$,*

$$\ker(I-T)^{n_0} = \ker(I-T)^n.$$

DEMOSTRACIÓN. Si no fuera así, podemos suponer $\ker(I-T)^n \subsetneq \ker(I-T)^{n+1}$ para todo $n \in \mathbb{N}$ (si no, razonamos análogamente tomando subsucesiones); en ese caso existiría una sucesión $(x_n)_n \subset X$ tal que, para todo $n \in \mathbb{N}$

$$x_n \in \ker(I-T)^{n+1} \setminus \ker(I-T)^n$$

y

$$1 = \inf_{y \in \ker(I-T)^n} \|x_n - y\| \leq \|x_n\| \leq 2 \tag{15}$$

(donde hemos usado el Lema de Riesz para la segunda afirmación).

Entonces, si $m < n$, se tiene

$$(I-T)^n((I-T)(x_n) + x_m - (I-T)(x_m)) = 0$$

por lo que

$$(I-T)(x_n)+x_m-(I-T)(x_m)\in \ker(I-T)^n$$

y por tanto

$$\|T(x_n)-T(x_m)\| = \|x_n-((I-T)(x_n)+x_m-(I-T)(x_m)\| \geq 1$$

por (15).

Por tanto la sucesión $(T(x_n))_n$ no puede tener subsucesiones convergentes, en contradicción con el hecho de que T sea compacto. $\square$

LEMA 8.20. *Sea X un espacio de Banach y $T\in\mathcal{K}(X)$. Entonces existe $n\in\mathbb{N}$ tal que los subespacios cerrados $N:=\ker(I-T)^n$ y $R:=Im(I-T)^n$ verifican*

1. *$codimR<\infty$*
2. *$N\oplus R=X$*
3. *$(I-T)(N)\subset N$*
4. *$(I-T)(R)\subset R$*
5. *$(I-T)_R:=(I-T)_{|R}\in\mathcal{L}(R)$ es invertible*
6. *$(I-T)_N:=(I-T)_{|N}\in\mathcal{L}(N)$ es nilpotente, en concreto $(I-T)_N^n=0$*

DEMOSTRACIÓN. Aplicando el Lema 8.19 a $(I-T)$, a $(I-T)^*=I-T^*$ y a $(I-T)^{**}=I-T^{**}$ tenemos que existe $n\in\mathbb{N}$ tal que, para todo $m\geq n$,

$$\ker(I-T)^n=\ker(I-T)^m,\quad \ker(I-T^*)^n=\ker(I-T^*)^m,\quad \text{y}$$
$$\ker(I-T^{**})^n=\ker(I-T^{**})^m.$$

Usando que $Im(I-T)^m$ es cerrado para todo $m\in\mathbb{N}$, ya hemos visto en la demostración del Lema 8.18 que

$$Im(I-T)^m=(\ker(I-T^*)^m)^\perp.$$

Por tanto, para todo $m\geq n$ se tiene

$$Im(I-T)^n=(\ker(I-T^*)^n)^\perp=(\ker(I-T^*)^m)^\perp=Im(I-T)^m.$$

Sean entonces N y R como en el enunciado. Ambos son cerrados y $\dim N<\infty$ y $codimR<\infty$ (esto ya está probado en el Lema 8.18). Además

$$(I-T)(N)=(I-T)(\ker(I-T)^{n+1})\subset\ker(I-T)^n=N$$

y

$$(I-T)(R) = (I-T)(Im(I-T)^n) = Im(I-T)^{n+1} = R.$$

Por la elección de n, $(I-T)^n_N = 0$. Acabamos de ver que $(I-T)_R$ es sobre. Veamos que también es inyectivo: Si

$$(I-T)_R(x) = 0,$$

entonces se tiene que $x \in \ker(I-T)^{n+1} = \ker(I-T)^n$, es decir,

$$(I-T)^n(x) = 0.$$

Es decir, $(I-T)_R$ es inyectiva y por tanto biyectiva. De aquí se sigue que $N \cap R = \{0\}$ -en efecto, si $x \in N$ entonces $(I-T)^n(x) = 0$; además, si $x \in R$, por ser $(I-T)^n$ biyectiva sobre R se tiene que $(I-T)^n(x) = 0$ si y sólo si $x = 0$- y por tanto

$$\dim N \leq codim R.$$

Aplicando un razonamiento análogo a T^* en lugar de T, se sigue que

$$\dim \ker(I-T^*)^n \leq codim(I-T^*)^n(X).$$

Como en la prueba del Lema 8.18 podemos ver que

$$\ker(I-T^*)^n = X/Im(I-T)^n$$

y por tanto

$$codim R = \dim(X/Im(I-T)^n) = \dim \ker(I-T^*)^n$$

y puesto que $Im(I-T^*)^n$ es cerrado, de nuevo razonando como en el Lema 8.18 tenemos que

$$\begin{aligned}(\ker(I-T)^n)^* &= X^*/(\ker(I-T)^n)^{\perp} = \\ &= X^*/((I-T)^n)^*(X) = X^*/(I-T^*)^n(X)\end{aligned}$$

y por tanto

$$\dim N = \dim \ker(I-T)^n = \dim(\ker(I-T)^n)^* = codim(I-T^*)^n(X)$$

y juntando todo se tiene que

$$\dim N = codim(I-T^*)^n(X) \geq \dim \ker(I-T^*)^n = codim R$$

Así pues se tiene que

$$\dim N = codim R$$

de donde se sigue que $X = N \oplus R$ □

Proposición 8.21. *Sea X un espacio de Banach de dimensión infinita y sea $T \in \mathcal{K}(X)$. Entonces, para todo $k \in \sigma(T) \setminus \{0\}$ existen dos subespacios de X topológicamente complementarios N_k y R_k invariantes por $kI - T$ y que verifican*

1. $(kI - T)_{|R_k}$ *es un isomorfismo de R_k en sí mismo.*
2. *Existe $n_k \in \mathbb{N}$ tal que*
$$(kI - T)_{|N_k}^{n_k} \equiv 0,$$
es decir, $(kI - T)_{|N_k}^{n_k}$ es nilpotente.
3. $\{0\} \neq \ker(kI - T) \subset N_k$ *y* $\dim N_k < \infty$.

Demostración. Puesto que $\frac{1}{k}T$ es compacto para todo $k \neq 0$, podemos aplicar la Proposición 8.20 a $(I - \frac{1}{k}T) = \frac{1}{k}(kI - T)$ para obtener $n_k, N_k, R_k, (I - T)_{N_k}$ e $(I - T)_{R_k}$ con las propiedades allí establecidas. Entonces R_k y N_k son topológicamente complementarios y (1) y (2) quedan probados.

Puesto que $(kI - T)_{|R_k}$ es un isomorfismo (esto es (1)) y dado que $X = N_k \oplus R_k$, si N_k fuera $\{0\}$, tendríamos que $(kI - T)$ sería inversible, en contradicción con que $k \in \sigma(T)$. Por tanto $N_k \neq \{0\}$. Ahora, dado que $(kI - T)_{|N_k}$ es nilpotente, no puede ser inyectivo, por lo que existe $0 \neq x_0 \in N_k$ tal que $(kI - T)(x_0) = 0$, es decir

$$x_0 \in \ker(kI - T)$$

lo que prueba que $\ker(kI - T) \neq \{0\}$.

Puesto que $N_k = \ker(kI - T)^{n_k}$ (ver Lema 8.20), está claro que $\ker(kI - T) \subset \ker(kI - T)^{n_k}$, y que

$$\dim N_k < \infty$$

se sigue del Lema 8.20. □

Finalmente podemos probar el resultado principal de la Teoría de Riesz-Schauder del espectro de operadores compactos.

Teorema 8.22. *Sea X un espacio de Banach de dimensión infinita, y $T \in \mathcal{K}(X)$. Entonces*

(i) $0 \in \sigma(T)$
(ii) *Todo $k \in \sigma(T) \setminus \{0\}$ es un autovalor, y el autoespacio correspondiente E_k tiene dimensión finita.*
(iii) *Existe una sucesión convergente a cero $(k_n)_n \in \mathbb{N}$ (quizás eventualmente constante) tal que $\sigma(T) = \{0\} \cup \{k_n; n \in \mathbb{N}\}$.*
(iv) *Para todo $k \in \mathbb{K} \setminus \{0\}$ se tiene*
$$\dim \ker(kI - T) = codim(kI - T)(X).$$

Demostración. (i) y (ii) ya han sido probados en la Proposición 8.21.

Para probar (iii), probemos en primer lugar el siguiente

Aserto: Para todo $k \in \mathbb{K} \setminus \{0\}$ existe un entorno abierto U_k de k tal que $kI - T$ es un isomorfismo de X para todo $z \in U_k \setminus \{k\}$.

Demostración del Aserto: Puesto que $\sigma(T)$ es cerrado, su complementario es abierto, por lo que dado $k \in \mathbb{K} \setminus \sigma(T)$ existe un entorno U_k de k totalmente contenido en $\mathbb{K} \setminus \sigma(T)$.

Ahora, si $k \in \sigma(T) \setminus \{0\}$ y R_k y N_k son como en la Proposición 8.21, entonces

$$(kI - T)_{|_{R_k}} = kI_{R_k} - T_{|_{R_k}}$$

es inversible, es decir $k \notin \sigma(T_{|_{R_k}})$. Puesto que $T_{|_{R_k}}$ es compacto, de nuevo $\sigma(T_{|_{R_k}})$ es compacto y $\mathbb{K} \setminus \sigma(T_{|_{R_k}})$ es un abierto. Por tanto, existe $U_k \setminus \mathbb{K}$ un entorno de k tal que $(zI - T)$ es inversible para todo $z \in U_k$. Como $(kI - T)_{|_{N_k}}$ es nilpotente, se tiene que $\sigma((kI - T)_{|_{N_k}}) = \{0\}$ (esto es un ejercicio), es decir, para todo $z \neq 0$ se verifica que $(zI - T)_{|_{N_k}}$ es inversible (como elemento de $\mathcal{L}(N_k)$).

Juntando ambas cosas tenemos que para todo $z \in U_k$ $(zI - T)_{|_{N_k}}$ y $(zI - T)_{|_{R_k}}$ son inversibles. Como $X = N_k \oplus R_k$ tenemos que $(zI - T)$ es inversible (esto es un ejercicio) lo que prueba el aserto.

Ahora, puesto que $\sigma(T)$ es compacto, para todo $\epsilon > 0$ el conjunto

$$M_\epsilon = \{k \in \sigma(T); |k| \geq \epsilon\} = \sigma(T) \cap \{k \in \mathbb{K}; \epsilon \leq |k| \leq \|T\|\}$$

es compacto. Del Aserto se sigue que ningún punto de M_ϵ es de acumulación, por lo que M_ϵ ha de ser finito. De aquí se sigue que $\sigma(T)$ es (a lo sumo) numerable.

Finalmente probemos (iv). Si $k \notin \sigma(T)$ entonces $kI - T$ es inversible, $\ker(kI - T) = 0$ y $Im(kI - T) = X$ y (iv) es trivial.

Si $k \in \sigma(T) \setminus \{0\}$ entonces $(kI - T)_{|_{R_k}}$ es inversible y $\dim N_k < \infty$. Entonces

$$\dim \ker(kI - T) = \dim \ker(kI - T)_{|_{N_k}} =$$

$$= codimIm(kI - T)_{|_{N_k}} = codimIm(kI - T).$$

□

Como corolario resaltamos de forma explícita la llamada Alternativa de Fredholm, que enunciamos más abajo. Antes de enunciarla, intentaremos motivarla brevemente.

Pensemos en el sistema de n ecuaciones lineales con n incógnitas

$$Ax = y$$

y sea A^t la matriz traspuesta de A. Ya hemos visto que si pensamos en A como un operador $T_A : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ entonces A^t es la matriz asociada al operador $(T_A)^*$. Con esta notación, los siguientes resultados del Álgebra Lineal son bien conocidos

1. El sistema $Ax = y$ tiene una única solución para todo $y \in \mathbb{K}^n$ si y sólo si el sistema homogéneo asociado $Ax = 0$ tiene $x = (0, \ldots, 0)$ como única solución.
2. El sistema homogéneo $Ax = 0$ tiene una solución distinta de $(0, \ldots, 0)$ si y sólo si el sistema traspuesto $A^t x = 0$ tiene una solución distinta de $(0, \ldots, 0)$. Además, en ese caso el máximo número de soluciones linealmente independientes de ambos sistemas es el mismo.

Es fácil ver que este resultado *no* se puede extender al caso de una colección numerable de ecuaciones lineales con una cantidad numerable de incógnitas. Por ejemplo, considérese el sistema infinito con matriz asociada dada por

$$A = (a_{i,j})_{i,j \in \mathbb{N}} \text{ con } a_{j+1,j} = \frac{1}{j} \text{ para todo } j \text{ y } a_{i,j} = 0 \text{ si } i \neq j+1.$$

No es difícil ver que el sistema homogéneo asociado $Ax = 0$ sólo admite la solución $x_1 = x_2 = \cdots = 0$. En cambio, si $y = (1, 0, 0, \ldots)$ el sistema $Ax = y$ no tiene solución. Además, el sistema traspuesto $A^t x = 0$ admite la solución $(1, 0, 0, \ldots)$. El resultado probado por Fredholm nos muestra precisamente que el resultado del Álgebra Lineal arriba mencionado *sí* se mantiene si el operador A es de la forma $I - T$, con T un operador compacto. En concreto tenemos

Teorema 8.23 (Alternativa de Fredholm). *Sea X un espacio de Banach y sea $T \in \mathcal{K}(X)$. Entonces*

1. *Se verifica una y sólo una de las siguientes posibilidades*
   a) *Para todo $y \in X$ existe un único $x \in X$ tal que $(I-T)(x) = y$.*
   b) *Existe $0 \neq x \in X$ tal que $(I - T)(x) = 0$. Además, en este caso el máximo número de soluciones linealmente independientes de la ecuación homogénea $(I - T)(x) = 0$ es finito.*

2. *La ecuación homogénea* $(I-T)(x) = 0$ *tiene una solución distinta de* 0 *en* X *si y sólo si la ecuación traspuesta* $(I-T^*)(x^*) = 0$ *tiene una solución distinta de* 0 *en* X^*. *Además el máximo número de soluciones linealmente independientes de ambas ecuaciones es el mismo.*

OBSERVACIÓN 8.24. *Sean* X *y* T *como en el Teorema y para todo* $k \in \mathbb{K}$ *consideremos el operador* kT. *Este operador es compacto, por serlo* T. *Por lo tanto, para la ecuación*

$$(I - kT)(x) = y$$

se tienen las dos posibilidades enunciadas en el Apartado 1 del Teorema. Nótese que la segunda posibilidad se tiene precisamente cuando $\frac{1}{k}$ *es un autovalor de* T, *y que sabemos que el conjunto de autovalores es a lo sumo numerable.*

Veamos cómo se puede aplicar la Alternativa de Fredholm a la resolución de sistemas numerables de ecuaciones lineales y a la resolución de ecuaciones integrales, el problema al que se enfrentaba Fredholm y que motivó el desarrollo de la teoría.

TEOREMA 8.25. *Sea* $A = (a_{i,j})_{i,j\in\mathbb{N}}$ *una matriz infinita y sea* $1 \leq p < \infty$. *Si* $p = 1$, *supongamos que* $\gamma(j) = \sum_{i=1}^{\infty} |a_{i,j}| < \infty$ *para todo* $j \in \mathbb{N}$ *y que* $\gamma(j) \to 0$ *cuando* $j \to \infty$. *Si* $1 < p < \infty$ *y* $\frac{1}{p} + \frac{1}{q} = 1$, *supongamos que*

$$\sum_{i=1}^{\infty} \left(\sum_{j=1}^{\infty} |a_{i,j}|^q \right) \frac{p}{q} < \infty.$$

Entonces

1. *Se tiene una de las dos alternativas siguientes*
   a) *O bien para todo* $y \in \ell_p$ *existe un único* $x \in \ell_p$ *tal que*

$$\begin{aligned} (1 - a_{1,1})x_1 - a_{1,2}x_2 - a_{1,3}x_3 - \cdots &= y_1 \\ -a_{2,1}x_1 + (1 - a_{2,2})x_2 - a_{2,3}x_3 - \cdots &= y_2 \\ -a_{3,1}x_1 - a_{3,2}x_2 + (1 - a_{3,3})x_3 - \cdots &= y_3 \\ \vdots \qquad \vdots \qquad \vdots & \end{aligned}$$

   b) *O bien existe* $0 \neq x \in \ell_p$ *tal que*

$$(1 - a_{1,1})x_1 = a_{1,2}x_2 + a_{1,3}x_3 + \cdots \tag{16}$$

$$(1 - a_{2,2})x_2 = a_{2,1}x_1 + a_{2,3}x_3 + \cdots \tag{17}$$

$$(1 - a_{3,3})x_3 = a_{3,1}x_1 + a_{3,2}x_2 + \cdots \tag{18}$$

$$\vdots \qquad \vdots \qquad \vdots \tag{19}$$

En este caso no *pueden existir un número infinito de vectores $x \in \ell_p$ linealmente independientes que verifiquen (16)*

2. *El máximo número de vectores $x \in \ell_p$ linealmente independientes que verifiquen (16) es igual al máximo número de vectores $x \in \ell_q$ linealmente independientes que verifiquen el sistema traspuesto*

$$\begin{aligned}
(1-a_{1,1})x_1 &= a_{2,1}x_2 + a_{3,1}x_3 + \cdots \\
(1-a_{2,2})x_2 &= a_{1,2}x_1 + a_{3,2}x_3 + \cdots \\
(1-a_{3,3})x_3 &= a_{1,3}x_1 + a_{2,3}x_2 + \cdots \\
&\vdots \qquad\quad \vdots \qquad\quad \vdots
\end{aligned}$$

Demostración. Ya hemos visto en el Ejemplo 7.10 que la matriz A define un operador compacto $T : \ell_p \longrightarrow \ell_p$. Es un ejercicio comprobar que A^t define precisamente el operador traspuesto $T^* : \ell_q \longrightarrow \ell_q$. Por lo tanto no hay más que aplicar la alternativa de Fredholm. $\square$

Veamos el análogo continuo del ejemplo de arriba, lo que se conoce como ecuación integral de Fredholm de segunda clase.

Teorema 8.26. *Sea $k(\cdot,\cdot) \in L_2([0,1] \times [0,1])$. Entonces*

1. *Se tiene una y sólo una de las dos siguientes alternativas*
   a) *O bien para todo $y \in L_2[0,1]$ existe un único $x \in L_2[0,1]$ tal que para casi todo $s \in [0,1]$*

$$x(s) - \int_0^1 k(s,t)x(t)dt = y(s)$$

   b) *O bien existe $0 \neq y \in L_2[0,1]$ tal que para casi todo $s \in [0,1]$*

$$x(s) = \int_0^1 k(s,t)x(t)dy \tag{20}$$

   Además, en este caso el máximo número de $x \in L_2[0,1]$ linealmente independientes que verifican (20) es finito.

2. *La ecuación (20) tiene una solución distinta de 0 si y sólo si la ecuación traspuesta*

$$z(s) = \int_0^1 k(t,s)z(t)dt$$

tiene una solución distinta de 0 en $L_2[0,1]$

DEMOSTRACIÓN. Ya hemos visto en el Ejemplo 7.11 que el operador $T : L_2[0,1] \longrightarrow [0,1]$ dado por

$$T(x)(s) = \int_0^1 k(s,t)x(t)dt$$

es compacto. Una vez que identifiquemos adecuadamente su operador traspuesto, bastará con aplicar de nuevo la alternativa de Fredholm. $\square$

Prácticas sugeridas

EJERCICIO 8.1. *Sea X un espacio de Banach y $T \in \mathcal{L}(X)$. Demostrar que el operador $exp(T) := \sum_{n=0}^{\infty} \frac{T^n}{n!}$ (está bien definido y) es inversible.* **Sugerencia:** $(exp(T))^{-1} = exp(-T)$.

Demostrar además que

$$\sigma(exp(T)) = exp(\sigma(T))$$

EJERCICIO 8.2 (Lomonosov). *Sea X un espacio de Banach de dimensión infinita y sea $T \in \mathcal{K}(X)$. Demostrar que X posee un subespacio invariante no trivial.* [**24**, Ej. 18 p. 143].

EJERCICIO 8.3. *Sea X un espacio de Banach y sea $T \in \mathcal{L}(X)$. Sean $\lambda_1, \ldots, \lambda_n$ autovalores distintos de T y sea e_i un autovector de λ_i $(1 \leq i \leq n)$. Demostrar que los vectores $\{e_1, \ldots, e_n\}$ son linealmente independientes.*

Capítulo 9

Espacios de Hilbert

Los espacios hoy conocidos como *espacios de Hilbert* estuvieron históricamente entre los primeros espacios de Banach estudiados y pronto se comprendió que ocupan un lugar destacado entre todos ellos. La diferencia radical entre un espacio de Hilbert y un espacio de Banach cualquiera es que en un espacio de Hilbert tenemos un gran conocimiento de la "geometría" del espacio, y además esta geometría coincide con nuestra intuición geométrica finito dimensional. Esto se traduce en una serie de teoremas como el Teorema de Representación de Riesz, el Teorema de la Proyección Ortogonal y otros que permiten que trabajar en un espacio de Hilbert sea en general mucho más cómodo que trabajar en otro espacio de Banach. Una de las consecuencias de esta facilidad en el estudio de los espacios de Hilbert es que muchos de los "grandes" teoremas del Análisis Funcional se pueden probar de manera mucho más sencilla en el caso de los espacios de Hilbert. Hemos incluido algún ejemplo de esta situación en el desarrollo del capítulo y planteamos otros ejemplos de esto como ejercicios.

Comenzamos estudiando el *producto escalar*, a continuación la *Desigualdad de Schwartz* y la definición de norma asociada a un producto escalar, lo que nos lleva a la definición de *espacio de Hilbert*. Seguidamente vemos la *Identidad de Polarización*, que nos muestra como la norma caracteriza al producto escalar, la crucial noción de *ortogonalidad* y el *Teorema de Pitágoras* y la *Ley del Paralelogramo*, todas ellas herramientas sencillas de probar pero imprescindibles en la Teoría de espacios de Hilbert.

A continuación probamos el primero de los resultados "grandes" del capítulo, probando que la distancia de un punto a un convexo cerrado en un espacio de Hilbert se alcanza en un único punto. De ahí deducimos el *Teorema de la Proyección Ortogonal*, lo que nos lleva a estudiar la *complementación* en espacios de Hilbert.

Seguidamente comenzamos el estudio de los *conjuntos ortogonales*, lo que nos lleva a probar la *Desigualdad de Bessel* y el *Teorema de Riesz-Fischer*, a definir *Bases Hilbertianas* y a presentar seguidamente el *desarrollo en serie de Fourier* en espacios de Hilbert y la *Fórmula de Parseval*.

No hemos visto necesario definir explícitamente *familias sumables* en todo este desarrollo, sino que comprobamos que ciertos sumatorios formales que aparecen indexados por conjuntos quizás no numerables tienen a lo sumo una cantidad numerable de términos no nulos, por lo que los podemos reducir a series. Creemos que este enfoque es algo más sencillo, pero no descartamos definir familias sumables si con la experiencia nos pareciera mejor alternativa. Otra opción desde luego es limitarse a espacios de Hilbert *separables*.

A continuación se puede, aunque no es fundamental para la asignatura, definir *dimensión hilbertiana* y probar que todo espacio de Hilbert es (isométrico a) un espacio $\ell_2(I)$ para algún conjunto de índices I.

Terminamos el capítulo con el fundamental *Teorema de Riesz* y algunas aplicaciones de la Teoría de espacios de Hilbert.

Como ya hemos dicho, los espacios de Hilbert son centrales en la teoría de espacios de Banach, y muchos conceptos definidos y estudiados en esta memoria en el contexto general de los espacios normados resultan mucho más sencillos si nos restringimos a los espacios de Hilbert. Resulta por ello una opción didáctica a tener en cuenta el comenzar el estudio de la asignatura con este capítulo (adecuadamente modificado, claro está, para no presuponer resultados no conocidos) y proceder a continuación con el estudio de los espacios de Banach generales. Este es el esquema seguido en, por ejemplo, [**13**], donde el autor dice que el número de demostraciones que deben aparecer casi duplicadas por seguir este esquema es mucho menor de lo que él esperaba.

No nos hemos decidido a seguir ese esquema, pero tampoco descartaríamos experimentarlo en alguna ocasión finalmente impartiéramos la asignatura desarrollada en esta memoria.

Hemos seguido en la presentación de este Capítulo los libros [**13**] y [**31**].

Espacios de Hilbert

DEFINICIÓN 9.1. *Sea H un espacio vectorial sobre $\mathbb{K}$. Un* producto escalar *en H es una función*

$$\langle \cdot, \cdot \rangle : H \times H \longrightarrow \mathbb{K}$$

que verifica

1. *Es definida positiva, es decir, para todo* $x \in H$
$$\langle x, x\rangle \geq 0$$
y
$$\langle x, x\rangle = 0 \text{ si y sólo si } x = 0$$
2. *Linealidad en la primera variable, es decir, para todos* $x, y, z \in H$*,* $\alpha, \beta \in \mathbb{K}$*,*
$$\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$$
3. *Es antisimétrica, es decir, para todos* $x, y \in H$
$$\langle x, y\rangle = \overline{\langle y, x\rangle}$$
4. *Como consecuencia de 2 y 3 se tiene que un producto escalar siempre es conjugado-lineal en la segunda variable, es decir, para todos* $x, y, z \in H$*,* $\alpha, \beta \in \mathbb{K}$*,*
$$\langle x, \alpha y + \beta z\rangle = \overline{\alpha}\langle x, y\rangle + \overline{\beta}\langle x, z\rangle$$

Nótese que si estamos trabajando sobre el cuerpo de los reales, un producto escalar es una forma bilineal, simétrica y definida positiva.

Ejemplo 9.2. *Enumeramos a continuación algunos ejemplos sencillos de espacios con un producto escalar.*

- *El ejemplo más elemental es precisamente el producto escalar en* $\mathbb{R}^n$ *que nuestros alumnos ya conocen. La versión compleja de este ejemplo es el producto escalar en* $\mathbb{C}^n$ *que viene dado por*
$$\langle (z_i), (w_i)\rangle = \sum_{i=1}^{n} z_i \overline{w_i}.$$
- *Usando la desigualdad de Hölder se comprueba que la función*
$$\langle \cdot, \cdot\rangle : \ell_2 \times \ell_2 \longrightarrow \mathbb{K}$$
dada por
$$\langle (z_i), (w_i)\rangle = \sum_{i=1}^{\infty} z_i \overline{w_i}$$
está bien definida y es un producto escalar en ℓ_2*.*
- *Si nuestros alumnos ya conocen la medida de Lebesgue, usando la desigualdad de Hölder para integrales se comprueba que la aplicación*
$$\langle \cdot, \cdot\rangle : L_2[0, 1] \times L_2[0, 1] \longrightarrow \mathbb{K}$$

dada por

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}dt$$

está bien definida y es un producto escalar en $L_2[0,1]$.

- *Sea* H *el conjunto de las funciones* $f : [0,1] \longrightarrow \mathbb{K}$ *absolutamente continuas en* $[0,1]$ *tales que* $f(0) = 0$ *y* $f' \in L_2[0,1]$. *Si definimos*

$$\langle f, g\rangle = \int_0^1 f'(t)\overline{g'(t)}dt$$

para todo $f, g \in H$ *entonces* H *es un espacio de Hilbert.*

Si bien nuestros alumnos probablemente ya conocerán la desigualdad de Schwarz, probablemente no esté de más recordarles brevemente su enunciado y demostración.

PROPOSICIÓN 9.3 (Desigualdad de Schwarz). *Si* $\langle\cdot,\cdot\rangle$ *es un producto escalar en* H *entonces para todo* $x, y \in H$ *se tiene que*

$$|\langle x, y\rangle|^2 \le \langle x, x\rangle\langle y, y\rangle$$

y se tiene la igualdad si y sólo si x *e* y *son linealmente dependientes.*

DEMOSTRACIÓN. Sean $x, y \in H$ y sea $z = \langle y, y\rangle x - \langle x, y\rangle y$. Entonces

$$0 \le \langle z, z\rangle =$$

$$= \langle y, y\rangle^2\langle x, x\rangle - \langle y, y\rangle\overline{\langle x, y\rangle}\langle x, y\rangle - \langle x, y\rangle\langle y, y\rangle\langle y, x\rangle + \langle x, y\rangle\overline{\langle x, y\rangle}\langle y, y\rangle =$$

$$= \langle y, y\rangle\left(\langle x, x\rangle\langle y, y\rangle - |\langle x, y\rangle|^2\right)$$

Entonces si $y = 0$ el resultado se sigue trivialmente. Si $y \neq 0$, entonces, puesto que $\langle y, y\rangle > 0$, se tiene que

$$\langle x, x\rangle\langle y, y\rangle - |\langle x, y\rangle|^2 \ge 0$$

y de aquí se sigue la desigualdad.

Si x e y son linealmente dependientes, si uno de ellos es 0, la igualdad se sigue trivialmente. Si no, existe $\alpha \in \mathbb{K}$ tal que $x = \alpha y$, y se ve fácilmente que se da la igualdad. Recíprocamente, si se da la igualdad, tenemos que, con las notaciones anteriores,

$$\langle z, z\rangle = 0$$

de forma que $z = 0$ y por tanto $\langle y, y\rangle x = \langle x, y\rangle y$, de donde se sigue que x e y son linealmente dependientes. □

A partir de la desigualdad de Schwarz se ve fácilmente que el producto escalar permite definir una norma asociada a él

COROLARIO 9.4. *Sea H un espacio vectorial en el que está definido un producto escalar $\langle\cdot,\cdot\rangle$. Entonces la función*

$$\|\cdot\| : H \longrightarrow [0,\infty)$$

definida como

$$\|x\| = (\langle x, x\rangle)^{\frac{1}{2}}$$

es una norma en H.

DEMOSTRACIÓN. La única dificultad reside en la desigualdad triangular. Observemos que para todos $x, y \in H$,

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + \|y\|^2 + 2\Re\langle x, y\rangle.$$

Por otro lado, de la desigualdad de Schwarz se sigue que

$$\Re\langle x, y\rangle \leq |\langle x, y\rangle| \leq \|x\|\|y\|$$

y por tanto

$$\|x+y\|^2 \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2$$

de donde se sigue la desigualdad triangular. □

Probamos a continuación la *Identidad de Polarización* para el producto escalar. Dicha identidad es cierta en el contexto más general de las formas sesquilineales si los espacios son complejos y en de las formas bilineales *simétricas* si los espacios son reales, y así lo podemos mencionar a los alumnos, pero puesto que no la usaremos más que en el producto escalar, sólo la probamos en ese caso.

LEMA 9.5 (Identidad de Polarización). *Sea H un espacio vectorial complejo con un producto escalar definido en él. Entonces*

$$\langle x, y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + \imath\|x+\imath y\|^2 - \imath\|x-\imath y\|^2\right)$$

y

$$\Re\langle x, y\rangle = \frac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right)$$

por lo que en el caso real

$$\langle x, y\rangle = \frac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right)$$

DEMOSTRACIÓN. Inmediata. □

Obsérvese que $\langle x, y\rangle = 0$ para todo $y \in H$ si y sólo si $x = 0$; claramente, si $x = 0$ entonces $\langle x, y\rangle = 0$ para todo $y \in H$. Recíprocamente, si $\langle x, y\rangle = 0$ para todo $y \in H$ en particular $\langle x, x\rangle = 0$ lo que implica que $x = 0$.

Al igual que ocurría en el caso de dimensión finita, la gran ventaja de la geometría de los espacios de Hilbert es que la existencia de producto escalar nos permite definir el concepto de ortogonalidad.

DEFINICIÓN 9.6. *Sea H un espacio de Hilbert, $x, y \in H$. Decimos que x e y son ortogonales, y lo escribimos*

$$x \perp y,$$

si

$$\langle x, y\rangle = 0.$$

Decimos que dos subconjuntos $A, B \subset H$ son ortogonales, $A \perp B$ si para todo $a \in A$, $b \in B$ se tiene que $a \perp b$.

Los alumnos ya saben del caso finito-dimensional que esta noción se corresponde precisamente con la noción geométrica de ortogonalidad.

PROPOSICIÓN 9.7 (Pitágoras). *Sea H un espacio de Hilbert. Si $x_1, \ldots, x_n \in H$ son dos a dos ortogonales (es decir $x_i \perp x_j$ para todo $i \neq j$) entonces*

$$\left\| \sum_{i=1}^{n} x_i \right\|^2 = \sum_{i=1}^{n} \|x_i\|^2$$

DEMOSTRACIÓN. Aplicando inducción, es fácil ver que sólo es necesario probar el caso $n = 2$. Supongamos que $x_1 \perp x_2$. Entonces

$$\|x_1+x_2\|^2 = \langle x_1+x_2, x_1+x_2\rangle = \langle x_1, x_1\rangle + \langle x_2, x_2\rangle + \langle x_1, x_2\rangle + \langle x_2, x_1\rangle =$$

$$\langle x_1, x_1\rangle + \langle x_2, x_2\rangle = \|x_1\|^2 + \|x_2\|^2.$$

□

PROPOSICIÓN 9.8 (Ley del Paralelogramo). *Sea H un espacio de Hilbert, $x, y \in H$. Entonces*

$$\|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$$

DEMOSTRACIÓN. Dados x, y se sigue de la identidad de polarización que

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\Re\langle x, y\rangle$$

y

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\Re\langle x, y\rangle.$$

Ahora se suman las dos ecuaciones. □

La Ley del Paralelogramo caracteriza de hecho a los espacios de Hilbert, y proponemos esto como ejercicio.

Los dos siguientes resultados son fundamentales en la Teoría de espacios de Hilbert, y son los que permiten que trabajar en un Hilbert sea cómodo, en el sentido de que nuestra intuición geométrica finito dimensional se mantiene razonablemente intacta.

TEOREMA 9.9. *Sea H un espacio de Hilbert, $K \subset H$ un conjunto cerrado convexo no vacío y $x \in H$. Entonces existe un único $k_0 \in K$ tal que*

$$\|x - k_0\| = dist(x, K) := \inf_{k \in K} \|x - k\|.$$

DEMOSTRACIÓN. Supongamos inicialmente $x = 0$. Buscamos probar la existencia y unicidad de un $k_0 \in K$ tal que

$$\|k_0\| = dist(0, K) = \inf_{k \in K} \|k\|.$$

Sea $d = dist(0, K)$. Por definición existe una sucesión $(k_n) \subset K$ tal que $\|k_n\| \to d$. De la Ley del Paralelogramo se sigue que

$$\left\|\frac{k_n - k_m}{2}\right\|^2 = \frac{1}{2}(\|k_n\|^2 + \|k_m\|^2) - \left\|\frac{k_n + k_m}{2}\right\|^2. \tag{21}$$

Puesto que K es convexo, $\frac{k_n+k_m}{2} \in K$ y por tanto

$$\left\|\frac{k_n + k_m}{2}\right\|^2 \geq d^2.$$

Sea $\epsilon > 0$. Entonces existe $N \in \mathbb{N}$ tal que para todo $n \geq N$ $\|k_n\|^2 < d^2 + \frac{1}{4}\epsilon^2$. Por tanto, para todos $m, n \geq N$ se sigue de (21) que

$$\left\|\frac{k_n - k_m}{2}\right\|^2 < \frac{1}{2}(2d^2 + \frac{1}{2}\epsilon^2) - d^2 = \frac{1}{4}\epsilon^2.$$

Por tanto $(k_n)_n$ es una sucesión de Cauchy y puesto que H es completo y K es cerrado sabemos que existe $k_0 \in K$ tal que $k_n \to k_0$. Además

$$d \leq \|k_0\| = \|k_0 - k_n + k_n\| \leq \|k_0 - k_n\| + \|k_n\| \to d$$

y por tanto

$$\|k_0\| = d.$$

Ya hemos probado la existencia de k_0, veamos la unicidad. Supongamos que existe $k_1 \in K$ tal que $\|k_1\| = d$. Por convexidad $\frac{k_0+k_1}{2} \in K$ y por tanto

$$d \leq \left\| \frac{k_0 + k_1}{2} \right\| \leq \frac{1}{2}(\|k_0\| + \|k_1\|) = d$$

de forma que

$$d = \left\| \frac{k_0 + k_1}{2} \right\|.$$

Pero la Ley del Paralelogramo implica que

$$d^2 = \left\| \frac{k_0 + k_1}{2} \right\|^2 = d^2 - \left\| \frac{k_0 - k_1}{2} \right\|^2$$

de donde $k_0 = k_1$.

Si $x \neq 0$, basta aplicar lo anterior a $0 = x - x$ y el cerrado convexo no vacío $x - K$ para obtener un único $x - k_0$ que cumple las condiciones pedidas. Entonces es un ejercicio ver que k_0 es el único elemento de K en el que se alcanza la distancia. □

Los subespacios son siempre trivialmente convexos, por lo que podemos aplicar el teorema anterior al caso de un subespacio cerrado. Pero en ese caso podemos decir aún más.

Teorema 9.10 (Proyección ortogonal). *Sea H un espacio de Hilbert y $M \subset H$ un subespacio cerrado. Sea $x \in H$ y sea $y_0 \in M$ el único elemento de M tal que $\|x - y_0\| = dist(x, M)$. Entonces*

$$x - y_0 \perp M.$$

Recíprocamente, si $y_0 \in M$ es tal que $x - y_0 \perp M$ entonces

$$\|x - y_0\| = dist(x, M).$$

Demostración. Hacemos la demostración en el caso complejo. El caso real es esencialmente igual aunque algo más sencillo y queda como ejercicio para los alumnos. Supongamos que $y_0 \in M$ verifica $\|x - y_0\| = dist(x, M)$. Si $y \in M$, entonces $y + y_0 \in M$ y por tanto, usando la identidad de polarización,

$$\|x - y_0\|^2 \leq \|x - (y_0 + y)\|^2 = \|(x - y_0) - y\|^2 =$$
$$= \|x - y_0\|^2 - 2\Re\langle x - y_0, y\rangle + \|y\|^2.$$

Por lo tanto

$$2\Re\langle x - y_0, y\rangle \leq \|y\|^2 \tag{22}$$

para todo $y \in M$.

Queremos probar que $\langle x - y_0, y\rangle = 0$. Fijemos $y \in M$ y llamemos $\langle x - y_0, y\rangle = re^{\imath\theta}$, con $r \geq 0$, $\theta \in [0, 2\pi)$. Consideremos para cada $t \in \mathbb{R}_+$ el vector $te^{\imath\theta}y \in M$. La ecuación (22) aplicada a este vector nos dice que

$$2\Re(te^{-\imath\theta}re^{\imath\theta}) \leq t^2\|y\|^2$$

es decir

$$2tr \leq t^2\|y\|^2.$$

Puesto que esto ocurre para todo $t \in \mathbb{R}_+$ se sigue que $r = 0$, de donde

$$\langle x - y_0, y\rangle = 0$$

como queríamos probar.

Recíprocamente, sea $y_0 \in M$ tal que $x - y_0 \perp M$. Entonces, para todo $y \in M$, también $y - y_0 \in M$ y por tanto

$$x - y_0 \perp y - y_0$$

y de ahí se sigue, por el Teorema de Pitágoras, que

$$\|x - y\|^2 = \|(x - y_0) + (y_0 - y)\|^2 =$$

$$= \|x - y_0\|^2 + \|y_0 - y\|^2 \geq \|x - y_0\|^2$$

y por tanto

$$\|x - y_0\| = dist(y, M).$$

□

El Teorema anterior no es el fin de la historia. Vemos que dado un subespacio vectorial M de H existe una aplicación bien definida

$$P : H \longrightarrow M$$

que lleva x a y_0, donde y_0 es el único elemento de M tal que $\|x - y_0\| = dist(x, M)$ y simultáneamente es el único elemento de M tal que

$$x - y_0 \perp M.$$

Por supuesto, dado que estamos trabajando con espacios normados, nos gustaría que esta aplicación fuera lineal y continua. Este es el contenido principal del siguiente teorema, que en realidad es una continuación del anterior. Necesitamos previamente un concepto nuevo.

DEFINICIÓN 9.11. *Sea $A \subset H$ un subconjunto. Se define el ortogonal de A, y lo llamamos $A^{\perp}$, como*
$$A^{\perp} := \{x \in H \text{ tales que } x \perp a \text{ para todo } a \in A\}.$$

Más adelante, cuando veamos el Teorema de Representación de Riesz veremos que esta notación es coherente con el $A^{\perp} \subset H^*$ definido anteriormente.

TEOREMA 9.12 (Proyección ortogonal). *Sea H un espacio de Hilbert y $M \subset H$ un subespacio vectorial cerrado. Sea $P : H \longrightarrow M \hookrightarrow H$ la aplicación definida en el párrafo anterior. Entonces*
1. *P es lineal y continua, y $\|P\| = 1$ si $M \neq \{0\}$.*
2. *$P^2 = P$ (pensando P como aplicación de H en H, para que P^2 tenga sentido).*
3. *$\ker P = M^{\perp}$ e $Im(P) = M$*

DEMOSTRACIÓN. 1. Sean $x_1, x_2 \in H$, $\alpha, \beta \in \mathbb{K}$. Entonces, para todo $y \in M$ se tiene que
$$\langle \alpha x_1 + \beta x_2 - (\alpha P(x_1) + \beta P(x_2)), y\rangle =$$
$$= \alpha\langle x_1 - P(x_1), y\rangle + \beta\langle x_2 - P(x_2), y\rangle = 0 + 0 = 0.$$
Puesto que $P(\alpha x_1 + \beta x_2)$ es el único elemento de M que verifica la condición de ortogonalidad, se tiene que
$$P(\alpha x_1 + \beta x_2) = \alpha P(x_1) + \beta P(x_2).$$

Para ver que P es continua y de norma 1, sea $x \in H$. Entonces
$$x = (x - P(x)) + P(x)$$
donde $P(x) \in M$ y $x - P(x) \in M^{\perp}$. Por tanto $P(x) \perp (x - P(x))$ y del Teorema de Pitágoras se sigue que
$$\|x\|^2 = \|x - P(x)\|^2 + \|P(x)\|^2 \geq \|P(x)\|^2,$$
es decir
$$\|P(x)\| \leq \|x\|$$
y por tanto $\|P\| \leq 1$. Por otro lado, considerando $0 \neq x \in M$, $P(x) = x$ (esto es trivial) y por tanto $\|P\| \geq 1$.

2. Como $P(y) = y$ para todo $y \in M$, tomando $x \in H$ tenemos que $P(x) \in M$ y por tanto
$$P^2(x) = P(P(x)) = P(x)$$
lo que prueba lo pedido.

3. Si $P(x) = 0$ entonces $x = x - P(x) \in M^\perp$. Recíprocamente, si $x \in M^\perp$, sea $P(x) = y \in M$ el único vector de M tal que $x - y \in M^\perp$. Esto implica en particular que

$$0 = \langle x - y, y\rangle = \langle x, y\rangle - \langle y, y\rangle = 0 - \|y\|^2$$

de donde $y = 0$.

Para ver que $Im(P) = M$, puesto que $Im(P) \subset M$, sólo es necesario observar que $P(y) = y$ para todo $y \in M$. □

Definición 9.13. *Si $M \subset H$ es un subespacio vectorial cerrado, a la proyección P anteriormente se le denomina Proyección Ortogonal sobre M.*

Corolario 9.14. *Sea H un espacio de Hilbert, $M \subset H$ un subespacio vectorial. Entonces M está complementado en H. De hecho un complementario de M en H es precisamente $M^\perp$.*

El Teorema de la Proyección Ortogonal es uno de los resultados más utilizados en espacios de Hilbert, y uno de los motivos de que dichos espacios sean mucho más "cómodos" que otros espacios de Banach a la hora de trabajar con ellos. De hecho, el último corolario enunciado caracteriza a los Hilbert, en el sentido de que cualquier espacio de Banach X en el que todo subespacio cerrado esté complementado es isomorfo a un espacio de Hilbert [**32**]

Podemos obtener como fácil corolario una versión débil en espacios de Hilbert del Teorema Bipolar

Corolario 9.15. *Si $M \subset H$ es un subespacio vectorial cerrado, entonces $(M^\perp)^\perp = M$.*

Demostración. Si P es la proyección ortogonal sobre M, entonces es fácil ver que $I - P$ es la proyección ortogonal sobre $M^\perp$. Por tanto, del apartado 3 del Teorema de la Proyección ortogonal se sigue que

$$(M^\perp)^\perp = \ker I - P.$$

Pero $x \in \ker I - P$ si y sólo si $P(x) = x$, es decir, $x \in M$. □

Aplicación: Como aplicación de la proyección ortogonal podemos mostrar una construcción "elemental" de la esperanza condicional.

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad (es decir un espacio Ω con una σ-álgebra $\mathcal{F}$ definida en él y una probabilidad $P : \mathcal{F} \longrightarrow [0, 1]$). Supongamos para fijar ideas que $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}$, la σ-álgebra de Borel y $P = \lambda$ es la medida de Lebesgue.

Para todo $1 \leq p \leq \infty$ sea $L_p([0,1], \mathcal{B}, \lambda)$ el espacio de Lebesgue habitual.

Consideremos una sub-σ-álgebra $\mathcal{G} \subset \mathcal{B}$. Tiene entonces sentido hablar de $L_p([0,1], \mathcal{G}, \lambda)$, como el espacio de las funciones $\mathcal{G}$-medibles con $\|\cdot\|_p$ finita. Es fácil ver que $L_p([0,1], \mathcal{G}, \lambda)$ es un subespacio cerrado de $L_p([0,1], \mathcal{B}, \lambda)$. Se puede ver que de hecho $L_p([0,1], \mathcal{G}, \lambda)$ está complementado en $L_p([0,1], \mathcal{B}, \lambda)$, y que podemos una definir una proyección

$$\mathbf{E}(\cdot|\mathcal{G}) : L_p([0,1], \mathcal{B}, \lambda) \longrightarrow L_p([0,1], \mathcal{G}, \lambda)$$

que denominaremos *esperanza condicional* con propiedades especiales que lo hacen muy importante en la Teoría de Probabilidad y Estadística. Las propiedades interesantes de la esperanza condicional son consecuencia de ser una proyección que verifica que para todo $G \in \mathcal{G}$ y para todo $f \in L_p([0,1], \mathcal{B}, \lambda)$

$$\int_G \mathbf{E}(f|\mathcal{G}) = \int_G f d\lambda \tag{23}$$

La interpretación probabilística que se suele hacer de esto es la siguiente. La σ-álgebra $\mathcal{G}$ (o más bien las funciones $\mathcal{G}$-medibles) representa "lo que sabemos", o más correctamente los *observables* en unas circunstancias dadas, o un instante dado. Entonces $\mathbf{E}(f|\mathcal{G})$ representa "lo mejor que podemos decir" acerca de una función f con los conocimientos representados en $\mathcal{G}$.

La construcción de la Esperanza Condicional se hace habitualmente utilizando el Teorema de Radon-Nikodym. Vamos a ver una construcción "elemental".

Consideramos primeramente el caso $p = 2$. En ese caso, es un ejercicio ver que la proyección ortogonal es una proyección que verifica (23).

Recordemos que si $p \geq q$, $L_p[0,1] \subset L_q[0,1]$. Por tanto para todo $p > 2$ podemos definir el operador esperanza condicional simplemente restringiendo la esperanza condicional en L_2. De nuevo es fácil ver que el operador así definido termina en $L_p([0,1], \mathcal{G}, \lambda)$ y verifica (23).

El problema viene cuando tratamos de extender la esperanza condicional a todo L_1. No incluimos todos los detalles, que requieren de un uso más o menos cuidadoso de técnicas estandar de Teoría de la Medida. El lector interesado puede consultar por ejemplo [**35**].

Quizás sea interesante mencionar que siguiendo en esta dirección se puede, a partir de esta definición de esperanza condicional y utilizando un teorema de convergencia de martingalas, demostrar el Teorema de

Radon-Nikodym, cerrando así el círculo de ideas. Los detalles se pueden ver por ejemplo en [**28**].

Conjuntos ortogonales

Buscamos ahora ver la versión infinito-dimensional de la noción, ya conocida por los alumnos en dimensión finita, de base ortonormal de un espacio de Hilbert. Puesto que previsiblemente no trataremos bases de Schauder en el programa de la asignatura, damos aquí una presentación autocontenida. Si hubiéramos visto bases de Schauder en espacios de Banach, entonces relacionaríamos esta sección con aquellas.

DEFINICIÓN 9.16. *Un subconjunto $E \subset H$ se dice ortogonal si para todo $x, y \in E$, $x \neq y$, se tiene que*

$$x \perp y.$$

Si además $\|x\| = 1$ para todo $x \in X$ decimos que X es ortonormal.

PROPOSICIÓN 9.17. *Sea $E \subset H$ un subconjunto ortogonal. Si $0 \notin E$ entonces E es linealmente independiente. Si además E es ortonormal entonces*

$$\|x - y\| = \sqrt{2}$$

para todo $x, y \in E$ con $x \neq y$.

DEMOSTRACIÓN. Sean $x_1, \ldots, x_n \in E$ y sean $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tales que

$$\sum_{i=1}^{n} \alpha_i x_i = 0.$$

Entonces, para todo $1 \leq j \leq n$ se tiene

$$0 = \langle 0, x_j \rangle = \langle \sum_{i=1}^{n} \alpha_i x_i, x_j \rangle = \sum_{i=1}^{n} \alpha_i \langle x_i, x_j \rangle = \alpha_j \langle x_j, x_j \rangle.$$

Por tanto, puesto que $x_j \neq 0$ se tiene que $\alpha_j = 0$ para todo $1 \leq j \leq n$ y E es linealmente independiente. Si E es ortonormal entonces para todo $x \neq y \in E$

$$\|x - y\|^2 = \langle x - y, x - y \rangle = \langle x, x \rangle + \langle y, y \rangle = 2.$$

□

Puesto que los sistemas ortonormales son linealmente independientes, dado un sistema ortonormal podemos, utilizando el Álgebra Lineal, extender dicho sistema a una base de Hamel de H. Sin embargo habitualmente las bases de Hamel no son un objeto de trabajo interesante en la Teoría de Espacios Normados, y en particular en la Teoría de Espacios de Hilbert, puesto que no sacan partido de la estructura topológica y esto las hace ser demasiado grandes: ya vimos en el Ejercicio 4.5 que los espacios de Banach infinito dimensionales no pueden tener bases de Hamel numerables.

En la Teoría de Espacios Normados de dimensión Infinita el objeto interesante son las bases de Schauder, y en la Teoría de Espacios de Hilbert son las *bases ortonormales* que iremos construyendo en los resultados siguientes.

Vamos a empezar viendo el proceso de ortogonalización de Gram-Schmidt que nos dice que, dado un conjunto contable linealmente independiente de H podemos construir inductivamente un conjunto ortonormal que genera el mismo subespacio vectorial en cada paso.

TEOREMA 9.18 (Proceso de Ortonormalización de Gram-Schmidt). *Sea H un espacio de Hilbert y $\{x_n; n \in \mathbb{N}\} \subset H$ un subconjunto linealmente independiente. Entonces existe un conjunto ortonormal $\{e_n; n \in \mathbb{N}\}$ tal que, para todo $n \in \mathbb{N}$,*

$$span[e_1, \dots, e_n] = span[x_1, \dots, x_n]$$

DEMOSTRACIÓN. Definimos inductivamente la sucesión $(e_n)_n$. En realidad vamos definiendo una sucesión ortogonal (y_n) y a continuación la normalizamos. En primer lugar,

$$y_1 = x_1$$

y

$$e_1 = \frac{y_1}{\|y_1\|}.$$

Obviamente $span[e_1] = span[x_1]$

Supongamos que ya hemos definido los $n-1$ primeros términos de la sucesión. Definimos entonces

$$y_n = x_n - \langle x_n, e_1\rangle e_1 - \cdots - \langle x_n, e_{n-1}\rangle e_{n-1}$$

y

$$e_n = \frac{y_n}{\|y_n\|}.$$

Todo es como en el caso finito dimensional que ya conocen así que dejamos como ejercicio la conclusión de la demostración. $\square$

EJEMPLO 9.19. 1. *En ℓ_2, la base canónica (e_n) es ortonormal.*

2. *Si en* ℓ_2 *ortonormalizamos la "base sumante"* $(x_n)_n$ *con* $x_n = (1, \overset{(n)}{\dots}, 1, 0, 0, \dots)$ *obtenemos la base canónica.*
3. *Si ortonormalizamos la sucesión* $1, x, x^2, \dots,$ *en* $L_2[0,1]$ *obtenemos la sucesión*
$$e_n = \sqrt{\frac{2n+1}{2}} P_n(x)$$
donde los polinomios
$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}(x^2 - 1)$$
son los Polinomios de Legendre.

Enunciamos y probamos a continuación la desigualdad de Bessel en su forma habitual, es decir para conjuntos ortonormales *numerables*. Más adelante vemos como corolario que la desigualdad también es cierta para conjuntos ortonormales arbitrarios.

TEOREMA 9.20 (Desigualdad de Bessel). *Sea* H *un espacio de Hilbert,* $\{e_n; n \in \mathbb{N}\} \subset H$ *un conjunto ortonormal y* $x \in H$. *Entonces*
$$\sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 \leq \|x\|^2.$$
Además se tiene la igualdad si y sólo si
$$x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n.$$

DEMOSTRACIÓN. Fijemos $n \in \mathbb{N}$ y sea $x_n = x - \sum_{k=1}^{n} \langle x, e_k \rangle e_k$. De la ortogonalidad de los (e_n) se sigue que para todo $n \in \mathbb{N}$ y para todo $1 \leq k \leq n$,
$$x_n \perp e_k.$$
Puesto que $x = x_n + \sum_{k=1}^{n} \langle x, e_k \rangle e_k$, del teorema de Pitágoras se tiene que
$$\|x\|^2 = \|x_n\|^2 + \sum_{k=1}^{n} \|\langle x, e_k \rangle e_k\|^2 = \|x_n\|^2 + \sum_{k=1}^{n} |\langle x, e_k \rangle|^2$$
de lo que se sigue que
$$\sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \leq \|x\|^2.$$
Puesto que esto es cierto para un $n \in \mathbb{N}$ arbitrario, se tiene la desigualdad buscada.

Si se tiene la igualdad, esto implica que

$$\|x_n\|^2 \to 0$$

y de aquí se sigue que

$$x = \lim_{n\to\infty} \sum_{k=1}^{n} \langle x, e_k \rangle e_k = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$$

Recíprocamente, si $x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$ entonces $\|x_n\| \to 0$ y se sigue la igualdad buscada. □

COROLARIO 9.21. *Si $E \subset H$ es un conjunto ortonormal y $x \in H$ entonces $\langle x, e \rangle = 0$ excepto en a lo sumo una cantidad numerable de elementos $e \in E$.*

DEMOSTRACIÓN. Para todo $n \in \mathbb{N}$, sea

$$E_n := \{e \in E \text{ tales que } \langle x, e \rangle \geq \frac{1}{n}\}$$

Por la desigualdad de Bessel E_n es finito para cada n, y

$$\cup_{n\in\mathbb{N}} E_n = \{e \in E \text{ tales que } \langle x, e \rangle \neq 0\}.$$

□

COROLARIO 9.22 (Desigualdad de Bessel). *Sea H un espacio de Hilbert, $E \subset H$ un conjunto ortonormal y $x \in H$. Entonces*

$$\sum_{e\in E} |\langle x, e \rangle|^2 \leq \|x\|^2.$$

Estudiamos ahora la convergencia de las series $\sum_n k_n e_n$ donde los k_n son escalares y $\{e_n; n \in \mathbb{N}\}$ forman un sistema ortonormal.

TEOREMA 9.23 (Riesz-Fischer). *Sea H un espacio de Hilbert, sea $\{e_n; n \in \mathbb{N}\} \subset H$ un sistema ortonormal y sea $(k_n)_{n\in\mathbb{N}} \subset \mathbb{K}$. Si $\sum_n k_n e_n$ converge a $x \in H$ entonces, para todo $n \in \mathbb{N}$,*

$$\langle x, e_n \rangle = k_n$$

y

$$\sum_n |k_n|^2 < \infty.$$

Recíprocamente, si $\sum_n |k_n|^2 < \infty$ entonces

$$\sum_n k_n e_n$$

converge en H.

Demostración. Si $x = \sum_n k_n e_n$ entonces de la ortogonalidad de los (e_n) se sigue que $\langle x, e_n\rangle = k_n$ para todo $n \in \mathbb{N}$ y la desigualdad de Bessel implica que

$$\sum_n |k_n|^2 < \infty.$$

Recíprocamente, para todo $m \in \mathbb{N}$ sea

$$x_m = \sum_{n=1}^{m} k_n e_n.$$

Entonces para todo $1 \leq j \leq m$ se tiene que

$$x_m - x_j = \sum_{n=j+1}^{m} k_n e_n$$

y de nuevo la ortonormalidad de los (e_n) implica que

$$\|x_m - x_j\|^2 = \langle x_m - x_j, x_m - x_j\rangle = \sum_{n=j+1}^{m} |k_n|^2.$$

Por tanto, si $\sum_{n=1}^{\infty} |k_n|^2 < \infty$ se sigue que la sucesión $(x_m)_m$ es de Cauchy y por tanto convergente. □

Con la desigualdad de Bessel y el Teorema de Riesz-Fischer estamos ya bastante cerca de la noción de *base Hilbertiana* o *base ortonormal.*

Definición 9.24. *Un sistema ortonormal $E \subset H$ es una* base ortonormal*, o* base Hilbertiana*, si es un sistema ortonormal maximal para el orden dado por la inclusión. Es decir, si existe un sistema ortonormal E' tal que $E \subset E'$ entonces $E = E'$.*

Podemos garantizar la existencia de bases ortonormales mediante el Lema de Zorn (3.3).

Proposición 9.25. *Sea $H \neq \{0\}$ un espacio de Hilbert. Entonces existe una base ortonormal en H. De hecho, dado un sistema ortonormal $E' \subset H$, E' se puede ampliar a una base ortonormal E.*

Demostración. Notemos en primer lugar que si $0 \neq x \in H$, el sistema $\{\frac{x}{\|x\|}\}$ es trivialmente ortonormal. Veamos que un sistema ortonormal E' se puede extender a una base ortonormal. Sea $\mathcal{E}$ el conjunto de todos los sistemas ortonormales de H que contienen a E'. Este es un conjunto dirigido por la inclusión conjuntista. Si consideramos un subconjunto totalmente ordenado (una cadena) es fácil ver que la unión de todos los elementos de la cadena es una cota superior de esta. Por tanto el Lema de Zorn nos dice que existe un elemento maximal de $\mathcal{E}$ al que llamamos E. Claramente $E' \subset E$ y E es base ortonormal. $\square$

Necesitaremos ahora un lema que nos permite reducir las sumas que aparecen en los sistemas ortonormales a series.

Lema 9.26. *Sea $E = \{e_\alpha; \alpha \in A\} \subset H$ un sistema ortonormal y $x \in H$. Sea*

$$E_x := \{e_\alpha \in E \text{ tales que } \langle x, e_\alpha\rangle \neq 0\}.$$

Entonces E_x es un conjunto numerable y llamando $E_x = \{e_n; n \in \mathbb{N}\}$ se tiene

$$\lim_n \langle x, e_n\rangle = 0.$$

Además

$$\sum_n \langle x, e_n\rangle e_n$$

converge a un elemento $y \in H$ tal que

$$x - y \perp e_\alpha \text{ para todo } \alpha \in A.$$

Demostración. Si $x = 0$ no hay nada que demostrar. Sea entonces $x \neq 0$. Que E_x es contable es el Corolario 9.21. Llamando $E_x = \{e_n; n \in \mathbb{N}\}$ de la desigualdad de Bessel se sigue que

$$\sum_n |\langle x, e_n\rangle|^2 \leq \|x\|^2 < \infty$$

y por tanto $|\langle x, e_n\rangle|^2 \to 0$.

Puesto que ya hemos visto que $\sum_n |\langle x, e_n\rangle|^2 < \infty$, del Teorema de Riesz-Fischer (9.23) se sigue que

$$\sum_n \langle x, e_n\rangle e_n$$

converge a un elemento $y \in H$. Además, para todo $\alpha \in A$,

$$\langle y, e_\alpha\rangle = \left\langle \sum_n \langle x, e_n\rangle e_n, e_\alpha \right\rangle = \sum_n \langle x, e_n\rangle \langle e_n, e_\alpha\rangle = \langle x, e_\alpha\rangle$$

y de ahí se sigue fácilmente que

$$x - y \perp e_\alpha \text{ para todo } \alpha \in A.$$

$\square$

Teorema 9.27. *Sea $E = \{e_\alpha; \alpha \in A\} \subset H$ un sistema ortonormal. Entonces son equivalentes*

(1) *E es una base ortonormal de H*

(2) **(Desarrollo en serie de Fourier)** *Para todo* $x \in H$

$$x = \sum_{\alpha \in A} \langle x, e_\alpha \rangle e_\alpha$$

entendiendo este sumatorio como una serie, puesto que a lo sumo una cantidad numerable de términos son no nulos, es decir

$$x = \sum_{n} \langle x, e_n \rangle e_n$$

donde $\{e_n; n \in \mathbb{N}\} = \{e_\alpha \in E$ *tales que* $\langle x, e_\alpha \rangle \neq 0\}$.

(3) **(Fórmula de Parseval)** *Para todo* $x \in H$

$$\|x\|^2 = \sum_{\alpha \in A} |\langle x, e_\alpha \rangle|^2$$

entendiendo este sumatorio igual que en el apartado anterior.

(4) $span[E]$ *es denso en* H

(5) *Si* $x \in H$ *es tal que* $\langle x, e_\alpha \rangle = 0$ *para todo* $\alpha \in A$ *entonces* $x = 0$.

Demostración. Supongamos que (1) es cierto, es decir E es una base ortonormal, es decir un sistema ortonormal maximal. Sea $x \in H$. Por el Lema 9.26 existe $y \in H$ tal que

$$\sum_{\alpha} \langle x, e_\alpha \rangle e_\alpha = y$$

y además

$$x - y \perp E.$$

Si $y \neq x$, sea

$$e = \frac{x - y}{\|x - y\|}.$$

Entonces $\|e\| = 1$ y $e \perp E$ por lo que $E \cup \{e\}$ es un sistema ortonormal, en contradicción con que E sea maximal. Por tanto

$$x = y = \sum_{\alpha} \langle x, e_\alpha \rangle e_\alpha$$

y (2) es cierto.

La equivalencia de (2) y (3) se sigue de la condición descrita para la igualdad en la desigualdad de Bessel (Teorema 9.20).

Fácilmente (2) implica (4), ya que $\sum_{n=1}^{m} \langle x, e_n \rangle e_n \in span[E]$.

Para ver que (4) implica (5), sea $x \in H$ tal que $\langle x, e_\alpha \rangle = 0$ para todo $\alpha \in A$, y sea $(x_n)_{n\in\mathbb{N}} \subset span[E]$ una sucesión tal que $x_n \to x$. Entonces $\langle x_n, x\rangle = 0$ para todo n y por tanto

$$0 = \langle x_n, x\rangle \to \langle x, x\rangle$$

de donde se sigue que $x = 0$

Finalmente, para ver que (5) implica (1), sea $E' \subset H$ un sistema ortonormal tal que $E \subset E'$. Sea $e \in E' \setminus E$. De la ortogonalidad de E' se sigue que $\langle e, e_\alpha\rangle = 0$ para todo $\alpha \in A$. De (5) se sigue que $e = 0$, una contradicción puesto que si E' es ortonormal y por tanto $\|e\| = 1$. □

Podemos ahora definir un concepto de *dimensión Hilbertiana*, de forma análoga a la definición de la dimensión algebraica.

PROPOSICIÓN 9.28. *Si H es un espacio de Hilbert, todas sus bases ortonormales tienen el mismo cardinal.*

DEMOSTRACIÓN. Sean E y F dos bases ortonormales de H, sean $\epsilon = card(E)$, $\eta = card(F)$. Si ϵ o η son finitos, entonces $\epsilon = \eta$, usando razonamientos totalmente análogos a los finito-dimensionales. Supongamos entonces que ambos son infinitos. Dado $e \in E$, sea

$$F_e := \{f \in F \text{ tales que } \langle e, f\rangle \neq 0\}.$$

Ya hemos visto que F_e es contable. Además todo $f \in F$ debe de pertenecer al menos a uno de los F_e, puesto que si no $\langle e, f\rangle = 0$ para todo $e \in E$ y de la condición (5) del Teorema 9.27 se seguiría que $f = 0$. Por tanto

$$F = \cup_{e\in E} F_e$$

y por tanto

$$\eta \leq \epsilon\aleph_0 = \epsilon$$

Análogamente se puede ver que

$$\epsilon \leq \eta\aleph_0 = \eta$$

de donde se sigue que $\epsilon = \eta$. □

Ya podemos definir la dimensión hilbertiana.

DEFINICIÓN 9.29. *Sea H un espacio de Hilbert. Se define su dimensión hilbertiana como el cardinal de una cualquiera de sus bases ortonormales.*

PROPOSICIÓN 9.30. *Sea H un espacio de Hilbert. Entonces H es separable si y sólo si* $\dim H = \aleph_0$.

DEMOSTRACIÓN. Si $\dim H = \aleph_0$, y $E = \{e_n; n \in \mathbb{N}\}$, es un ejercicio ver que el conjunto numerable

$$\left\{\sum_{n=1}^{m} q_n e_n; q_n \in \mathbb{Q} + \imath\mathbb{Q}, m \in \mathbb{N}\right\}$$

es denso en H.

Recíprocamente, sea E una base ortonormal de H. Ya vimos que para todo $e_1, e_2 \in E$

$$\|e_1 - e_2\|^2 = \|e_1\|^2 + \|e_2\|^2 = 2.$$

Por tanto, el conjunto

$$\{B_{e,\frac{1}{\sqrt{2}}}; e \in E\}$$

es un conjunto de bolas abiertas de H disjuntas dos a dos. De aquí se sigue que este conjunto, y por tanto E, deben de ser numerables. El razonamiento es el siguiente:

Si (X, d) es un espacio métrico separable y $\{B_i = B_{x_i,r_i}; i \in I\}$ es un conjunto de bolas abiertas disjuntas dos a dos, entonces I tiene que ser contable. En efecto, sea $D \subset X$ un subconjunto contable denso. Entonces para todo $i \in I$,

$$B_i \cap D \neq \emptyset$$

y por tanto existe $y_i \in B_i \cap D$. Por tanto $\{y_i; i \in I\}$ es un subconjunto de D con el cardinal de I. De aquí se sigue que

$$cardI \leq cardD = \aleph_0.$$

$\square$

EJEMPLO 9.31. *Sea $H = \ell_2$ y sea $E = \{e_n; n \in \mathbb{N}\} \subset \ell_2$ la base canónica. X es un sistema ortonormal, y dado $x \in \ell_2$, si*

$$\langle x, e_n \rangle = 0$$

para todo $n \in \mathbb{N}$ se sigue que $x = 0$. Por tanto E es una base ortonormal, y queda justificado el nombre de base canónica, y tenemos que ℓ_2 es separable.

En la categoría de espacios de Hilbert, el concepto natural de isomorfismo es aquel que preserva el producto escalar junto con la estructura lineal. Es decir, una aplicación lineal que preserve productos escalares. Veamos que los isomorfismos en este sentido son exactamente las isometrías.

PROPOSICIÓN 9.32. *Sean H, K espacios de Hilbert y $T : H \longrightarrow K$ una aplicación lineal. Entonces T es una isometría si y sólo si para todo $x, y \in H$*

$$\langle T(x), T(y)\rangle = \langle x, y\rangle.$$

DEMOSTRACIÓN. Si $\langle T(x), T(y)\rangle = \langle x, y\rangle$ entonces

$$\|T(x)\|^2 = \langle T(x), T(x)\rangle = \langle x, x\rangle = \|x\|^2$$

y T es una isometría.

Recíprocamente, si T es una isometría, se sigue de la identidad de polarización que se conserva el producto escalar. □

Decimos que dos espacios de Hilbert H, K son isomorfos como espacios de Hilbert si son isométricos como espacios de Banach, es decir, si existe una isometría $T : H \longrightarrow K$ sobreyectiva.

TEOREMA 9.33. *Dos espacios de Hilbert H, K son isomorfos como espacios de Hilbert si y sólo si tienen la misma dimensión.*

DEMOSTRACIÓN. Si $T : H \longrightarrow K$ es una isometría sobre y E es una base de H, entonces $T(E) = \{T(e); e \in E\}$ es un sistema ortonormal de K. Además, dado $y \in K$, sea $x \in H$ el único elemento de H tal que $T(x) = y$. Si $\langle T(e), y\rangle = 0$ para todo $T(e) \in T(E)$ tenemos que $\langle T(e), T(x)\rangle = \langle e, x\rangle = 0$ para todo $e \in E$. Puesto que E es una base ortonormal, esto implica que $x = 0$ y por tanto $y = 0$ y se tiene que $T(E)$ es una base ortonormal por el Teorema 9.27. De ahí se sigue que $\dim H = \dim K$.

Recíprocamente, sea H es un espacio de Hilbert y $E \subset H$ una base ortonormal. Consideramos el espacio

$$\ell_2(E) = \{(a_e)_{e\in E} \text{ tales que } \sum_e |a_e|^2 < \infty\}$$

De manera totalmente análoga a como vimos que ℓ_2 es un espacio de Hilbert, se comprueba que $\ell_2(E)$ es un espacio de Hilbert con el producto escalar obvio. (En algún momento de la demostración hay que observar, como ya hemos hecho anteriormente en más de una ocasión, que si $\sum_e |a_e|^2 < \infty$ entonces a lo sumo una cantidad numerable de "coordenadas" a_e son distintas de 0.)

Definimos entonces la aplicación

$$T : H \longrightarrow \ell_2(E)$$

como

$$T(x) = (\langle x, e\rangle)_{e\in E}$$

La linealidad de T está clara. De la desigualdad de Bessel, o del Teorema de Parseval, se sigue que T está bien definido y que es una isometría. Para ver que es sobre, si $(a_e)_{e\in E} \in \ell_2(E)$, entonces del Teorema de de Riez-Fischer (9.23) se sigue que la serie

$$\sum_{e\in E} a_e e$$

(serie porque sólo una cantidad numerable de términos son no nulos) converge en H a un elemento x. Es claro entonces que $T(x) = (a_e)$.

Luego hemos probado que todo espacio de Hilbert es isomorfo a $\ell_2(I)$ donde I es un conjunto tal que $card(I) = \dim H$, y de ahí se sigue lo pedido. □

OBSERVACIÓN 9.34. *Nótese que hemos probado algo más que lo anunciado; hemos probado que de hecho todo espacio de Hilbert es (isomorfo a) un espacio $\ell_2(I)$, y que tiene sentido hablar de las coordenadas de un vector de forma al menos similar a las coordenadas de los espacios vectoriales finito dimensionales.*

COROLARIO 9.35. *Dos espacios de Hilbert separables son isomorfos (e isomorfos a ℓ_2)*

Ahora podemos estudiar al Análisis de Fourier dentro del marco teórico que hemos preparado.

Tenemos el siguiente teorema, que se puede probar directamente desde la Variable Compleja, aunque nosotros lo probamos aquí como consecuencia del Teorema de Stone-Weierstrass.

TEOREMA 9.36. *Sea $\mathbb{T} = \{z \in \mathbb{C}$ tales que $|z| = 1\}$ y sea*

$$f : \mathbb{T} \longrightarrow \mathbb{C}$$

una función continua. Entonces existe una sucesión $(P_n(z, \overline{z}))_{n\in\mathbb{N}}$ de polinomios en z y $\overline{z}$ tales que $P_n(z, \overline{z})$ tiende a f uniformemente en $\mathbb{T}$.

DEMOSTRACIÓN. $\mathbb{T}$ es un espacio compacto. Consideramos el espacio de Banach $C(\mathbb{T})$ que además tiene estructura de álgebra. Sea $A \subset C(\mathbb{T})$ el álgebra de los polinomios en z y $\overline{z}$. Es un ejercicio verificar que A cumple las hipótesis del Teorema de Stone-Weierstrass y por tanto $\overline{A} = C(\mathbb{T})$. □

Observemos que para todo $z \in \mathbb{T}$, $\overline{z} = z^{-1}$ y $z = e^{\imath\theta}$ y por tanto todo polinomio $P(z, \overline{z})$ se puede escribir como una función de $z = e^{\imath\theta}$ dada por

$$P(z) = \sum_{k=-n}^{n} \alpha_k e^{\imath k\theta}$$

Por este motivo a estos polinomios se les llama *polinomios trigonométricos*. Veamos ahora que el conjunto

$$\left\{ \frac{e^{\imath n\theta}}{\sqrt{2\pi}}; n \in \mathbb{Z} \right\} \subset L_2[0, 2\pi]$$

es una base ortonormal de $L_2[0, 2\pi]$.

Es un ejercicio elemental de variable compleja verificar que dicho sistema es ortonormal.

Para verificar que es una base hay al menos dos posibilidades, ambas no triviales. Una es seguir por ejemplo [**13**, Theorem 5.7] y probarlo utilizando que $C[0, 2\pi]$ es denso en $L_2[0, 2\pi]$ y a continuación utilizar la caracterización, no probada pero fácilmente probable como ejercicio que dice que un sistema ortonormal $E \subset H$ es una base ortonormal si y sólo si la clausura topológica de $span[E] = H$. La otra opción es seguir [**31**, p. 392], pero para esto es necesario utilizar el Teorema de Fejer, lo cual quizá nos lleve demasiado lejos.

Una vez que sabemos que $\left\{ \frac{e^{\imath n\theta}}{\sqrt{2\pi}}; n \in \mathbb{Z} \right\}$ es una base ortonormal podemos asignar a cada función $f \in L_2[0, 2\pi]$ sus coordenadas respecto de dicha base.

Definición 9.37. *Sea $f \in L_2[0, 2\pi]$. Definimos su n-simo coeficiente de Fourier como*

$$\hat{f}(n) = \langle f, \frac{e^{\imath n\theta}}{\sqrt{2\pi}} \rangle = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(t) e^{-\imath nt} dt.$$

Por los resultados anteriores, lo que hemos probado es que

$$f = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{+\infty} \hat{f}(n) e^{\imath n\theta},$$

es decir que toda función $f \in L_2[0, 2\pi]$ admite un desarrollo en serie de Fourier que converge a f en norma 2. Queda probado como corolario también la Identidad de Parseval clásica. La cuestión de la convergencia en casi todo punto de la serie de Fourier de f a f es más fina y se mete más en el Análisis Armónico. Es un resultado profundo que si $1 < p \leq \infty$ y $f \in L_p[0, 2\pi]$ entonces el desarrollo en serie de Fourier de f (que existe) converge a f en casi todo punto. Ya hemos visto que existen funciones $f \in L_1[0, 2\pi]$ tal que su desarrollo en serie de Fourier en casi todo punto no converge a f.

Queda también probada como corolario de los resultados anteriores la versión en L_2 del lema de Riemann-Lebesgue: Si $f \in L_2[0, 2\pi]$ entonces sus coeficientes de Fourier convergen a 0.

También como corolario queda probado no sólo que $L_2[0,2\pi]$ es separable, sino que la transformada de Fourier define una isometría sobreyectiva entre $L_2[0,2\pi]$ y ℓ_2.

Teorema de Riesz

El Teorema de Representación de Riesz es otro de los resultados fundamentales de la Teoría de Espacios de Hilbert. Nos dice que los espacios de Hilbert no sólo son reflexivos, sino que de hecho podemos identificar de manera natural H y H^*. Como forma de fijar la intuición aquí, podemos recordar al alumno que ya hemos visto que $\ell_2^* = \ell_2$, y acabamos de ver que todo espacio de Hilbert es un $\ell_2(I)$ para algún conjunto de índices I. Recordemos que hay otro Teorema de Representación de Riesz, con el que no nos debemos confundir, el que representa el dual de $C[0,1]$.

TEOREMA 9.38 (de Representación de Riesz). *Sea H un espacio de Hilbert y $f \in H^*$. Entonces existe un único $y \in H$ tal que*

$$f(x) = \langle x, y \rangle$$

para todo $x \in H$. Además $\|f\| = \|y\|$. De hecho, si $0 \neq z \in H$ es un elemento ortogonal a $\ker f$ entonces

$$y = \frac{\overline{f(z)}z}{\langle z, z \rangle}.$$

DEMOSTRACIÓN. Lo demostraremos utilizando el Teorema de la Proyección Ortogonal, siguiendo [**31**, 24.3] o [**13**, 3.4]. Existe otra demostración posible utilizando bases ortonormales

Si $f = 0$ entonces $y = 0$ cumple lo pedido. Sea entonces $f \neq 0$. Puesto que f es continuo, $M := \ker f$ es un subespacio cerrado de H. Al ser $f \neq 0$, $M \neq H$ y por tanto $M^\perp \neq \{0\}$. Por tanto, existe $0 \neq z \in M^\perp$. Sea $x \in H$. Por ser M un hiperplano, x se puede escribir como

$$x = w + kz,$$

donde $w \in M$ y $k \in \mathbb{K}$.

Entonces

$$\langle x, z \rangle = \langle w, z \rangle + k\langle z, z \rangle = k\langle z, z \rangle$$

de forma que

$$k = \frac{\langle x, z \rangle}{\langle z, z \rangle}$$

y por tanto

$$f(x) = f(w) + kf(z) = kf(z) = \frac{\langle x, z\rangle}{\langle z, z\rangle} f(z) = \left\langle x, \frac{\overline{f(z)}}{\langle z, z\rangle} z \right\rangle$$

y por tanto

$$y = \frac{\overline{f(z)}}{\langle z, z\rangle} z$$

cumple lo pedido.

Para ver la unicidad de y, sea $y' \in H$ otro vector de H tal que

$$f(x) = \langle x, y'\rangle.$$

Entonces considerando $x = y - y'$ tenemos que

$$\langle y - y', y\rangle = f(y - y') = \langle y - y', y'\rangle$$

y por tanto

$$\langle y - y', y - y'\rangle = 0,$$

es decir, $y = y'$. □

El Teorema de Representación de Riesz nos proporciona de hecho una isometría sobre y conjugada-lineal de H en H^* y nos permite dotar a H^* de estructura de espacio de Hilbert.

Teorema 9.39. *Sea H un espacio de Hilbert y sea $\theta : H^* \longrightarrow H$ la aplicación que a todo $f \in H^*$ le asigna el único $y_f \in H$ que verifica que*

$$f(x) = \langle x, y_f\rangle$$

para todo $x \in H$. Entonces θ es una isometría sobreyectiva y conjugada-lineal.

Demostración. Dados $f, g \in H^*$, para todo $x \in H$ se tiene que

$$(f + g)(x) = f(x) + g(x) = \langle x, y_f\rangle + \langle x, y_g\rangle = \langle x, y_f + y_g\rangle$$

y por tanto

$$\theta(f + g) = y_f + y_g = \theta(f) + \theta(g).$$

Análogamente, dados $f \in H^*$ y $k \in \mathbb{K}$ para todo $x \in H$ se tiene que

$$kf(x) = kf(x) = k\langle x, y_f\rangle = \langle x, \overline{k} y_f\rangle$$

y por tanto

$$\theta(kf) = \overline{k} y_f = \overline{k}\theta(f).$$

Por tanto θ es conjugado-lineal. Para ver que es sobre, sea $y \in H$ y sea $f : H \longrightarrow \mathbb{K}$ la función definida como

$$f(x) = \langle x, y\rangle$$

para todo $x \in H$. Entonces $f \in H^*$ y $\theta(f) = y$. Para ver que es una isometría,

$$\|f\| = \sup_{x \in B_H} |f(x)| = \sup_{x \in B_H} |\langle y_f, x\rangle| = \left\langle y_f, \frac{y_f}{\|y_f\|} \right\rangle =$$
$$= \frac{\langle y_f, y_f\rangle}{\|y_f\|} = \frac{\|y_f\|^2}{\|y_f\|} = \|y_f\|.$$

□

Por tanto podemos definir en H^* un producto escalar simplemente transportando el producto escalar de H de manera que H^* también es un espacio de Hilbert.

TEOREMA 9.40. *Con la notación del teorema anterior, sean $f, g \in H^*$. Definimos*

$$\langle f, g\rangle = \langle \theta(f), \theta(g)\rangle.$$

Entonces así definida, $\langle \cdot, \cdot \rangle$ es un producto escalar en H^* tal que $\|f\|^2 = \langle f, f\rangle$ para todo $f \in H^*$ y por tanto H es un espacio de Hilbert.

DEMOSTRACIÓN. Trivial. □

Veamos finalmente que todo espacio de Hilbert es reflexivo.

COROLARIO 9.41. *Todo espacio de Hilbert es reflexivo.*

DEMOSTRACIÓN. Sea $J : H \longrightarrow H^{**}$ la aplicación canónica dada por

$$J(x)(f) = f(x).$$

Sea $\theta : H^* \longrightarrow H$ la isometría sobre descrita más arriba y $\theta' : H^{**} \longrightarrow H^*$ la isometría respectiva construida sobre los duales. Tenemos simplemente que ver que J es sobre. Sea $x^{**} \in H$. Consideramos el elemento $x = \theta(\theta'(x^{**}) \in H$. Veamos que $J(x) = x^{**}$. Para todo $f \in H^*$

$$x^{**}(f) = \langle f, \theta'(x^{**})\rangle = \langle \theta(\theta'(x^{**}), \theta(f)\rangle = \langle x, \theta(f)\rangle = f(x) = J(x)(f)$$

de donde se sigue lo pedido. □

Para terminar este capítulo, podemos ver cómo en el caso particular de espacios de Hilbert tanto el Teorema de Hahn-Banach como la consecuencia del Principio de Acotación Uniforme que dice que débilmente acotado implica acotado se pueden probar de una manera más sencilla, incluso diciendo más que en al caso general. En el camino estudiaremos la convergencia débil de sucesiones en espacios de Hilbert, uno de los problemas que justificaron la introducción de estos espacios por Hilbert. Hemos incluido el resultado de extensión en el cuerpo del texto y los otros resultados en los ejercicios, pero todo ello es susceptible de modificación.

TEOREMA 9.42 (Extensión de Hahn-Banach con unicidad). *Sea H un espacio de Hilbert, $G \subset H$ un subespacio (no necesariamente cerrado) y $g : G \longrightarrow \mathbb{K}$ una forma lineal y continua. Entonces existe una única extensión $f : H \longrightarrow \mathbb{K}$ de g tal que*

$$f_{|G} = g \text{ y } \|f\| = \|g\|.$$

DEMOSTRACIÓN. Sea $M = \overline{G}$. Entonces g se extiende por continuidad de forma única a una función que seguimos llamando $g : M \longrightarrow \mathbb{K}$ que sigue siendo lineal y continua. Esto es un ejercicio. Puesto que M es un espacio de Hilbert, sabemos que existe $y_g \in M$ tal que, para todo $x \in M$

$$g(x) = \langle x, y_g \rangle$$

Sea entonces $f : H \longrightarrow \mathbb{K}$ el elemento de H^* dado por

$$f(x) = \langle x, y - g \rangle \text{ para todo } x \in H$$

Entonces f extiende a g y $\|f\| = \|y_g\| = \|g\|$.

Para probar la unicidad, sea $h \in H^*$ otra extensión de g con la misma norma. Sea y_h el elemento de H representante de h. Entonces

$$\|y_h\| = \|h\| = \|g\| = \|y_g\|$$

y

$$\langle y_g, y_h \rangle = h(y_g) = g(y_g) = \langle y_g, y_g \rangle = \|y_g\|^2$$

Por lo tanto, la Identidad de Polarización nos dice que

$$\|y_g - y_h\|^2 = \|y_g\|^2 - 2\Re\langle y_g, y_h \rangle + \|y_h\|^2 = 2\|y_g\|^2 - 2\Re\langle y_g, y_h \rangle = 0$$

y por tanto $y_g = y_h$ y de ahí $f = h$. $\square$

Aplicaciones.

Teorema Isoperimétrico: Como consecuencia de la Fórmula de Parseval para el sistema trigonométrico real se tiene

TEOREMA 9.43. *De todas las curvas simples, cerradas y diferenciables a trozos de longitud L en el plano, el círculo es la que encierra mayor área.*

Una demostración de este resultado se puede ver en [**22**, p. 178].

Teorema de Müntz: En [**22**, p. 181] se puede ver una demostración del siguiente hecho

TEOREMA 9.44. *Sea $1 \leq n_1 < n_2 < \cdots$ una sucesión de números naturales y sea $A = \{t^{n_1}, t^{n_2}, \ldots\} \subset L_2[0,1]$. Entonces span$[A]$ es denso en L_2 si y sólo si*

$$\sum_{i=1}^{\infty} \frac{1}{n_i} = \infty.$$

Prácticas sugeridas.

EJERCICIO 9.1. *Si queremos enfatizar la relación entre el Análisis Funcional y la Variable compleja podemos proponer el siguiente ejercicio* [**13**, Prop. 1.13, p. 6].

Sea $G \subset \mathbb{C}$ un abierto y sea $L_a^2(G)$ el espacio de las funciones analíticas $f : G \longrightarrow \mathbb{C}$ tales que

$$\int\int_G |f(x+\imath y)|^2 dxdy < \infty$$

Verificar que es un espacio de Hilbert con el producto escalar heredado. La única dificultad reside en verificar que $L_a^2(G)$ es un subespacio cerrado de $L_2(G,\mu)$ donde μ es la medida de Lebesgue en $\mathbb{C}$.

EJERCICIO 9.2. *Demostrar que la Ley del Paralelogramo caracteriza a los espacios de Hilbert, es decir, probar que si $\|\cdot\|$ es una norma en un espacio de Banach X que satisface la identidad del paralelogramo, esto es, para todos $x, y \in X$*

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$$

entonces dicha norma proviene de un producto escalar, y por tanto X es un espacio de Hilbert. **Sugerencia.** *Definir*

$$\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2 + \imath\|x+\imath y\|^2 - \imath\|x-\imath y\|^2).$$

Aún es necesario dar alguna sugerencia más para acabar el ejercicio.

EJERCICIO 9.3. *Utilizando la Ley del paralelogramo, probar que ℓ_p o L_p es un espacio de Hilbert si y sólo si $p = 2$.*

EJERCICIO 9.4. *Sea $\{e_1, \ldots, e_n\} \subset H$ un sistema ortonormal y sea $M = span[e_1, \ldots, e_n]$. Si $P : H \longrightarrow M$ es la proyección ortogonal entonces para todo $h \in H$*

$$P(h) = \sum_{k=1}^{n} \langle h, e_k\rangle e_k.$$

EJERCICIO 9.5. *Sea $f \in L_2[0,1]$ y sea $n \in \mathbb{N}$. Demostrar que existe un único polinomio de grado menor o igual que n que es la mejor aproximación polinómica de orden n en el sentido de L_2.*

EJERCICIO 9.6. *Sea H un espacio de Hilbert, sea $\{x_1, \dots, x_n\} \subset H$ un conjunto ortonormal y sea $x \in H$. Demostrar que de todas las sumas de la forma*

$$\sum_i a_i x_i$$

la distancia

$$\|x - \sum_i a_i x_i\|_2$$

es mínima precisamente cuando los a_i son los coeficientes de Fourier de x con respecto a $\{x_1, \dots, x_n\}$.

Pueden ser ejercicios interesantes (con suficientes sugerencias) comprobar que ciertos resultados profundos y necesitados de herramienta más o menos sofisticada en el contexto de espacios de Banach se pueden probar de manera elemental, aunque laboriosa, en el contexto de los espacios de Hilbert. En este sentido se pueden proponer.

EJERCICIO 9.7. *Sea $(x_n) \subset H$ una sucesión en un espacio de Hilbert. Entonces $x_n \to x$ en norma si y sólo si $x_n \to x$ en la topología débil y*

$$\limsup_{n\to\infty} \|x_n\| \leq \|x\|$$

EJERCICIO 9.8. *Se puede probar el Teorema de Eberlein para espacios de Hilbert: Sea $(x_n) \subset H$ una sucesión en un espacio de Hilbert. Entonces (x_n) tiene una subsucesión débilmente convergente.*

EJERCICIO 9.9. *También se puede probar directamente una de las consecuencias del PAU: Sea H un espacio de Hilbert y $B \subset H$. Entonces B es acotado si y sólo si B es débilmente acotado, es decir, si para todo $x^* \in H^*$, $x^*(B)$ está acotado.*

Capítulo 10

Teoría espectral en espacios de Hilbert: Operadores compactos normales

La Teoría Espectral alcanza su máxima perfección en los operadores definidos en espacios de Hilbert. En particular los operadores *normales*, aquellos que conmutan con su traspuesto, son "diagonalizables" en un sentido amplio gracias al Teorema Espectral. Nosotros no llegamos a desarrollar en este programa el Teorema Espectral para operadores normales, ya que eso requiere más tiempo y conocimientos de los que disponemos en esta asignatura. En cambio, sí desarrollamos en este capítulo el Teorema Espectral para operadores normales y *compactos*. En presencia de la compacidad el espectro del operador se reduce a una cantidad a lo sumo numerable de puntos, todos ellos excepto quizás el 0 autovalores, y eso permite que lo que en el Teorema Espectral sea una integral respecto de una medida vectorial, la *medida espectral*, se convierta en una serie mucho más fácil de construir con técnicas "elementales". Es frecuente en muchos libros de Análisis Funcional comenzar demostrando el Teorema Espectral para operadores compactos *autoadjuntos* de una manera razonablemente sencilla y a continuación demostrar el Teorema Espectral para operadores compactos *normales* usando razonamientos sensiblemente más delicados.

Nosotros hemos preferido seguir [**39**], con algunas modificaciones, donde el autor demuestra de manera bastante sencilla directamente el resultado para operadores compactos normales, y de ahí es fácil luego obtener la versión para operadores autoadjuntos como corolario.

Hemos incluido también para terminar el capítulo el Teorema espectral para operadores compactos no necesariamente normales, que no es otra cosa que la versión para operadores compactos en espacios infinito dimensionales del resultado de álgebra lineal que nos dice que toda forma bilineal (entre espacios finito dimensionales) admite una representación diagonal como forma bilineal (es decir, permitiendo elegir dos bases diferentes, una en cada espacio). La demostración de este resultado la hemos obtenido de [**16**].

Antes de obtener los Teoremas Espectrales que forman el punto principal del capítulo necesitamos definir el *adjunto* de un operador (y lo relacionamos con el *traspuesto* que hemos venido utilizando hasta ahora. También hacemos un estudio de los operadores *normales* y *autoadjuntos*, y sus espectros.

Al final del capítulo se puede estudiar como aplicación uno de los problemas que motivaron el comienzo del desarrollo de la Teoría Espectral, los llamados problemas de Sturm-Liouville, aunque su estudio requiere un cierto tiempo.

Operadores normales y autoadjuntos

Definición 10.1. *Sean H, K dos espacios de Hilbert y sea $T : H \longrightarrow K$ un operador. Se define su adjunto en el sentido de espacio de Hilbert, al que llamaremos T^*, como el único operador $T^* : K \longrightarrow H$ que verifica que para todos $x \in H$, $y \in K$,*

$$\langle x, T^*(y)\rangle = \langle T(x), y\rangle.$$

Es necesario ver que T^* está bien definido y es lineal y continuo. Para ver que está bien definido nótese que, dado $y \in K$, la aplicación $x^* : X \longrightarrow \mathbb{K}$ dada por

$$x^* = \langle T(x), y\rangle$$

está bien definida, es lineal y continua. Por tanto $x^* \in H^*$. Por el Teorema de Representación de Riesz sabemos que existe $z \in H$ tal que $x^*(x) = \langle x, z\rangle$ para todo $x \in H$. Definimos entonces $T^*(y) = z$. Ahora es un ejercicio ver que T^* es lineal y continuo y que es único. También es un ejercicio comprobar que para todos $x, y \in H$ se tiene

$$\langle y, T(x)\rangle = \langle T^*(y), x\rangle.$$

Se puede sugerir como ejercicio clarificar la relación entre este traspuesto en el sentido de operador hilbertiano y el traspuesto habitual que habíamos definido anteriormente. No hay más que utilizar el teorema de Representación de Riesz y dibujar un diagrama para comprobar que ambos traspuestos son (esencialmente) el mismo.

Ejercicio 10.1. *Sean K, H espacios de Hilbert, $T : H \longrightarrow K$ un operador. Sea $T^B : K^* \longrightarrow H^*$ su traspuesto en el sentido de espacios de Banach y sea $T^* : K \longrightarrow H$ su adjunto en el sentido de espacios de Hilbert. Sean además $R_H : H \longrightarrow H^*$ y $R_K : K \longrightarrow K^*$ las isometrías dadas por el Teorema de Representación de Riesz. Comprobar que $T^B \circ R_K = R_H \circ T^*$.*

En todo este capítulo entenderemos por T^* el adjunto en el sentido de espacios de Hilbert.

Además, de ahora en adelante consideraremos exclusivamente operadores $T : H \longrightarrow H$.

EJERCICIO 10.2. *Sea H un espacio de Hilbert, sean $S, T \in \mathcal{L}(H)$, $\alpha \in \mathbb{K}$. Entonces*

1. $(S + T)^* = S^* + T^*$
2. $(\alpha T)^* = \overline{\alpha} T^*$
3. $(ST)^* = T^* S^*$
4. $(T^*)^* = T$
5. *T es inversible si y sólo si T^* es inversible, y en ese caso* $(T^*)^{-1} = (T^{-1})^*$

Todo es fácil de probar, y además son esencialmente las mismas cuentas que ya hemos hecho para los adjuntos en el sentido anterior. Otra posibilidad es considerar los resultados conocidos para los adjuntos anteriores y hacer el diagrama conmutativo. Nótese que la aparición del conjugado de α en el punto 2 está relacionada con el hecho de que la isometría que aparece en el Teorema de Representación de Riesz no es lineal sino conjugada-lineal.

PROPOSICIÓN 10.2. *Si $T \in \mathcal{L}(H)$ entonces*

$$\|T^*\| = \|T\| \text{ y } \|T^*T\| = \|TT^*\| = (\|T\|)^2.$$

DEMOSTRACIÓN. Se deja como ejercicio probar que $\|T^*\| = \|T\|$. Para todos $x, y \in H$ se tiene

$$\|T(x)\|^2 = \langle T(x), T(x) \rangle = \langle x, T^*T(x) \rangle \leq \|T^*T\| \|x\|^2.$$

Tomando supremos en $x \in B_H$ se tiene

$$\|T\|^2 \leq \|T^*T\|.$$

Como además

$$\|T^*T\| \leq \|T^*\| \|T\| = \|T\|^2$$

se tiene la igualdad $\|T\|^2 = \|T^*T\|$. La otra igualdad se prueba análogamente. $\square$

Veamos algunos ejemplos de operadores adjuntos.

EJEMPLO 10.3. *Si $T_k : L_2[0, 1] \longrightarrow L_2[0, 1]$ un operador integral de Fredholm con núcleo $k(\cdot, \cdot) : [0, 1] \times [0, 1] \longrightarrow \mathbb{K}$ como en el Ejemplo 7.11. Entonces $(T_k)^*$ es un operador integral de Fredholm con núcleo $k^*(x, y) = \overline{k(y, x)}$. De nuevo nótese la necesidad de tomar el conjugado complejo.*

Ejemplo 10.4. *Si* $S : \ell_2 \longrightarrow \ell_2$ *es el desplazamiento a la derecha dado por*

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

entonces S^* *es el desplazamiento a la izquierda dado por*

$$S^*(x_1, x_2, \ldots) = (x_2, x_3, \ldots)$$

Ejercicio 10.3. *Si* $T \in \mathcal{L}(H)$ *entonces*

$$\ker(T) = (Im(T^*))^{\perp} \text{ y } \ker(T^*) = (Im(T))^{\perp}$$

donde $M^{\perp}$ *es el ortogonal de* M *en el sentido hilbertiano.*

En general la composición de operadores no es conmutativa. Veamos que los operadores con cierto grado de conmutatividad van a ser precisamente aquellos para los que podemos dar una descomposición espectral.

Definición 10.5. *Sea* $T \in \mathcal{L}(H)$. *Entonces*

1. T *se dice* normal *si* $T^*T = TT^*$.
2. T *se dice* unitario *si* $T^*T = TT^* = I$, *es decir, si* $T^* = T^{-1}$
3. T *se dice* autoadjunto *si* $T^* = T$.

Claramente los operadores autoadjuntos o unitarios son normales. En cambio un operador normal no tiene por qué ser unitario ni autoadjunto. Por ejemplo, sea $H = \mathbb{C}^2$ y sea

$$T(x) = (x_1 - x_2, x_1 + x_2)$$

Entonces

$$T^*(x) = (x_1 + x_2, -x_1 + x_2)$$

y es un ejercicio ver que T no es unitario ni autoadjunto.

Observación 10.6. *Si* T *es normal y* S *es tal que* $S^*S = I$ *(para esto* S *no tiene por qué ser necesariamente unitario, piénsese en el operador desplazamiento a la derecha del Ejemplo 10.4) entonces* $U = STS^*$ *es normal. En efecto, en ese caso* $U^* = ST^*S^*$ *y por tanto*

$$UU^* = STS^*ST^*S^* = STT^*S^* = ST^*TS^* = ST^*S^*STS^* = U^*U.$$

Proposición 10.7. *Sea* $T \in \mathcal{L}(H)$. *Entonces*

1. T *es normal si y sólo si para todos* $x, y \in H$
$$\langle T(x), T(y)\rangle = \langle T^*(x), T^*(y)\rangle$$
2. T *es unitario si y sólo si para todos* $x, y \in H$
$$\langle T(x), T(y)\rangle = \langle x, y\rangle = \langle T^*(x), T^*(y)\rangle$$

3. *T es autoadjunto si y sólo si para todos $x, y \in H$*

$$\langle T(x), y\rangle = \langle x, T(y)\rangle$$

Demostración. Si T es normal entonces para todos $x, y \in H$

$$\langle T(x), T(y)\rangle = \langle x, T^*T(y)\rangle = \langle x, TT^*(y)\rangle = \langle T^*(x), T^*(y)\rangle$$

Recíprocamente, si para todos $x, y \in H$

$$\langle T(x), T(y)\rangle = \langle T^*(x), T^*(y)\rangle$$

entonces

$$\langle x, T^*T(y)\rangle = \langle T(x), T(y)\rangle = \langle T^*(x), T^*(y)\rangle = \langle x, TT^*(y)\rangle$$

de donde se sigue que $TT^*(y) = T^*T(y)$ para todo $y \in H$, es decir T es normal. Los otros dos apartados se prueban análogamente y la demostración se deja como ejercicio. □

Ejemplo 10.8. *Sea $H = L_2[0,1]$, sea $f \in L_\infty[0,1]$ y sea*

$$T : L_2[0,1] \longrightarrow L_2[0,1]$$

el operador definido como

$$T(g) = fg.$$

T está bien definido y es lineal y continuo, ya que para todo $g \in H$

$$\|T(g)\|_2 = \left(\int_0^1 |fg|^2\right)^{\frac{1}{2}} \leq \|f\|_\infty \|g\|_2.$$

Además para todos $g, h \in H$

$$\langle g, T^*(h)\rangle = \langle T(g), h\rangle = \int_0^1 fg\overline{h} = \int_0^1 g(\overline{\overline{f}h}) = \langle g, \overline{f}h\rangle,$$

y por lo tanto

$$T^*(g) = \overline{f}g \text{ para todo } g \in H.$$

Entonces

$$T^*T(g) = |f|^2 g = TT^*(g)$$

y por tanto T es normal. En cierto sentido, todo operador normal es así (factorizándolo vía una isometría adecuada), aunque esto no es nada obvio y es una consecuencia (o más bien una reformulación) del Teorema Espectral. El lector interesado en este resultado y en general en el Teorema Espectral debería consultar **[25]**.

Sea H un espacio de Hilbert separable. Ya hemos visto que en ese caso H es (isométrico a) ℓ_2. Como consecuencia de los resultados que probamos en el capítulo anterior referidos a bases hilbertianas, queda claro que todo operador $T : \ell_2 \longrightarrow \ell_2$ queda completamente determinado si conocemos la matriz doblemente infinita $A = (a_{i,j})_{i,j\in\mathbb{N}}$ definida como

$$a_{i,j} = \langle T(e_j), e_i\rangle$$

ya que en ese caso sabemos que para todo $x \in \ell_2$

$$T(x) = \sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty} a_{i,j}x_j\right)e_i.$$

Análogamente a lo visto en el Ejercicio 5.1 se puede probar con facilidad que el operador T^* lleva asociado la matriz $\overline{A}^t$, la matriz conjugada y traspuesta de A de manera que para todo $x \in \ell_2$

$$T^*(x) = \sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty} \overline{a_{j,i}}x_j\right)e_i.$$

(De nuevo la razón de que aparezca el conjugado complejo es que estamos considerando el adjunto hilbertiano de T).

Entonces resulta claro que T es autoadjunto si y sólo si $A = \overline{A}^t$, es decir, si A es conjugada-simétrica.

Además T^*T viene representado por

$$\overline{A}^t A = \sum_n \overline{k_{n,i}}k_{n,j}$$

y TT^* viene representado por

$$A\overline{A}^t = \sum_n k_{i,n}\overline{k_{j,n}}.$$

Así pues T es unitario si y sólo si $\overline{A}^t A = A\overline{A}^t = I$, es decir, si

$$\sum_n \overline{k_{n,i}}k_{n,j} = \sum_n k_{i,n}\overline{k_{j,n}} = \delta_{i,j}. \tag{24}$$

Nótese que, en ese caso, si consideramos el vector $k^i \in \ell_2$ definido como $k^i(n) = k_{n,i}$ entonces (24) nos dice que

$$\langle k^i, k^j\rangle = \delta_{i,j},$$

es decir, T es unitario si y sólo si las columnas (y análogamente las filas) de la matriz A forman una base ortonormal de ℓ_2.

Veremos más adelante que si T es compacto, entonces T es normal si y sólo si A es diagonalizable. De momento, observando que T es normal si y sólo si $\overline{A}^t A = A\overline{A}^t$, queda claro que si A es diagonal entonces T es normal.

Veamos ahora algunas caracterizaciones de los operadores normales, unitarios y autoadjuntos que nos serán útiles más adelante.

Proposición 10.9. *Sea H un espacio de Hilbert* complejo *y sea $T \in \mathcal{L}(H)$. Entonces T es autoadjunto si y sólo si $\langle T(x), x\rangle \in \mathbb{R}$ para todo $x \in H$.*

Demostración. Si $T = T^*$ entonces, para todo $x \in H$,

$$\langle T(x), x\rangle = \langle x, T^*(x)\rangle) = \langle x, T(x)\rangle = \overline{\langle T(x), x\rangle}$$

y por tanto $\langle T(x), x\rangle \in \mathbb{R}$.

Recíprocamente, supongamos que $\langle T(x), x\rangle \in \mathbb{R}$ para todo $x \in H$. Entonces para todo $\alpha \in \mathbb{C}$ y para todos $x, y \in H$ se tiene

$$\langle T(x+\alpha y), x+\alpha y\rangle = \langle T(x), x\rangle + \overline{\alpha}\langle T(x), y\rangle +$$
$$+\alpha\langle T(y), x\rangle + |\alpha|^2\langle T(y), y\rangle \in \mathbb{R}$$

Por lo tanto, como $\langle T(x), x\rangle \in \mathbb{R}$ y $|\alpha|^2\langle T(y), y\rangle \in \mathbb{R}$, se sigue que

$$\overline{\alpha}\langle T(x), y\rangle + \alpha\langle T(y), x\rangle \in \mathbb{R},$$

es decir

$$\overline{\alpha}\langle T(x), y\rangle + \alpha\langle T(y), x\rangle = \overline{\overline{\alpha}\langle T(x), y\rangle + \alpha\langle T(y), x\rangle} =$$
$$= \alpha\langle y, T(x)\rangle + \overline{\alpha}\langle x, T(y)\rangle = \overline{\alpha}\langle T^*(x), y\rangle + \alpha\langle T^*(y), x\rangle.$$

Puesto que esto es cierto para todo $\alpha \in \mathbb{C}$, se sigue con facilidad (por ejemplo tomando $\alpha = 1$ y $\alpha = \imath$ y operando) que para todo $x, y \in H$

$$\langle T(x), y\rangle = \langle T^*(x), y\rangle,$$

de donde $T = T^*$. $\square$

Observación 10.10. *Obviamente el resultado anterior es falso si H es un espacio de Hilbert real, ya que en ese caso $\langle T(x), y\rangle \in \mathbb{R}$ para todos $x, y \in H$ y obviamente hay operadores no autoadjuntos (considérese cualquier matriz cuadrada no simétrica de por ejemplo orden 2).*

Proposición 10.11. *Sea $T \in \mathcal{L}(H)$. Si T es autoadjunto entonces*

$$\|T\| = \sup_{x\in B_H} |\langle T(x), x\rangle|.$$

Demostración. Sea $M = \sup_{x \in B_H} |\langle T(x), x\rangle|$. Claramente $M \leq \|T\|$. Para la otra desigualdad observemos que para todos $x, y \in H$ se tiene

$$\begin{aligned}\langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle &= 2\langle T(x), y\rangle + 2\langle T(y), x\rangle = \\ &= 2\langle T(x), y\rangle + 2\langle y, T^*(x)\rangle = 2\langle T(x), y\rangle + 2\langle y, T(x)\rangle = \\ &= 2\langle T(x), y\rangle + 2\overline{\langle T(x), y\rangle} = 4\Re\langle T(x), y\rangle.\end{aligned}$$

Por lo tanto

$$\begin{aligned}4\Re\langle T(x), y\rangle &= \langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle \leq \\ &\leq M(\|x+y\|^2 - \|x-y\|^2) = 2M(\|x\|^2 + \|y\|^2),\end{aligned}$$

donde en la última igualdad hemos usado la Ley del Paralelogramo. Por tanto, si $\|x\| = \|y\| = 1$ tenemos

$$4\Re\langle T(x), y\rangle \leq 4M.$$

Finalmente, si $\langle T(x), y\rangle = |\langle T(x), y\rangle| e^{\imath\theta}$, sustituyendo x por $xe^{-\imath\theta}$ obtenemos que para todo $x, y \in H$

$$|\langle T(x), y\rangle| \leq M$$

de donde $\|T\| \leq M$. □

Nótese que, en particular, si T es autoadjunto y $\langle T(x), x\rangle = 0$ para todo $x \in H$ entonces $T = 0$.

Proposición 10.12. *Sea $T \in \mathcal{L}(H)$. Entonces T es normal si y sólo si para todo $x \in H$*

$$\|T(x)\| = \|T^*(x)\|.$$

Además en ese caso

$$\|T^2\| = \|T^*T\| = \|T\|^2$$

Demostración. Si $x \in H$, se tiene

$$\begin{aligned}\|T(x)\|^2 - \|T^*(x)\|^2 &= \langle T(x), T(x)\rangle - \langle T^*(x), T^*(x)\rangle = \\ &= \langle T^*T(x), x\rangle - \langle TT^*(x), x\rangle = \langle (T^*T - TT^*)(x), x\rangle.\end{aligned}$$

Como $S = T^*T - TT^*$ es autoadjunto (esto se ve muy fácil), de la Proposición 10.11 nos dice que $S = 0$ si y sólo si

$$\langle S(x), x\rangle = 0 \text{ para todo } x \in H,$$

es decir, si y sólo si

$$\langle (T^*T - TT^*)(x), x\rangle = \|T(x)\|^2 - \|T^*(x)\|^2 = 0 \text{ para todo } x \in H,$$

es decir si y sólo si

$$\|T(x)\| = \|T^*(x)\| \text{ para todo } x \in H.$$

Además en ese caso

$$\|T^2(x)\| = \|T(T(x))\| = \|T^*(T(x))\| \text{ para todo } x \in H$$

y por tanto

$$\|T^2\| = \|T^*T\| = \|T\|^2,$$

donde la segunda igualdad se sigue de 10.2 □

PROPOSICIÓN 10.13. *Sea $T \in \mathcal{L}(H)$. Entonces T es unitario si y sólo si T es sobre y $\|T(x)\| = \|x\|$ para todo $x \in H$. Además en ese caso $\|T^{-1}(x)\| = \|x\|$ para todo $x \in H$ y*

$$\|T\| = \|T^{-1}\| = 1.$$

DEMOSTRACIÓN. Para todo $x \in H$ se tiene

$$\|T(x)\|^2 - \|x\|^2 = \langle T(x), T(x)\rangle - \langle x, x\rangle =$$

$$= \langle T^*T(x), x\rangle - \langle x, x\rangle = \langle (T^*T - I)(x), x\rangle.$$

Como $S = T^*T - I$ es autoadjunto se tiene que $S = 0$ si y sólo si

$$\|T(x)\| = \|x\| \text{ para todo } x \in H.$$

Por tanto, si $\|T(x)\| = \|x\|$ para todo $x \in H$ y T es sobreyectiva, entonces $T^*T = I$ y T es biyectivo, de forma que

$$TT^* = (TT^*)(TT^{-1}) = T(T^*T)T^{-1} = TT^{-1} = I$$

y por tanto T es unitario.

Recíprocamente, si T es unitario entonces $T^*T = I$ y $T^{-1} = T^*$ de forma que $\|T(x)\| = \|x\|$ para todo $x \in H$ y T es sobreyectiva.

En ese caso además

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

y tomando supremos en $x \in B_H$ se tiene

$$\|T\| = \|T^{-1}\| = 1.$$

□

Como quizás ya se haya podido apreciar en este momento puede haber una cierta analogía entre los operadores en $\mathcal{L}(H)$ y los números complejos, donde el conjugado de un complejo se asocia al adjunto de un operador. En esta analogía los operadores autoadjuntos se corresponderían con los números reales (coinciden con su conjugado). Eso motiva el siguiente

TEOREMA 10.14. *Sea H un espacio de Hilbert complejo y $T \in \mathcal{L}(H)$. Entonces existen dos únicos operadores autoadjuntos $R, S \in \mathcal{L}(H)$ tales que*

$$T = R + \imath S.$$

Además T es normal si y sólo si $RS = SR$, T es unitario si y sólo si $RS = SR$ y $R^2 + S^2 = I$ y T es autoadjunto si y sólo si $S = 0$.

También tendremos ocasión de usar más adelante que T es compacto si y sólo si R y S son compactos.

DEMOSTRACIÓN. Sean

$$R = \frac{T + T^*}{2} \text{ y } S = \frac{T - T^*}{2\imath}.$$

Es fácil ver que R y S son autoadjuntos y $T = R + \imath S$.

Si $T = R' + \imath S'$ con R' y S' autoadjuntos, entonces

$$T^* = (R' + \imath S')^* = R'^* - \imath S'^* = R' - \imath S'$$

y por tanto

$$R' = \frac{T + T^*}{2} = R \text{ y } S' = \frac{T - T^*}{2\imath} = S.$$

El resto de la demostración se deja como ejercicio. □

Con estos conocimientos previos acerca de operadores autoadjuntos y normales a nuestra disposición, podemos desarrollar la Teoría Espectral de operadores normales compactos.

LEMA 10.15. *Sea H un espacio de Hilbert complejo y sea $T \in \mathcal{L}(H)$. Entonces $\lambda \in \sigma(T)$ si y sólo si $\overline{\lambda} \in \sigma(T^*)$.*

Si además T es normal, entonces para todo $\lambda \in \mathbb{C}$ se tiene que

$$\ker(T - \lambda I) = \ker(T - \lambda I)^* = \ker(T^* - \overline{\lambda} I)$$

DEMOSTRACIÓN. La Proposición 8.3 nos dice que un operador es inversible si y sólo si su adjunto lo es. Por lo tanto $T - \lambda I$ es inversible si y sólo si $(T - \lambda I)^* = T^* - \overline{\lambda} I$ es inversible. A partir de ahí el primer enunciado se sigue inmediatamente.

Supongamos ahora que T es normal. Entonces es fácil ver que $T - \lambda I$ también es normal. Por lo tanto, la Proposición 10.12 nos dice que para todo $x \in H$

$$\|(T - \lambda I)(x)\| = \|(T - \lambda I)^*(x)\|,$$

y de ahí se sigue que $\ker(T - \lambda I) = \ker(T - \lambda I)^*$ □

TEOREMA 10.16. *Sea H un espacio de Hilbert y sea $0 \neq T \in \mathcal{L}(H)$ un operador compacto y autoadjunto. Entonces al menos uno de los números $\|T\|$ o $-\|T\|$ es un autovalor de T.*

DEMOSTRACIÓN. Sea T como en la hipótesis. De la Proposición 10.11 se sigue que existe una sucesión $(x_n) \subset B_H$ tal que

$$|\langle T(x_n), x_n\rangle| \to \|T\|.$$

Puesto que $\langle T(x), x\rangle \in \mathbb{R}$ para todo $x \in H$, tomando una subsucesión si fuera necesario podemos suponer que $\langle T(x_n), x_n\rangle \to \lambda$, donde $\lambda = \|T\|$ o $\lambda = -\|T\|$. Como $|\lambda| = \|T\|$, se tiene

$$0 \leq \|(T - \lambda I)(x_n)\|^2 = \|T(x_n)\|^2 - 2\lambda\langle T(x_n), x_n\rangle + \lambda^2\|x_n\|^2 \leq$$
$$\leq \|T\|^2\|x_n\|^2 + \lambda^2\|x_n\|^2 - 2\lambda\langle T(x_n), x_n\rangle \leq$$
$$\leq 2\lambda^2 - 2\lambda\langle T(x_n), x_n\rangle \to 0$$

de donde se obtiene que $\|(T - \lambda I)(x_n)\| \to 0$.

Como además T es compacto, existe una subsucesión $(x_{n_k})_k \subset (x_n)$ y un elemento $y \in H$ tales que $T(x_{n_k}) \to y$, y por tanto, como $T(x_{n_k}) - \lambda x_{n_k} \to 0$, se sigue que $x_{n_k} \to \frac{y}{\lambda}$ y

$$T(y) = \lambda y$$

lo que termina la demostración. □

Vamos a ver ahora que, usando el resultado anterior, podemos decir algo similar de los operadores compactos normales.

TEOREMA 10.17. *Sea H un espacio de Hilbert complejo y sea $0 \neq T \in \mathcal{L}(H)$ un operador normal y compacto. Entonces T admite al menos un autovalor distinto de 0.*

DEMOSTRACIÓN. Sea $T = R + \imath S$ con R, S compactos y autoadjuntos y $SR = RS$ (véase el Teorema 10.14). Supongamos inicialmente que $R \neq 0$. Entonces, el Teorema 10.16 nos dice que R tiene al menos un autovalor $\alpha \neq 0$. Sea $M = \ker(R - \alpha I)$. El subespacio M es invariante por S ya que para todo $x \in M$, puesto que $SR = RS$, se tiene

$$RS(x) = SR(x) = S(\alpha x) = \alpha S(x)$$

luego $S(x) \in \ker(R - \alpha I) = M$.

Entonces $S_{|_M} : M \longrightarrow M$ es un operador compacto y autoadjunto del espacio de Hilbert M en sí mismo. Aplicando de nuevo el Teorema

10.16 tenemos que existe $\beta \in \sigma_e(S_{|M})$. Sea ahora $0 \neq x \in M$ tal que $S(x) = \beta x$. Entonces

$$T(x) = (R + \imath S)(x) = (\alpha + \imath\beta)(x)$$

luego $\alpha + \imath\beta \in \sigma_e(T)$

Finalmente, si $R = 0$, entonces $S \neq 0$ y por tanto el Teorema 10.16 nos dice que S tiene un autovalor $\beta \neq 0$. Entonces $\imath\beta$ es autovalor de $T = \imath S$. □

Ya vimos en la Proposición 8.19 que si T es compacto entonces para todo $\lambda \in \mathbb{K}$ existe $n_0 \in \mathbb{N}$ tal que para todo $n \geq n_0$

$$\ker(T - \lambda I)^n = \ker(T - \lambda I)^{n_0}.$$

Dado un autovalor λ definimos entonces su *índice* $\nu(\lambda)$ como el primer $n_0 \in \mathbb{N}$ para el que ocurre eso, es decir $\nu(\lambda)$ es el único natural tal que

$$\ker(T - \lambda I)^n \subsetneq \ker(T - \lambda I)^{n+1} \text{ para todo } 0 \leq n < \nu(\lambda)$$

y

$$\ker(T - \lambda I)^n = \ker(T - \lambda I)^{n+1} \text{ para todo } n \geq \nu(\lambda)$$

Proposición 10.18. *Si $T \in \mathcal{L}(H)$ es normal entonces para todo $\lambda \in \sigma_e(T)$ se tiene que $\nu(\lambda) = 1$.*

Demostración. Sea $\lambda \in \sigma_e(T)$ y sea $x \in \ker(T - \lambda I)^2$. Sólo hay que ver que $x \in \ker(T - \lambda I)$.

Puesto que $T - \lambda I$ es normal, aplicando la Proposición 10.12 se tiene que

$$\|(T^* - \overline{\lambda} I)(T - \lambda I)(x)\| = \|(T - \lambda I)(T - \lambda I)(x)\| = 0.$$

Por lo tanto

$$\|(T - \lambda I)(x)\|^2 = \langle (T - \lambda I)(x), (T - \lambda I)(x) \rangle =$$
$$= \langle (T^* - \overline{\lambda} I)(T - \lambda I)(x), x \rangle = 0$$

lo que implica que $x \in \ker(T - \lambda I)$ y termina la demostración. □

Definición 10.19. *Sea $T \in \mathcal{L}(H)$ un operador normal. Dado $\lambda \in \sigma_e(T)$ definimos su* multiplicidad *como la dimensión de* $\ker(T - \lambda I)$.

Observación 10.20. *En realidad se define la* multiplicidad algebraica *como* $\dim \ker(T - \lambda I)^{\nu(\lambda)}$ *y la* multiplicidad geométrica *como* $\dim \ker(T - \lambda I)$. *En el caso de que T sea normal (el único que estudiaremos en esta memoria), la Proposición 10.18 nos garantiza que ambos números coinciden, y los denominamos* multiplicidad *simplemente.*

PROPOSICIÓN 10.21. *Si* $T \in \mathcal{L}(H)$ *es normal y* $\lambda, \mu \in \sigma_e(T)$ *entonces*

$$\ker(T - \lambda I) \perp \ker(T - \mu I)$$

DEMOSTRACIÓN. En efecto, para todos $x \in \ker(T - \lambda I)$, $y \in \ker(T - \mu I)$, usando que $\overline{\lambda} \in \sigma_e(T^*)$ (Lema 10.15), se tiene

$$\lambda\langle y, x\rangle = \langle y, \overline{\lambda}x\rangle = \langle y, T^*(x)\rangle = \langle T(y), x\rangle = \langle \mu y, x\rangle = \mu\langle y, x\rangle.$$

Puesto que $\lambda \neq \mu$, se sigue que $\langle y, x\rangle = 0$. □

Necesitamos un lema más antes de enunciar y demostrar el Teorema Espectral para operadores compactos normales.

LEMA 10.22. *Sea* $M \subset H$ *un subespacio cerrado y sea* $P : H \longrightarrow M$ *la proyección ortogonal. Entonces* $P^* = P$

DEMOSTRACIÓN. Para todo $x \in H$, $P(x)$ y $x - P(x)$ son ortogonales, es decir

$$\langle P(x), x - P(x)\rangle = 0$$

y por tanto

$$\langle P(x), x\rangle = \langle P(x), P(x)\rangle = \|P(x)\|^2.$$

Ahora la Proposición 10.9 nos garantiza que P es autoadjunto. □

Teoremas Espectrales para Operadores Compactos

TEOREMA 10.23. (Teorema espectral para operadores compactos normales) *Sea* H *un espacio de Hilbert complejo y* $T \in \mathcal{L}(H)$ *un operador compacto normal. Sea* $(\lambda_n)_n \subset \mathbb{C}$ *la sucesión (finita o infinita) de autovalores, ordenados por módulos decrecientes y contado cada uno de ellos tantas veces como su multiplicidad. Entonces existe una sucesión* $(e_n)_n \subset H$ *con tantos elementos como la sucesión* $(\lambda_n)_n$*, de manera que para cada* n *se tiene que* $T(e_n) = \lambda_n e_n$*, es decir,* e_n *es un autovalor de* λ_n*. Además la sucesión* $(e_n)_n$ *es un sistema ortonormal y se verifica que para todo* $x \in H$

$$T(x) = \sum_n \lambda_n \langle x, e_n\rangle e_n \tag{25}$$

DEMOSTRACIÓN. Sea H y T como en la hipótesis. Por el Teorema 10.17 sabemos que $\sigma_e(T) \neq \emptyset$, y por el Teorema 8.22 sabemos que $\sigma_e(T)$ es a lo sumo numerable. Sea $(\mu_n)_n$ la sucesión (finita o infinita) de autovalores. De nuevo por el Teorema 8.22 sabemos que si (μ_n) es infinita entonces $\mu_n \to 0$. Por tanto, siempre tiene podemos elegir la sucesión (μ_n) ordenada de tal forma que

$$|\mu_1| \geq |\mu_2| \geq |\mu_3| \geq \cdots ,$$

bien entendido que esta ordenación no tiene por qué ser única ya que pueden existir autovalores distintos con el mismo módulo. Para todo n, sea $m(n)$ la multiplicidad (algebraica o geométrica) de μ_n. Sea entonces (λ_n) la sucesión (finita o infinita) formada por $m(1)$ copias de μ_1, a continuación $m(2)$ copias de μ_2, etc. Está claro que (λ_n) es infinita si y sólo si (μ_n) es infinita, y que en ese caso también $\lambda_n \to 0$. Para todo n sea $N(n) = \ker(T - \mu_n I)$. Por definición de multiplicidad $\dim N(n) = m(n)$. Para cada n, sea

$$\{e_{n,1}, e_{n,2}, \ldots , e_{n,m(n)}\} \subset H$$

una base ortonormal de $N(n)$. Sea entonces $(e_n) \subset H$ la sucesión (finita o infinita) formada por

$$(e_{1,1}, e_{1,2}, \ldots , e_{1,m(1)}, e_{2,1}, e_{2,2}, \ldots , e_{2,m(2)}, \cdots$$

Por la construcción de (e_n) y (λ_n) queda claro que para todo n

$$T(e_n) = \lambda_n e_n.$$

Además, puesto que para todo $j \neq i$ se tiene que $N(j) \perp N(i)$ (Proposición 10.21) se sigue que la familia (e_n) es ortonormal.

Consideramos ahora el operador $T_0 : H \longrightarrow H$ dado por

$$T_0(x) = \sum_n \lambda_n \langle x, e_n\rangle e_n \text{ para todo } x \in H.$$

<u>Aserto:</u> T_0 está bien definido y es continuo.

Demostración del Aserto: Para todo $x \in H$, utilizando la desigualdad de Bessel se tiene

$$\sum_n |\langle \lambda_n \langle x, e_n\rangle|^2 \leq |\lambda_1|^2 \sum_n |\langle x, e_n\rangle|^2 \leq\leq |\lambda_1| \|x\|^2,$$

y por el Teorema de Riesz-Fischer se sigue el Aserto.

De hecho se sigue de la implicación fácil del Teorema 10.24, que podríamos haber probado previamente, que T_0 es compacto y normal, aunque no lo necesitamos.

Veamos que $T = T_0$.

Sea $M = \overline{span[(e_n)]}$, el subespacio vectorial cerrado generado por los (e_n). Por el Teorema de la Proyección Ortogonal, M está complementado en H. Sea $P : H \longrightarrow M$ la proyección ortogonal de H sobre M. Es claro que para todo n

$$T_0(e_n) = \lambda_n e_n = T(e_n).$$

Por tanto, por linealidad y continuidad se sigue que $T_{|M} = T_{0|M}$, es decir

$$T_0 P = TP$$

Además, usando que $P^* = P$ (Lema 10.22), se tiene que para todo $x \in H$

$$T_0(x) = \sum_n \lambda_n \langle x, e_n\rangle e_n = \sum_n \lambda_n \langle x, P(e_n)\rangle e_n =$$

$$= \sum_n \lambda_n \langle P(x), e_n\rangle e_n = T_0 P(x).$$

Por lo tanto $T_0 = T_0 P$ y de ahí $TP = T_0$.

Además, como $PT(x) \in M$ para todo $x \in H$ y como (e_n) es una base ortonormal de M (entendiéndola como base hilbertiana si es una colección infinita), se tiene

$$PT(x) = \sum_n \langle PT(x), e_n\rangle e_n = \sum_n \langle x, T^* P(e_n)\rangle e_n.$$

Además, por el Lema 10.15 se tiene que

$$T^* P(e_n) = T^*(e_n) = \overline{\lambda_n} e_n$$

y por tanto

$$PT(x) = \sum_n \lambda_n \langle x, e_n\rangle e_n = T_0(x)$$

Así pues $TP = PT = T_0$ y, tomando adjuntos,

$$PT^* = T^* P.$$

Por lo tanto, dos cualesquiera de T, T^* y P conmutan entre si y de aquí se deduce fácilmente que $T - T_0 = T - TP$ es normal.

Finalmente, supongamos que $T \neq T_0$. Como $T - T_0$ es normal, compacto y distinto de 0, el Teorema 10.17 nos dice que $T - T_0$ tiene al menos un autovalor $\mu \neq 0$. Sea $0 \neq e \in H$ un autovector de μ.

Como $(T - T_0)(e_n) = 0$ para todo n y $(T - T_0)e = \mu e$, se tiene que $e \perp e_n$ para todo n. En efecto,

$$\mu\langle e_n, e\rangle = \langle e_n, \overline{\mu} e\rangle = \langle e_n, (T - T_0)^* e\rangle = \langle (T - T_0)e_n, e\rangle = \langle 0, e\rangle = 0$$

Por lo tanto,

$$T_0(e) = \sum_n \lambda_n \langle e, e_n \rangle e_n = 0$$

y de ahí

$$T(e) = (T - T_0)e = \mu e.$$

De forma que μ es un autovalor de T. Puesto que la sucesión λ_n estaba formada por todos los los autovalores μ_n (contados con su multiplicidad), se tiene que $\mu = \mu_j$ para algún j, y por tanto $e \in N(j)$. Pero entonces es imposible que e sea distinto de 0 y simultáneamente sea ortogonal a todos los elementos de una base ortogonal de $N(j)$.

De esta contradicción se sigue que $T = T_0$ □

En el caso especial de operadores compactos autoadjuntos, el Teorema anterior queda igual, con la salvedad de que ahora todos los autovalores (μ_n) (y por tanto también los (λ_n)) son reales.

Es fácil ver, y se puede proponer como ejercicio, que el Teorema 10.23 (y su correspondiente versión para operadores compactos autoadjuntos) son caracterizaciones, de manera que dado un operador $T : H \longrightarrow H$, T es compacto y normal si y sólo si T se puede escribir en la forma 25, y T es compacto y autoadjunto si y sólo si T se puede escribir en la forma 25, con $(\lambda_n) \subset \mathbb{R}$.

Veamos para acabar este capítulo qué se puede decir si T es un operador compacto, no necesariamente normal. Sugerimos para ello la demostración de [**16**]. Se puede dar otra demostración, como en [**39**], utilizando el Teorema 10.23 y la *descomposición polar* de un operador.

Teorema 10.24. *Sea H un espacio de Hilbert. Entonces un operador $T \in \mathcal{L}(H)$ es compacto si y sólo si existen sucesiones (finitas o infinitas) $(\lambda_n) \subset \mathbb{C}$, $(e_n) \subset H$ y $(f_n) \subset H$ con el mismo número de índices tales que $\lambda_n \to 0$ si es infinita y (e_n) y (f_n) son familias ortonormales de forma que T se puede escribir en la forma*

$$T(x) = \sum_n \lambda_n \langle x, e_n \rangle f_n \tag{26}$$

Aplicaciones

Una aplicación clásica de la Teoría espectral de operadores compactos en espacios de Hilbert es la resolución de los llamados Problemas de Sturm-Liouville. En los libros de nuestra bibliografía, se pueden ver presentaciones de esta aplicación en [**31**, Apendix C] o en [**13**, Section II.6].

Prácticas sugeridas.

Ejercicio 10.4.
1. *Si S, T son autoadjuntos entonces $S + T$ es autoadjunto. Además ST es autoadjunto si y sólo si S y T conmutan.*
2. *Si S, T son unitarios entonces ST es unitario. Buscar alguna condición necesaria y suficiente sobre S y T para que $S+T$ sea unitario.*
3. *Si S y T son normales y además S conmuta con T^* y T conmuta con S^* entonces $S + T$ y ST son normales.*

Ejercicio 10.5. *Si $(T_n) \subset \mathcal{L}(H)$ es una sucesión de operadores autoadjuntos (resp. unitarios, normales) tal que $T_n \to T$ entonces T es autoadjunto (resp. unitario, normal).* **Sugerencia***: Si $T_n \to T$ entonces $T_n^* \to T^*$.*

Ejercicio 10.6 (Operador de Volterra). *Sea $k : [0, 1] \times [0, 1] \longrightarrow \mathbb{R}$ la función característica del conjunto $\{(x, y) \in [0, 1] \times [0, 1]$ tales que $y < x\}$. Sea $V : L_2 \longrightarrow L_2$ el operador integral dado por*

$$V(f)(s) = \int_0^1 f(t)k(s, t)dt.$$

Nótese que $V(f)(s) = \int_0^s f(t)dt$. Se pide
1. *Calcular V^**
2. *Demostrar que V no tiene autovalores.*
3. *Demostrar que V es compacto.*

Ejercicio 10.7. *Si $S : \ell_2 \longrightarrow \ell_2$ calcular S^*, SS^* y S^*S. Además demostrar*
1. $\sigma_e(S^*) = \{z \in \mathbb{C}; |z| < 1\}$
2. $\sigma_e(S) = \emptyset$
3. $\sigma(S) = \sigma(S^*) = \{z \in \mathbb{C}; |z| \leq 1\}$

Ejercicio 10.8. *Sea $D : \ell_2 \longrightarrow \ell_2$ el operador diagonal dado por*

$$D(a) = (d_n a_n)_n$$

con $d_n = 2^{-n}$, y sea $S : \ell_2 \longrightarrow \ell_2$ el desplazamiento a la derecha. Sea $T = SD$. Demostrar que T es compacto, $\sigma_e(T) = \emptyset$ y $\sigma(T) = \{0\}$.

Ejercicio 10.9 (Teorema de Hellinger-Töplitz). *Sea H un espacio de Hilbert y sea $T : H \longrightarrow H$ un operador lineal y autoadjunto. Entonces T es continuo.*

Ejercicio 10.10. *Sea H un espacio de Hilbert complejo y sea $T \in \mathcal{L}(H)$ un operador compacto normal no nulo. Entonces existe $\lambda \in \sigma_e(T)$ tal que $|\lambda| = \|T\|$.*

Bibliografía

[1] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis. A Hitchhiker guide,* Springer 1999.

[2] T. M. Apostol, *Análisis Matemático* Reverté 1991.

[3] K. Arrow and G. Debreu, Existence of an equilibrium for a competitive economy, *Econometrica* **22** (1954), 265–290.

[4] P. Artzner, F. Delbaen, J.M. Eber, D. Heath, Coherent measures of risk, *Math. Finance* **9** (1999), 203–228.

[5] G. Bachman y L. Narichi, *Functional Analysis*, Academic Press, 1966.

[6] S. Banach, Sur les opérations dans les ensembles abstraits et leurs applications aux équations intégrales, *Fundamenta Math.,* **3** (1922) 131-181.

[7] S. Banach, *Theorie des opérations linéaires*, Warszawa, 1932.

[8] B. Beauzamy, *Introduction to Banach spaces and their geometry*, North-Holland, 1985.

[9] F. Black and M. Scholes, The valuation of option contracts and a test of market efficiency, *J. Finance* **27** (1972), 399–417.

[10] F. Black and M. Scholes, The pricing of options and corporate liability, *J. Political Economy*, **81** (1973), 637–659.

[11] F. Bombal, Análisis Funcional: una perspectiva histórica, Proceedings of the Seminar of Mathematical Analysis 2002-2003, Secretariado de Publicaciones, Universidad de Sevilla (2993) 81–117.

[12] G. Choquet, *Cours d'Analyse. Tome II. Topologie.,* Masson et C$^{\text{ie}}$, Paris, 1964.

[13] J. B. Conway, *A Course in Functional Analysis* Springer 1990.

[14] F. Delbaen y W. Schachermayer, The fundamental theorem of asset pricing for unbounded stochastic processes, *Math. Ann.* **312** (1998) 215–250.

[15] J. Diestel, *Sequences and series in Banach Spaces,* Springer-Verlag 1984.

[16] J. Diestel, H. Jarchow and A. Tonge, *Absolutely Summing Operators,* Cambridge Stud. Adv. Math. **43**, Cambridge Univ. Press, Cambridge 1995.

[17] D. Duffie, *Security Markets. Stochastic models* Academic Press 1988.

[18] P. Enflo, On the invariant subspace problem for Banach spaces, *Acta Math* **158** (1987) 213–313.

[19] P. Enflo, A counterexample to the Approximation problem for Banach spaces, *Acta Math* **130** (1973) 309–317.

[20] P. Enflo y V. Lomonosov, *Some aspects of the invariant subspace problem. Handbook of the geometry of Banach spaces, Vol. I*, 533–559, North-Holland, Amsterdam, 2001.

[21] M. Fabian, P. Habala, P. Hajek, V. Montesinos Santalucia, J. Pelant, V. Zizler, *Functional Analysis and infinite dimensional geometry*, Springer, 2001.

[22] C. Goffman and G. Pedrick, *First Course in Functional Analysis,* Prentice-Hall 1965.
[23] A. Grothendieck, Résumé de la théorie métrique des produits tensoriels topologiques, *Bol. Soc. Mat. São Paulo* **8** (1953) 1–79.
[24] P. Habala, P. Hajek, V. Zizler, *Introduction to Banach Spaces I*, Matfyzpress Univerzity Karlovy, Prague, 1996.
[25] P. R. Halmos, What does the spectral theorem say?, *Amer. Math. Monthly* **70** (1963) 241–247.
[26] J. Horvath, *Topological Vector Spaces and Distributions,* Addison-Wesley Publishing Company 1966.
[27] R. C. James, A non-reflexive Banach space isometric with its second conjugate space, *Proc. Nat. Acad. Sci. USA,* **37** (1951) 174–177.
[28] P. E. Kopp, *Martingales and stochastic integrals,* Cambridge University Press 1984.
[29] S. Lang, *Real and Functional Analysis,* Springer-Verlag 1993.
[30] R. Larsen, *Functional Analysis: An Introduction,* Marcel Dekker, New York 1973.
[31] B. V. Limaye, *Functional Analysis*, Wiley Eastern Ltd., 1981.
[32] J. Lindenstrauss y L. Tzafriri, On the complemented subspace problem, *Israel J. Math* **9** (1971) 263–269.
[33] R. Meise y D. Vogt, *Introduction to Functional Analysis*, Oxford Science Publications, 1997.
[34] R. Merton, The theory of rational option pricing, *Bell. J. Econ. Manage. Sci.,* **4** (1973), 141–183.
[35] J. Neveu, *Discrete parameter martingales*, North-Holland 1975.
[36] R. I. Ovsepian y A. Pełczyński, On the existence of a fundamental total and bounded biorthogonal sequence in every separable Banach space, and related constructions of uniformly bounded orthonormal systems in L^2, *Studia Math.* **54** (1975), no. 2, 149–159.
[37] G. K. Pedersen, *Analysis Now,* Springer-Verlag 1989.
[38] A. M. Plichko y D. Yost, Complemented and uncomplemented subspaces of Banach spaces. III Congress on Banach Spaces (Jarandilla de la Vera, 1998). *Extracta Math.* **15** (2000), 335–371.
[39] J. R. Ringrose, *Compact Non-Self-Adjoint Operators* Van Nostrand Reinhold Company, London, 1971.
[40] W. Rudin, *Principios de Análisis Matemático*, McGraw-Hill 1978.
[41] W. Rudin, *Functional Analysis,* McGraw-Hill 1973.
[42] W. Rudin, *Real and Complex Analysis,* McGraw-Hill 1966.
[43] W. Shachermayer, A Hilbert space proof of the fundamental theorem of asset pricing in finite discrete time, *Insurance Math. Econ.* **11** (1992) 249–257.
[44] H. H. Schaefer, *Topological Vector Spaces,* Springer-Verlag 1971.