Fundamentos de Análisis Matemático

B. Cascales y S. Troyanski

2007 Universidad de Murcia

Índice general

Introducción

ESTE curso de doctorado está diseñado, tanto para iniciarse en la investigación en análisis funcional, como para completar la formación y cultura matemática de aquéllos que vayan a dedicarse a otras especializaciones en análisis matemático, álgebra, geometría y topología, etc. El estudiante recibirá una orientación específica según su progreso académico. Es objetivo primordial del curso introducir y familiarizar al alumno con técnicas profundas del análisis matemático que tienen aplicaciones a cuestiones de diferenciación en espacios de Banach, optimización, teoría general de aproximación, *tests* de convergencia sobre fronteras e integración vectorial. Nos centramos en aquellos aspectos del análisis que constituyen la base necesaria para el estudio de problemas de vigencia actual. Pretendemos que este curso proporcione, a quienes lo sigan con aprovechamiento, una base sólida sobre la que poder iniciar tareas de investigación en las líneas del grupo de *Análisis Funcional* de la Universidad de Murcia.

LAS notas que siguen están organizadas en cinco capítulos pensados como material de trabajo para el alumno. Cada capítulo cuenta con unos objetivos definidos, notas históricas y llamadas de atención marcadas con un signo STOP, en las que se aíslan comentarios que, en una primera lectura, pudieran haber pasado desapercibidos para el lector. Las citas bibliográficas son numerosas, y cada capítulo se cierra con un apartado **PARA SABER MÁS** que contiene comentarios sobre libros y artículos donde poder ampliar los temas desarrollados en el mismo. Los capítulos 3, 4 y 5 contienen una sección titulada *Conexión con el proyecto de investigación*, donde se exponen cuestiones actuales de investigación, algunas de ellas planteadas en el proyecto BFM2002-01719 (2002-2005) financiado por el Ministerio de Ciencia y Tecnología.

El primer capítulo contiene los preliminares de topología y teoría de dualidad que se necesitan para el resto del curso. El capítulo segundo se dedica al estudio y demostración de varias versiones del teorema del punto fijo (Banach, Brouwer, Schauder, Tychonoff), así como a sus aplicaciones a las ecuaciones diferenciales (teoremas de Peano y Picard) y a la existencia de subespacios invariantes para operadores compactos en espacios de Banach (teorema de Lomonosov). En el capítulo tercero estudiamos diversas cuestiones relacionadas con funcionales que alcanzan la norma en conjuntos convexos de un espacio de Banach, y demostramos los teoremas de Krein-Milman, Krein-Šmulian, Minkowski, Bishop-Phelps y el Principio Variacional de Ekeland. Presentamos algunas aplicaciones del teorema de James a la teoría general de la aproximación, y lo ligamos con ciertos *tests* de convergencia y compacidad débil. El cuarto capítulo está dedicado al estudio de la diferenciabilidad y el renormamiento en espacios de Banach. Habida cuenta de que se pueden obtener funciones diferenciables a partir de normas diferenciables, y que estas últimas se pueden construir a partir de normas duales *rotundas*, renormamiento y diferenciabilidad confluyen de forma natural. Demostramos resultados clásicos de renormamiento (Clarkson, Kadec, Klee) en espacios de Banach con dual separable, los cuales nos permiten probar que estos espacios tienen particiones de la unidad de clase C^1. El último capítulo se dedica al estudio de la integración en espacios de Banach (integrales de Bochner y de Pettis). En los capítulos anteriores, algunas de las demostraciones hacen uso, de forma encubierta, de resultados que se pueden aislar en términos de integración vectorial (existencia de baricentros). Aquí proporcionamos una introducción a la integración en espacios de Banach que deja al lector a las puertas de cuestiones de diferenciabilidad de medidas vectoriales que están ligadas, de forma natural, con los problemas de diferenciabilidad y renormamiento desarrollados en el capítulo anterior. Nos remitimos a los cuadros 2.1, 3.1, 4.1 y 5.1 de las páginas 62, 104, 140 y 174, respectivamente, en los que se muestra la interrelación existente entre los resultados centrales estudiados en cada uno de los capítulos.

EL material presentado aquí es, en su mayoría, autocontenido, partiendo de la base de que el alumno ha estudiado cursos básicos de topología conjuntista, cálculo de varias variables, teoría de la medida y análisis funcional. Referencias recomendadas donde encontrar los prerrequisitos del curso son [2, 23, 50, 69, 96].

QUEREMOS insistir en que el contacto de nuestros alumnos con la investigación no se debe reducir únicamente a los cursos de doctorado y al trabajo con su director de tesis. Ambos se deben complementar con la asistencia de los alumnos a las conferencias que, periódicamente, se organizan en el Departamento de Matemáticas y a las sesiones del Seminario de Análisis Funcional.

Finalmente, reseñamos que este curso de doctorado de *4 créditos* es parte del programa de doctorado *Matemáticas* que se imparte en la Universidad de Murcia, y que ha sido distinguido en el curso 2003-2004 con la *Mención de Calidad* del Ministerio de Educación y Ciencia que, en su informe final sobre la articulación y coherencia de los contenidos y estructura general del programa, argumentaba literalmente:

> *«Los objetivos de los cursos están formulados en general con mucha claridad. Los contenidos de los cursos guardan una buena concordancia con los objetivos del programa. Hay una excelente correspondencia entre los contenidos de los cursos y las líneas de investigación. Los contenidos de los cursos son muy adecuados en extensión y profundidad a los créditos asignados. La metodología didáctica propuesta es muy adecuada a los objetivos. Los criterios de evaluación están muy bien definidos.»*

Bernardo Cascales y Stanimir Troyanski

Nombres para la historia

Índice de figuras

Capítulo 1

Preliminares

«OBJETIVOS»

- Fijar una notación y terminología coherentes e inequívocas para el curso.
- Recordar las cuestiones básicas de topología y espacios vectoriales que el alumno debe conocer y que necesitaremos.
- Volver a introducir al alumno en la riqueza que la interrelación de técnicas de topología, análisis y álgebra confieren a los espacios vectoriales topológicos.
- Presentar los espacios vectoriales topológicos como el marco adecuado para estudiar cuestiones relativas a la dualidad, aún cuando éstas sólo se apliquen a los espacios de Banach.

En este capítulo repasamos algunas nociones básicas de topología, álgebra lineal, y análisis funcional, que deben ser conocidas por todos los alumnos, dado que son impartidas en asignaturas troncales de los planes de estudio de la Licenciatura en Matemáticas. En concreto, en esta Universidad, los resultados que se recuerdan en este capítulo preliminar han sido cursados, respectivamente, en las asignaturas:

- *Topología* y *Ampliación de Topología* (Asignaturas Troncales de 6 y 9 Créditos. Primer y segundo cursos de la «Licenciatura en Matemáticas»).
- *Álgebra Lineal y Geometría Euclídea* (Asignatura Troncal de 15 Créditos. Primer curso de la «Licenciatura en Matemáticas»).
- *Análisis Funcional* (Asignatura Troncal de 6 Créditos. 1^{er} Cuatrimestre. Cuarto curso de la «Licenciatura en Matemáticas»).

Nuestras referencias básicas para topología son [48, 69], para álgebra lineal [31, 87] y para análisis funcional [23, 50].

1.1 Espacios topológicos

UNA topología en un conjunto X es una una colección τ de subconjuntos de X (llamados *conjuntos abiertos*) que satisface las propiedades siguientes: *(a)* el total, X, y el conjunto vacío, $\emptyset$, son abiertos; *(b)* la intersección de una cantidad finita de conjuntos abiertos es un conjunto abierto; *(c)* la unión de cualquier colección de abiertos es un abierto. El conjunto X es el *espacio* de la topología τ y el par (X,τ) se llama espacio topológico. Algunas veces, cuando no haya confusión posible, no mencionaremos a τ, y simplemente diremos que X es un *espacio topológico*, tal y como hacemos en toda esta sección.

He aquí un resumen del vocabulario usual que utilizaremos al referirnos a topologías τ y a los espacios topológicos (X,τ).

Comparación de topologías Sean X un conjunto y $\mathfrak{T}$, τ dos topologías en X. Se dice que $\mathfrak{T}$ es más *fina* que τ si los conjuntos τ-abiertos de X son $\mathfrak{T}$-abiertos. En este caso también se dice que τ es más *gruesa* que $\mathfrak{T}$. Se dice que $\mathfrak{T}$ y τ son *comparables* si $\mathfrak{T}$ es más fina que τ, o al revés, τ es más fina que $\mathfrak{T}$.

Cerrados Un conjunto $F \subset X$ es *cerrado* si, y sólo si, su complementario es abierto. Es claro que tanto X como $\emptyset$ son cerrados. Por otra parte, dado que para cada colección $\{A_i\}_{i\in I}$ de subconjuntos de X se tienen las identidades

$$X \setminus \bigcup_{i\in I} A_i = \bigcap_{i\in I}(X\setminus A_i) \quad \text{y} \quad X \setminus \bigcap_{i\in I} A_i = \bigcup_{i\in I}(X\setminus A_i),$$

se obtiene que la intersección de cualquier colección de cerrados es un cerrado y la unión finita de conjuntos cerrados es también un cerrado.

Adherencia La *adherencia* de un subconjunto A de X, denotada por $\overline{A}$, es la intersección de todos los cerrados de X que contienen a A.

Interior El *interior* de un subconjunto A de X, denotado por $\mathring{A}$ o $\operatorname{int} A$, se define como el mayor conjunto abierto de X contenido en A.

De las definiciones anteriores se pueden deducir fácilmente las propiedades que siguen y que el lector comprobará sin dificultad.

Proposición 1.1.1. *Sean (X,τ) un espacio topológico y $A \subset X$. Entonces se verifican las siguientes propiedades:*

i) $\mathring{A} \subset A \subset \overline{A}$.

ii) El interior y la adherencia de A quedan caracterizados por las igualdades

$$\mathring{A} = \left\{x \in X : \text{ existe } V_x \in \tau \text{ tal que } x \in V_x \subset A\right\}$$

y

$$\overline{A} = \left\{x \in X : \text{ para cada } U \in \tau,\ x \in U, \text{ se tiene } U \cap A \neq \emptyset\right\}.$$

iii) *A es abierto si, y sólo si,* $A = \mathring{A}$.
iv) *A es cerrado si, y sólo si,* $A = \overline{A}$.

Topología inducida Si A es un subconjunto de X y τ' es la colección de todas las intersecciones $A \cap U$, con $U \in \tau$, entonces τ' es una topología sobre A, como puede comprobarse fácilmente. A τ' se le llama topología *inducida* en A por τ.

Lema 1.1.2. *Sea $\mathscr{B}$ una familia de subconjuntos de un conjunto X tal que, para cualesquiera miembros $B_1, B_2 \in \mathscr{B}$ y cada $x \in B_1 \cap B_2$, existe $B \in \mathscr{B}$ con $x \in B \subset B_1 \cap B_2$. Entonces, $\emptyset$, X y la familia de uniones de elementos de $\mathscr{B}$ forman una topología τ sobre X.*

Demostración. Es casi evidente que τ es una topología. Sólo probaremos que la intersección finita de miembros de τ está de nuevo en τ. Consideremos $A = \bigcup_{i \in I} A_i$ y $C = \bigcup_{j \in J} C_j$, donde cada $A_i, C_j \in \mathscr{B}$. Para cada $x \in A_i \cap C_j$, tomamos $B_{ij}^x \in \mathscr{B}$ tal que $x \in B_{ij}^x \subset A_i \cap C_j$. Es claro que

$$A \cap C = \bigcup \left\{ B_{ij}^x : i \in I,\, j \in J, x \in X \right\},$$

con lo que $A \cap C \in \tau$. □

Base de una topología Una colección $\mathscr{B} \subset \tau$ es una *base* para la topología τ si cada elemento de τ es unión de elementos de $\mathscr{B}$.

Entornos Un conjunto V se llama *entorno* de un punto $x \in X$ si x pertenece al interior de V. Una colección $\mathscr{B}_x$ de entornos de un punto $x \in X$ es una *base de entornos* para x si todo entorno de x contiene un elemento de $\mathscr{B}_x$. Si, para cada $x \in X$, $\mathscr{B}_x$ es una base de entornos de x, la colección $\{\mathscr{B}_x\}_{x \in X}$ se llama *sistema de entornos* para el espacio X. Todo sistema formado por entornos abiertos satisface las siguientes propiedades:

(E_1).- Para cada $x \in X$, $\mathscr{B}_x \neq \emptyset$, y para cada $U \in \mathscr{B}_x$, $x \in U$.
(E_2).- Si $x \in U \in \mathscr{B}_y$, existe $V \in \mathscr{B}_x$ tal que $V \subset U$.
(E_3).- Para cada $U_1, U_2 \in \mathscr{B}_x$, existe $U \in \mathscr{B}_x$ tal que $U \subset U_1 \cap U_2$.

Base de filtro Una *base de filtro* en un conjunto X es una familia de subconjuntos no vacíos $\mathscr{U}$ de X con la propiedad de que $U \cap V$ contiene un miembro de $\mathscr{U}$, para cada $U, V \in \mathscr{U}$. Si X es un espacio topológico y $\mathscr{B}_x$ es una base de entornos de $x \in X$, entonces $\mathscr{B}_x$ es una base de filtro.

Proposición 1.1.3. *Supongamos que en un conjunto X hemos dado una colección $\{\mathscr{B}_x\}_{x \in X}$ de familias de subconjuntos de X que tienen las propiedades (E_1), (E_2) y (E_3). Sea τ la familia de todos los subconjuntos de X que son uniones de subfamilias de $\bigcup_{x \in X} \mathscr{B}_x$. La familia τ es una topología en X y la colección $\{\mathscr{B}_x\}_{x \in X}$ es un sistema de entornos del espacio topológico (X, τ).*

Demostración. Véase [48, Section 1.2]. □

$\boxed{\textit{Espacios Hausdorff (separados)}}$ Se dice que (X, τ) es un *espacio Hausdorff* (τ es una *topología Hausdorff*) cuando puntos distintos de X tienen entornos disjuntos. Si (X, τ) es un espacio Hausdorff, es fácil convencerse de que cada punto $x \in X$ es un conjunto cerrado.

STOP En lo que sigue, salvo que especifiquemos lo contrario, siempre supondremos que nuestros espacios topológicos son Hausdorff.

Ejemplo 1.1.4 (Espacios métricos). Sea (M, d) un espacio métrico. Sea $\mathscr{B}$ la familia de todas las bolas abiertas de (M, d), *i.e.*, la familia de todos los subconjuntos

$$B(x,a) = \{y \in M : d(x,y) < a\},$$

con $a > 0$ y $x \in M$. La familia $\mathscr{B}$ satisface la condición exigida en el lema 1.1.2, ya que para cualesquiera x, y, $z \in M$ y a, b, $c \in \mathbb{R}$, si $z \in B(x,a) \cap B(y,b)$, se tiene que

$$B(z,c) \subset B(x,a) \cap B(y,b),$$

donde

$$0 < c \leq \text{mín}\{a - d(x,z), b - d(y,z)\}. \qquad \square$$

$\boxed{\textit{Espacios metrizables}}$ Un espacio topológico (X, τ) se dice *metrizable* si existe una métrica d sobre X tal que cada elemento de τ es unión de bolas abiertas. En este caso, diremos que la topología τ es *compatible* con la métrica d.

$\boxed{\textit{Funciones continuas}}$ Sean X e Y espacios topológicos y $f : X \longrightarrow Y$ una función. Se dice que f es *continua en* $x \in X$ si para todo entorno V de $y = f(x)$ se tiene que

$$f^{-1}(V) = \{u \in X : f(u) \in V\}$$

es un entorno de x. La aplicación f se dice *continua de* X *a* Y si es continua en cada $x \in X$. f es continua si, y sólo si, $f^{-1}(V)$ es abierto en X para cada abierto V de Y. Equivalentemente, f es continua si, y sólo si, para cada subconjunto C de X se tiene que $f(\overline{C}) \subset \overline{f(C)}$. La aplicación f se dice que es *abierta de* X *sobre* Y si $f(V)$ es abierto en Y para cada abierto $V \subset X$. La aplicación f se dice que es un *homeomorfismo* si es biyectiva, y tanto ella como su inversa son continuas.

H **Espacios topológicos arbitrarios, 1879-1914.** El gran desarrollo de la topología general estuvo, en sus orígenes, asociado al desarrollo y necesidad de formalizar cuestiones de análisis y geometría. Según R. Engelking, [48, p. 35]: *La topología general debe sus comienzos a una serie de artículos publicados por G. Cantor entre 1879-1884. Al discutir la unicidad de problemas para series trigonométricas, Cantor se concentró en el estudio de conjuntos de puntos excepcionales, donde uno puede quitar algunas hipótesis a un teorema sin* *que éste deje de ser cierto. Después, se dedicó exclusivamente a la investigación de conjuntos, dando lugar, de esta forma, al desarrollo de la teoría de conjuntos y la topología. Cantor introdujo y estudió, en el marco de los espacios Euclídeos, algunas nociones fundamentales de topología. Otros conceptos importantes, también en espacios Euclídeos, fueron introducidos entre 1893-1905 por C. Jordan, H. Poincaré, E. Borel, R. Baire y H. Lebesgue*. En el marco de la teoría general fue decisivo avanzar, desde los espacios Euclídeos hasta los espacios abstractos: nombres como los de B. Riemann, G. Ascoli, C. Arzelá, V. Volterra, D. Hilbert e I. Fredholm, aparecen entre los precursores de conceptos como los de *variedad*, *conjunto de curvas*, *conjunto de funciones*, etc. Los espacios abstractos con una estructura topológica fueron primero introducidos por M. Fréchet y F. Riesz, 1906-1908. Con F. Hausdorff, en 1914, comienza la topología general en el sentido que se le da hoy. Volviendo a la noción de entorno y utilizando las propiedades (E_1), (E_2) y (E_3) introducidas en la página 9, Hausdorff da la primera definición satisfactoria de espacio topológico.

1.1.1 Conjuntos dirigidos y redes

Conjunto dirigido Una relación binaria $\geq$ *dirige* a un conjunto D si D es no vacío y se satisfacen las siguientes propiedades:

a) Si $i, j, k \in D$ son tales que $i \geq j$ y $j \geq k$, entonces $i \geq k$ (propiedad transitiva).
b) Si $i \in D$, entonces $i \geq i$ (propiedad reflexiva).
c) Para cada $i, j \in D$, existe $k \in D$ tal que $k \geq i$ y $k \geq j$.

Un *conjunto dirigido* es un par $(D, \geq)$ tal que $\geq$ dirige a D. Un subconjunto $L \subset D$ se dice *cofinal* en D si, para cada $i \in D$, existe $j \in L$ tal que $j \geq i$; si L es cofinal en D, entonces $(L, \geq)$ es un conjunto dirigido con la relación binaria inducida.

Redes Una *red* en un conjunto X es una aplicación $\phi : D \longrightarrow X$, con dominio un conjunto dirigido $(D, \geq)$. La red se suele denotar mediante $(x_i)_{i \in D}$, donde $x_i = \phi(i)$ para $i \in D$.

Límites Sea $(x_i)_{i \in D}$ una red en X y $A \subset X$. Se dice que $(x_i)_{i \in D}$ está:

– *frecuentemente* en A, si para cada $i \in D$ existe $j \in D$ tal que $j \geq i$ y $x_j \in A$;
– *eventualmente* en A, si no está frecuentemente en $X \setminus A$.

La red $(x_i)_{i \in D}$ en el espacio topológico X se dice *convergente* a un punto $x \in X$ si está eventualmente en cada entorno de x. No es difícil comprobar que X va a ser un espacio Hausdorff si, y sólo si, las redes convergentes convergen a un único punto. Si X es un espacio Hausdorff y $(x_i)_{i \in D}$ es una red convergente a x, se escribe

$$\lim_{i \in D} x_i := x,$$

y se dice que x es el *límite de la red*. Las sucesiones son un caso particular de las redes, tomando $\mathbb{N}$, con su orden total natural, como conjunto dirigido.

Utilizando redes convergentes, se pueden caracterizar las funciones continuas: sean X e Y espacios topológicos y $f : X \longrightarrow Y$ una aplicación. f es continua en $x \in X$ si, y sólo si, para cada red $(x_i)_{i\in D}$ en X convergente a x, se verifica que $\lim_{i\in D} f(x_i) = f(x)$. También podemos caracterizar los subconjuntos abiertos y cerrados en los términos que siguen.

Proposición 1.1.5. *Sea X un espacio topológico. Entonces:*

- *Para cada $F \subset X$ se tiene que*

$$\overline{F} = \big\{x \in X : \text{ existe una red } (x_i)_{i\in D} \text{ en } F \text{ convergente hacia } x\big\}. \tag{1.1}$$

- *Un conjunto $F \subset X$ es cerrado si, y sólo si, ninguna red en F converge a un punto de $X \setminus F$.*
- *$G \subset X$ es abierto si, y sólo si, cada red en X que converge a un punto de G está eventualmente en G.*

Demostración. Sólo probaremos la igualdad (1.1). El resto de propiedades quedan como ejercicio. Es claro que el conjunto descrito en la parte derecha de la igualdad (1.1) está contenido en $\overline{F}$. Al revés, sea $x \in \overline{F}$. Para cada entorno V de x, se tiene que $V \cap F \neq \emptyset$. Elegimos $x_V \in V \cap F$. Dirigimos ahora la familia de entornos $\mathscr{U}_x$ del punto x mediante la relación binaria que sigue: diremos que

$$U \geq V \text{ cuando } U \subset V, \quad U, V \in \mathscr{U}_x.$$

Es claro que la red $(x_V)_{V\in\mathscr{U}_x}$ converge y tiene límite x, y así termina la prueba. □

$\boxed{\textit{Punto de aglomeración}}$ Un punto $x \in X$ se dice que es un *punto de aglomeración* de una red $(x_i)_{i\in D}$ si $(x_i)_{i\in D}$ está frecuentemente en cada entorno de x. Obsérvese que el conjunto de los puntos de aglomeración de $(x_i)_{i\in D}$ (que puede ser vacío) se describe por la intersección

$$C\big((x_i)_{i\in D}\big) = \bigcap_{i\in D} \overline{\{x_j : j \geq i\}}.$$

$\boxed{\textit{Subredes}}$ Sea $(x_i)_{i\in D}$ una red en X. Otra red $(y_\ell)_{\ell\in L}$ se dice que es una *subred* de $(x_i)_{i\in D}$ si existe una aplicación $\sigma : L \longrightarrow D$ tal que

(i) $y_\ell = x_{\sigma(\ell)}$ para cada $\ell \in L$.
(ii) Para cada $i_0 \in D$, existe $\ell_0 \in L$ tal que, si $\ell \geq \ell_0$, entonces $\sigma(\ell) \geq i_0$.

Si una red $(x_i)_{i\in D}$ es convergente hacia $x \in X$, entonces todas sus subredes convergen también al mismo punto.

Lema 1.1.6. *Sea $\mathscr{S}$ una familia de partes de X con la propiedad de que la intersección de dos miembros de $\mathscr{S}$ contiene un miembro de $\mathscr{S}$. Si $(x_i)_{i\in D}$ es una red que está frecuentemente en cada miembro $S \in \mathscr{S}$, entonces existe una subred $(y_\ell)_{\ell\in L}$ de la red $(x_i)_{i\in D}$ que está eventualmente en cada miembro $S \in \mathscr{S}$.*

Demostración. El conjunto de los pares ordenados

$$D \times \mathscr{S} := \big\{(i,S) : i \in D, S \in \mathscr{S}\big\}$$

se dirige por la relación binaria

$$(i,S) \geq (j,T) \text{ si, y sólo si, } i \geq j \text{ y } S \subset T.$$

El conjunto $L = \big\{(i,S) \in D \times \mathscr{S} : x_i \in S\big\}$ es cofinal en $D \times \mathscr{S}$, gracias a que $(x_i)_{i\in D}$ está frecuentemente en cada $S \in \mathscr{S}$. Si consideramos ahora la aplicación $\sigma : L \longrightarrow D$ dada por $\sigma(\ell) = i$, para $\ell = (i,S) \in L$, y definimos $y_\ell = x_{\sigma(\ell)}$, entonces $(y_\ell)_{\ell\in L}$ es una subred de $(x_i)_{i\in D}$ que está eventualmente en cada $S \in \mathscr{S}$. □

Del lema anterior se obtiene inmediatamente el siguiente resultado.

Proposición 1.1.7. *Para una red $(x_i)_{i\in D}$ en el espacio topológico X, las siguientes afirmaciones son equivalentes:*

(i) $x \in X$ es punto de aglomeración de $(x_i)_{i\in D}$.
(ii) Existe una subred $(y_\ell)_{\ell\in L}$ de $(x_i)_{i\in D}$ convergente hacia x.

Convergencia según Moore-Smith, 1922-1955. El concepto de red fue introducido por E. H. Moore y H. L. Smith en 1922. La noción de convergencia en espacios topológicos generales fue descrita por G. Birkhoff en 1937, si bien su presentación tenía algunas deficiencias. La definición correcta de convergencia de redes y subredes, que hoy utilizamos, se debe a J. L. Kelley, quien la presentó en 1950. La teoría de convergencia de filtros, introducida por H. Cartan en 1937 y desarrollada por N. Bourbaki en 1940, es una teoría, a la postre, equivalente a la teoría de convergencia de redes, como fue probado por R. Bartle en 1955. A partir de redes y filtros, los conceptos de ultra-redes y ultra-filtros son herramientas valiosas para demostrar, entre otras cosas, en su total generalidad y de forma elegante y contundente, el teorema de Tychonoff relativo a la compacidad de un producto arbitrario de espacios compactos.

1.1.2 Compacidad

Cubrimiento (o recubrimiento) Sea $A \subset X$. Una familia $\mathscr{C}$ de subconjuntos de X es un *cubrimiento* de A si $A \subset \bigcup\{C : C \in \mathscr{C}\}$. El cubrimiento $\mathscr{C}$ de A se dice abierto si cada $C \in \mathscr{C}$ es abierto en X. Un *subcubrimiento* (o *subrecubrimiento*) de $\mathscr{C}$ es una subfamilia de $\mathscr{C}$ que también es un cubrimiento de A.

Compactos Un subconjunto K de un espacio topológico X se dice que es *compacto* si todo recubrimiento abierto de K en X tiene un subrecubrimiento finito. Si $K \subset X$ es compacto y

$x \notin K$, entonces existen abiertos $U, V \subset X$, con $x \in U$ y $K \subset V$, tales que $U \cap V = \emptyset$. Consecuentemente, los subconjuntos compactos de los espacios Hausdorff son cerrados. Recíprocamente, los subconjuntos cerrados de conjuntos compactos son también compactos.

Proposición 1.1.8. *Sea F un subconjunto cerrado de un conjunto compacto K. Entonces F también es compacto.*

Demostración. Sea $F \subset \bigcup_{i\in I} G_i$, donde cada G_i es abierto. Escribimos $G = X \setminus F$. Entonces, el conjunto G es abierto y $K \subset G \cup (\bigcup_{i\in I} G_i)$. Dado que K es compacto, existen $i_1, \dots, i_n \in I$ tales que $K \subset G \cup (\bigcup_{k=1}^n G_{i_k})$; luego $F \subset \bigcup_{k=1}^n G_{i_k}$, y la prueba concluye. □

$\boxed{\textit{Propiedad de la intersección finita}}$ Se dice que una familia $\{A_i\}_{i\in I}$ de subconjuntos de X tiene la *propiedad de la intersección finita* si, para cada $J \subset I$ finito, se tiene que $\bigcap_{j\in J} A_j \neq \emptyset$.

Cambiando abiertos por cerrados, la compacidad puede ser caracterizada en términos de familias con la propiedad de la intersección finita.

Proposición 1.1.9. *Un subconjunto K de un espacio topológico X es compacto si, y sólo si, cada familia $\{F_i\}_{i\in I}$ de conjuntos cerrados de K con la propiedad de la intersección finita tiene intersección no vacía,* i.e., $\bigcap_{i\in I} F_i \neq \emptyset$.

Demostración. Supongamos que K es compacto y que existe una familia $\{F_i\}_{i\in I}$ de conjuntos cerrados de K, con la propiedad de la intersección finita, tal que $\bigcap_{i\in I} F_i = \emptyset$. Para cada $i \in I$, escribimos $G_i = X \setminus F_i$. Se tiene

$$K \subset X = X \setminus \bigcap_{i\in I} F_i = \bigcup_{i\in I} G_i.$$

Como K es compacto, existen $i_1, \dots, i_n \in I$ tales que $K \subset \bigcup_{k=1}^n G_{i_k}$. Por tanto,

$$\bigcap_{k=1}^n F_{i_k} = X \setminus \bigcup_{k=1}^n G_{i_k} \subset X \setminus K.$$

Por otro lado, cada $F_i \subset K$, y así, $\bigcap_{k=1}^n F_{i_k} = \emptyset$, lo que proporciona una contradicción que termina esta parte de la prueba.

Recíprocamente, supongamos que, para cada familia $\{F_i\}_{i\in I}$ de conjuntos cerrados de K con la propiedad de la intersección finita, se tiene que $\bigcap_{i\in I} F_i \neq \emptyset$. Sea $\{G_i\}_{i\in I}$ un cubrimiento de K por abiertos. Definimos $F_i = K \cap (X \setminus G_i)$. Cada F_i es cerrado y se verifica que

$$\bigcap_{i\in I} F_i = K \cap \Big(\bigcap_{i\in I} (X \setminus G_i) \Big) = K \cap \Big(X \setminus \bigcup_{i\in I} G_i \Big) = \emptyset.$$

Por la hipótesis hecha sobre K, deben existir $i_1, \dots, i_n \in I$ tales que $\bigcap_{k=1}^n F_{i_k} = \emptyset$, y así,

$$\emptyset = \bigcap_{k=1}^n F_{i_k} = K \cap \Big(\bigcap_{k=1}^n (X \setminus G_{i_k}) \Big) = K \cap \Big(X \setminus \bigcup_{k=1}^n G_{i_k} \Big),$$

lo cual implica claramente que $K \subset \bigcup_{k=1}^n G_{i_k}$, y la prueba termina. □

A través de redes, la compacidad puede ser caracterizada en los términos que siguen.

Teorema 1.1.10. *Un subconjunto K de un espacio topológico X es compacto si, y sólo si, cada red de K tiene un punto de aglomeración que pertenece a K.*

Demostración. Supongamos que todas las redes de K tienen un punto de aglomeración que pertenece a K y que K no es compacto. Entonces, existe un cubrimiento abierto $\{G_i\}_{i\in I}$ de K tal que, para cada subconjunto finito $J \subset I$, se tiene que $K \not\subset \bigcup_{j\in J} G_j$. Para cada $J \subset I$ finito, elegimos $x_J \in K \setminus \bigcup_{j\in J} G_j$. Dirigimos la familia de subconjuntos finitos $\mathscr{P}_0(I)$ de I por la siguiente relación binaria:

$$J_1 \geq J_2 \text{ en } \mathscr{P}_0(I) \text{ cuando } J_2 \subset J_1.$$

Por hipótesis, la red $(x_J)_{J\in\mathscr{P}_0(I)}$ tiene un punto de aglomeración $x \in K$. Por un lado, como $\{G_i\}_{i\in I}$ es un cubrimiento de K, existe $k \in I$ tal que $x \in G_k$. Por otro lado, dado que x es punto de aglomeración de $(x_J)_{J\in\mathscr{P}_0(I)}$, existe $J \in \mathscr{P}_0(I)$ tal que $J \geq \{k\}$ (es decir, $k \in J$), con $x_J \in G_k$. Obsérvese que, por construcción, $x_J \notin \bigcup_{j\in J} G_j$ y, sin embargo, como $k \in J$, se tiene que $x_J \notin G_k$; esto proporciona una contradicción que concluye esta parte de la prueba.

Supongamos ahora que K es compacto y sea $(x_i)_{i\in D}$ una red en K. Veamos que el conjunto de los puntos de aglomeración $C\big((x_i)_{i\in D}\big)$ de $(x_i)_{i\in D}$ es no vacío: la familia de cerrados $\big\{\overline{\{x_j : j \geq i\}}\big\}_{i\in D}$ tiene la propiedad de la intersección finita y, consecuentemente, la proposición 1.1.9 puede utilizarse para obtener que

$$\emptyset \neq C\big((x_i)_{i\in D}\big) = \bigcap_{i\in D} \overline{\{x_j : j \geq i\}}. \qquad \square$$

Proposición 1.1.11. *La imagen continua de un espacio topológico compacto es compacta.*

Demostración. Sean X e Y espacios topológicos, con X compacto, y $f : X \longrightarrow Y$ una aplicación continua y sobreyectiva. Sea $\{G_i\}_{i\in I}$ un cubrimiento abierto de Y. Como f es continua, cada $f^{-1}(G_i)$ es abierto en X. Además, $X = \bigcup_{i\in I} f^{-1}(G_i)$. La compacidad de X nos asegura que existen $i_1, \ldots, i_n \in I$ tales que $X = \bigcup_{k=1}^n f^{-1}(G_{i_k})$. Entonces, $Y \subset \bigcup_{k=1}^n G_{i_k}$. $\square$

Corolario 1.1.12. *Sean X e Y espacios topológicos y $f : X \longrightarrow Y$ continua y biyectiva. Si X es compacto, entonces f es un homeomorfismo.*

Demostración. Lo único que hay que probar es que f lleva cerrados a cerrados, lo que se sigue inmediatamente de las proposiciones 1.1.8 y 1.1.11. $\square$

El corolario anterior conduce de inmediato al siguiente.

Corolario 1.1.13. *Sea X un conjunto con dos topologías $\mathfrak{T}$ y τ tal que $(X, \mathfrak{T})$ es compacto. Si $\mathfrak{T}$ es más fina que τ, entonces $\mathfrak{T} = \tau$.*

Proposición 1.1.14. *Sean K_1 y K_2 dos espacios topológicos compactos. Entonces el espacio producto $K_1 \times K_2$ es un espacio topológico compacto.*

Demostración. Sea $(z_i)_{i\in D}$ una red en $K_1 \times K_2$. Para cada $i \in D$ podemos escribir $z_i = (x_i, y_i)$, con $x_i \in K_1$ e $y_i \in K_2$. La red $(x_i)_{i\in D}$ tiene una subred convergente a un punto $x \in K_1$, es decir, existen un conjunto dirigido L_1 y $\sigma_1 : L_1 \longrightarrow D$ con la propiedad de que, para cada $i_0 \in D$, existe $\ell_0 \in L_1$ tal que, si $\ell \geq \ell_0$, entonces $\sigma(\ell) \geq i_0$, de forma que la red $(x_{\sigma(\ell)})_{\ell\in L_1}$ es convergente a x. Si consideramos ahora la red $(y_\ell)_{\ell\in L_1}$, podemos encontrar otro conjunto dirigido L_2 y $\sigma_2 : L_2 \longrightarrow L_1$ con la propiedad de cofinalidad adecuada, de forma que la subred $(y_\ell)_{\ell\in L_2}$ de $(y_i)_{i\in D}$ converge a un punto $y \in K_2$. Es claro que $(x_\ell, y_\ell)_{\ell\in L_2}$ es una subred de $(z_i)_{i\in D}$ que converge hacia (x,y) en $K_1 \times K_2$, que es, consecuentemente, un espacio compacto. □

La demostración anterior vale para un producto finito de espacios compactos. Observamos que de hecho el producto arbitrario de espacios compactos es un espacio compacto, después del teorema de Tychonoff, [69, p. 166-167].

Compacidad, 1894-1923. La génesis de la noción de compacidad está conectada con el teorema de Borel (1894) el cual establece que cada cubrimiento abierto numerable de un intervalo cerrado y acotado tiene un subcubrimiento finito, y con la observación de H. Lebesgue (1903) de que lo mismo se satisface para cualquier cubrimiento abierto no necesariamente numerable. Muchas veces, los conceptos generales de topología tienen un claro antecedente en la correspondiente propiedad de subconjuntos de $\mathbb{R}$. Esto pasó con la noción de compacidad. Durante algún tiempo no estuvo claro si el concepto de compacidad general debía extender las ideas subyacentes detrás del teorema de Borel, la noción de compacidad sucesional o la noción de compacidad numerable. El concepto de espacio (regular) compacto fue introducido por L. Vietoris, en 1921. La noción de compacto que hemos utilizado aquí se debe a P. Alexandroff y P. S. Uryshon, 1923, quienes demostraron muchos de los resultados que se siguen estudiando hoy en día en los textos básicos de topología. Mencionamos que la relación entre compacidad y redes puede encontrarse en [69].

1.1.3 Teorema de Baire

Conjunto raro Sean X un espacio topológico y R un subconjunto de X. Se dice que R es *denso en ninguna parte* o *raro* si su clausura tiene interior vacío, *i.e.*, $\operatorname{int}\overline{R} = \emptyset$.

Conjuntos de Primera y Segunda Categoría Las uniones numerables de conjuntos raros en X se llaman conjuntos de *primera categoría* en X. Los conjuntos que no son de primera categoría en X se llaman de *segunda categoría* en X. Obsérvese que el espacio X es de segunda categoría en sí mismo si, y sólo si, la intersección numerable de abiertos densos es no vacía.

Conjunto G_δ A las intersecciones numerables de abiertos se les da, en topología, el nombre de *conjuntos G_δ*.

$\boxed{\textit{Espacio de Baire}}$ Un espacio topológico se llama *de Baire* si la intersección de cualquier sucesión de abiertos densos es un conjunto denso. Todo espacio de Baire es de segunda categoría en sí mismo.

Teorema 1.1.15 (Teorema de la Categoría de Baire, 1899)**.** *Si (M,d) es un espacio métrico completo, entonces M es un espacio de Baire.*

Demostración. Sea $(G_n)_{n\in\mathbb{N}}$ una sucesión de abiertos densos. Para demostrar que $\bigcap_{n\in\mathbb{N}} G_n$ es denso, probaremos que, para cualquier abierto no vacío $V \subset M$, se tiene que $V \cap \left(\bigcap_{n\in\mathbb{N}} G_n\right)$ es no vacío. Por la densidad de G_1 en M, $G_1 \cap V$ es un abierto no vacío. Tomemos $x_1 \in M$ y $r_1 < 1$ de modo que $B[x_1,r_1] \subset V \cap G_1$. Por la densidad de G_2 en M, existen $x_2 \in M$ y $r_2 < 1/2$ tales que $B[x_2,r_2] \subset G_2 \cap B(x_1,r_1) \subset G_1 \cap G_2 \cap V$. Por inducción, se construyen sucesiones $(x_n)_n$ en M y $(r_n)_n$ con $0 < r_n < 1/n$, de modo que

$$B[x_n,r_n] \subset G_n \cap B(x_{n-1},r_{n-1}) \subset G_1 \cap G_2 \cap \ldots \cap G_n \cap V.$$

La sucesión $(x_n)_n$ así construida es de Cauchy, puesto que $B(x_n,r_n) \subset B[x_n,r_n] \subset B(x_{n-1},r_{n-1})$ y $\lim_n r_n = 0$. Y si denotamos por x el límite de $(x_n)_n$, se tiene que $x \in B[x_n,r_n] \subset G_n \cap V$ para cada n, debido a que $B[x_n,r_n]$ es cerrado. Es decir, $x \in V \cap \left(\bigcap_{n\in\mathbb{N}} G_n\right)$. $\square$

Las mismas ideas que aparecen en la demostración del teorema anterior sirven para probar que todo espacio topológico localmente compacto es de Baire.

Conviene observar que, si el espacio M no es completo, el teorema anterior no es cierto. Por ejemplo, para $M = \mathbb{Q} = \{q_n : n \in \mathbb{N}\}$ con la métrica inducida por la de $\mathbb{R}$, se cumple que $G_n := M \setminus \{q_n\}$ es un abierto denso, mientras que $\bigcap_{n\in\mathbb{N}} G_n = \emptyset$.

Teorema de la Categoría, 1899-1937. R. Baire demostró el teorema de la Categoría 1.1.15 para la recta real en 1899. F. Hausdorff extendió el resultado a los espacios completamente metrizables, 1914. E. Čech estableció, en 1937, que un espacio es completamente metrizable si, y sólo si, es metrizable y es un G_δ en alguna de sus compactificaciones. Los espacios que son G_δ en alguna de sus compactificaciones son conocidos como espacios Čech-completos; en los espacios Čech-completos se satisface el teorema de la Categoría de Baire. El teorema de la Categoría aplicado a espacios de funciones y a espacios de subconjuntos cerrados es una herramienta eficiente para establecer la existencia de algunos objetos matemáticos, como por ejemplo, la existencia de funciones continuas en $\mathbb{R}$ que no son derivables en ningún punto. El teorema de la Categoría es la herramienta que permite demostrar los teoremas de Gráfica Cerrada, Acotación Uniforme y Aplicación Abierta para espacios de Banach en análisis funcional.

1.2 Espacios vectoriales

MEDIANTE $\mathbb{K}$ representaremos, o bien el cuerpo de los números reales $\mathbb{R}$, o bien el cuerpo de los números complejos $\mathbb{C}$. En tal caso, se define un *escalar* como un elemento del *cuerpo de escalares* $\mathbb{K}$.

Espacio vectorial Un *espacio vectorial sobre* $\mathbb{K}$ es un conjunto E, cuyos elementos se llaman *vectores*, sobre el que se definen dos operaciones: la *suma* de vectores (a cada par de vectores (x,y) le corresponde un vector $x+y$) y la *multiplicación* de vectores por escalares (a cada par (α,x), con $\alpha \in \mathbb{K}$ y $x \in E$, le corresponde un vector αx), satisfaciendo las propiedades:

$$x+y = y+x, \text{ para cada } x,y \in E;$$
$$x+(y+z) = (x+y)+z, \text{ para cada } x,y,z \in E;$$
$$E \text{ contiene un único vector } 0 \text{ tal que } x+0 = x, \text{ para cada } x \in E;$$
$$\text{para cada } x \in E, \text{ existe un único } -x \in E \text{ tal que } x+(-x) = 0;$$
$$1x = x, \text{ para cada } x \in E;$$
$$\alpha(\beta x) = (\alpha\beta)x, \text{ para cada } \alpha,\beta \in \mathbb{K} \text{ y } x \in E;$$
$$\alpha(x+y) = \alpha x+\alpha y, \text{ para cada } \alpha \in \mathbb{K} \text{ y } x,y \in E;$$
$$(\alpha+\beta)x = \alpha x+\beta x, \text{ para cada } \alpha,\beta \in \mathbb{K} \text{ y } x \in E.$$

A 0 se le llama elemento *neutro* y $-x$ se denomina elemento *opuesto* de x.
Un *espacio vectorial real* es un espacio vectorial sobre $\mathbb{K} = \mathbb{R}$. Un *espacio vectorial complejo* es un espacio vectorial sobre $\mathbb{K} = \mathbb{C}$.

STOP Debe sobrentenderse que toda afirmación sobre espacios vectoriales en la que no se especifique el cuerpo de escalares, es válida para espacios vectoriales reales y complejos.

Suma de Minkowski Si E es un espacio vectorial, $A,B \subset E$, $x \in E$ y $\lambda \in \mathbb{K}$, usamos la notación:

- $x+A = \{x+a : a \in A\}$;
- $x-A = \{x-a : a \in A\}$;
- $A+B = \{a+b : a \in A,\, b \in B\}$ (*suma de Minkowski*);
- $\lambda A = \{\lambda a : a \in A\}$.

Mediante $-A := (-1)A$ designamos el conjunto de los opuestos de los elementos de A.

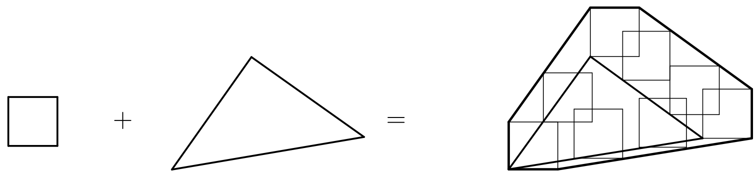


Figura 1.1: Suma de Minkowski

$\boxed{\textit{Subespacio vectorial}}$ Se dice que un subconjunto $F \subset E$ es un *subespacio* vectorial de E si F es un espacio vectorial respecto de las mismas operaciones inducidas por E. Se comprueba fácilmente que esto ocurre si, y sólo si,

$$\alpha F + \beta F \subset F, \text{ para cualesquiera } \alpha, \beta \in \mathbb{K}.$$

$\boxed{\textit{Conjuntos equilibrados}}$ Se dice que un conjunto $A \subset E$ es *equilibrado* si $\alpha A \subset A$ para todo $\alpha \in \mathbb{K}$ tal que $|\alpha| \leq 1$.

$\boxed{\textit{Conjuntos convexos}}$ Se dice que un conjunto $A \subset E$ es *convexo* si

$$\lambda A + (1-\lambda)A \subset A, \text{ para cada } 0 \leq \lambda \leq 1.$$

El conjunto $A \subset E$ se dice que es *absolutamente convexo* si A es convexo y equilibrado, o, equivalentemente, si

$$\alpha A + \beta A \subset A \text{ para cada } \alpha, \beta \in \mathbb{K} \text{ con } |\alpha| + |\beta| \leq 1.$$

La intersección de conjuntos convexos (respectivamente, absolutamente convexos) es convexo (respectivamente, absolutamente convexo). Así, en cualquier espacio vectorial E, si $A \subset E$, entonces existe un conjunto convexo (respectivamente, absolutamente convexo) más pequeño en E que contiene a A (*i.e.*, la intersección de todos los convexos –respectivamente, absolutamente convexos– que contienen a A) y que denotaremos por $\operatorname{co}(A)$ (respectivamente, $\Gamma(A)$). Es fácil comprobar que si A es equilibrado, entonces su *envoltura convexa* $\operatorname{co}(A)$ es un conjunto absolutamente convexo.

$\boxed{\textit{Combinación lineal}}$ Si $x_1, x_2, \ldots, x_n$ son vectores en E, diremos que un elemento de la forma $\lambda_1 x_1 + \ldots + \lambda_n x_n$, donde $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$, es una *combinación lineal* de los vectores $x_i \in E$, $i = 1, \ldots, n$. Las sumas anteriores suelen escribirse de forma más breve como $\sum_{i=1}^{n} \lambda_i x_i$.

$\boxed{\textit{Espacio generado}}$ Si A es un subconjunto de E, entonces $\operatorname{span} A$ representa el conjunto de todas las combinaciones lineales de elementos de A. Es claro que $\operatorname{span} A$ es un subespacio de E, el cual se denomina *subespacio vectorial generado* por A.

$\boxed{\textit{Conjuntos linealmente independientes}}$ Un subconjunto A de E se dice *linealmente independiente* si, para todo subconjunto finito no vacío $\{x_i : i = 1, \ldots, n\}$ de A, la igualdad

$$\sum_{i=1}^{n} \lambda_i x_i = 0$$

implica que

$$\lambda_i = 0, \text{ para todo } i = 1, \ldots, n.$$

$\boxed{\textit{Conjuntos ordenados}}$ Un elemento m de un conjunto $(A, \leq)$ con orden parcial se llama *maximal* si, cuando $\alpha \in A$, con $m \leq \alpha$, se tiene que $m = \alpha$. Un conjunto $(A, \leq)$ con una relación de orden se dice *totalmente ordenado*si, para cada $\alpha, \beta \in A$, se tiene que $\alpha \leq \beta$ ó $\beta \leq \alpha$.

Lema 1.2.1 (Kuratowski-Zorn, [69, p. 45-46])**.** *Sea A un conjunto con orden parcial tal que cada subconjunto B de A totalmente ordenado tiene un elemento maximal. Entonces A tiene, al menos, un elemento maximal.*

$\boxed{\textit{Base}}$ Un subconjunto B de un espacio vectorial E se dice que es una *base* (base de Hamel) de E si es linealmente independiente y $E = \operatorname{span} B$. Equivalentemente, B es base de E si es linealmente independiente y maximal con respecto a la inclusión de conjuntos. Todo espacio vectorial tiene una base como consecuencia del lema de Zorn 1.2.1.

$\boxed{\textit{Dimensión}}$ Dos bases cualesquiera de E tienen el mismo cardinal, que se llama *dimensión* de E sobre $\mathbb{K}$. Un espacio vectorial E tiene *dimensión* $n \in \mathbb{N}$ ($\dim E = n$) si tiene una *base* $\{u_1, \ldots, u_n\}$ con n elementos. Esto quiere decir que cada $x \in E$ tiene una única representación de la forma

$$x = \alpha_1 u_1 + \ldots + \alpha_n u_n, \text{ con } \alpha_i \in \mathbb{K}, \; i = 1, \ldots, n.$$

Si $\dim E = n$ para algún $n \in \mathbb{N}$, se dice que E tiene *dimensión finita*. Si $E = \{0\}$, entonces $\dim E = 0$.

$\boxed{\textit{Dual algebraico}}$ El *espacio dual* de un espacio vectorial E es el conjunto $E^{\#}$ cuyos elementos son las formas lineales en E, *i.e.*, las aplicaciones $g : E \longrightarrow \mathbb{K}$ satisfaciendo

$$g(x+y) = g(x) + g(y), \qquad g(\alpha x) = \alpha g(x), \text{ para cada } x, y \in E \text{ y } \alpha \in \mathbb{K}.$$

Nótese que $E^{\#}$ dotado de la adición y multiplicación por escalares definidas mediante

$$(g_1 + g_2)(x) = g_1(x) + g_2(x), \qquad (\alpha g_1)(x) = \alpha g_1(x),$$

se convierte en un espacio vectorial.

Lema 1.2.2. *Sea E un espacio vectorial. Entonces $E^{\#}$ tiene dimensión finita si, y sólo si, E tiene dimensión finita, y* $\dim E = \dim E^{\#}$.

Demostración. Sea $\dim E = n$. Entonces, existe $\{e_1, \ldots, e_n\}$ tal que cada $x \in E$ tiene una única representación de la forma $x = \sum_{i=1}^n f_i(x) e_i$. Por la unicidad de la representación, se tiene que

$$f_i(ax + by) = a f_i(x) + b f_i(y), \text{ para cada } a, b \in \mathbb{K} \text{ y } x, y \in E.$$

Así, se tiene que $f_i \in E^{\#}$, para $i = 1, \ldots, n$. Si tomamos $f \in E^{\#}$, entonces, para cada $x \in E$,

$$f(x) = \sum_{i=1}^n f(e_i) f_i(x),$$

luego $f = \sum_{i=1}^n f(e_i) f_i$. En consecuencia, $\dim E^{\#} \leq n = \dim E$.

Recíprocamente, supongamos que $\dim E^{\#} = n$. Como $E \subset (E^{\#})^{\#}$, obtenemos, usando lo que acabamos de probar, que $\dim E \leq \dim (E^{\#})^{\#} \leq \dim E^{\#}$, y así acaba la demostración. $\square$

Lema 1.2.3. *Sea $\{f_i : i = 1,\dots,n\}$ un subconjunto linealmente independiente de $E^\#$. Entonces, existe $\{u_j : j = 1,\dots,n\}$ en E tal que $f_i(u_j) = \delta_{ij}$, para $i, j = 1,\dots,n$.*

Demostración. Haremos la demostración por inducción sobre n.

Si $n = 1$ y $f_1 \not\equiv 0$, tomamos $e_1 \in E$ tal que $f_1(e_1) \neq 0$, y $u_1 = e_1/f_1(e_1)$.

Tomemos ahora $f_1,\dots,f_n \in E^\#$ linealmente independientes, y supongamos, por hipótesis de inducción, que la tesis del lema es cierta para $f_1,\dots,f_{n-1} \in E^\#$. Fijemos $e_1,\dots,e_{n-1} \in E$ tales que $f_i(e_j) = \delta_{ij}$, $i, j = 1,\dots,n-1$, y para cada $x \in E$ escribamos $y_x = \sum_{i=1}^{n-1} f_i(x)e_i$. Si $f_n(x - y_x) = 0$ para cada $x \in E$, entonces $f_n \in \operatorname{span}\{f_i : i = 1,\dots,n-1\}$, con lo que $f_1,\dots,f_n$ no serían linealmente independientes, contradiciendo nuestra hipótesis. Existe por tanto un $x \in E$ tal que $f_n(x - y_x) \neq 0$. Si definimos

$$u_n := \frac{x - y_x}{f_n(x - y_x)},$$

se tiene entonces que $f_i(u_n) = 0$, si $i < n$, y $f_n(u_n) = 1$. La prueba concluye tomando, para cada $i = 1,\dots,n-1$, $u_i = e_i - f_n(e_i)u_n$. □

Núcleo Para $f \in E^\#$, se define el *núcleo* de f como

$$\ker f := \{x \in E : f(x) = 0\}.$$

Es claro que $\ker f$ es un subespacio de E. Si f es no nula, entonces $\ker f$ es un subespacio propio maximal, o, equivalentemente, $\dim(E/\ker f) = 1$.

Lema 1.2.4. *Sean $f_1,\dots,f_n, f$ formas lineales en E que satisfacen*

$$\bigcap_{i=1}^{n} \ker f_i \subset \ker f.$$

Entonces $f \in \operatorname{span}\{f_1,\dots,f_n\}$.

Demostración. Si $f_i = 0$ para $i = 1,\dots,n$, entonces $\ker f_i = E$ para $i = 1,\dots,n$. Así, se tiene que $E = \bigcap_{i=1}^{n} \ker f_i \subset \ker f$, y por tanto, $f = 0$.

Supongamos ahora que $f_1,\dots,f_k$ son linealmente independientes y que $f_j \in \operatorname{span}\{f_1,\dots,f_k\}$ para $j = k+1,\dots,n$. Entonces,

$$\bigcap_{i=1}^{k} \ker f_i \subset \bigcap_{i=1}^{n} \ker f_i \subset \ker f. \tag{1.2}$$

Supongamos que $f \notin \operatorname{span}\{f_1,\dots,f_k\}$. Entonces, $f, f_1,\dots,f_k$ son linealmente independientes. El lema 1.2.3 nos asegura la existencia de $u \in E$ tal que $f(u) = 1$ y $f_i(u) = 0$, para cada $i = 1,\dots,k$. La inclusión (1.2) nos dice que $f(u) = 0$, llegando a una contradicción que termina la prueba. □

Espacios vectoriales, 1640-1932. La discusión de $\mathbb{R}^2$ y $\mathbb{R}^3$ como espacios de coordenadas se remonta hasta R. Descartes y P. Fermat. La noción de vector ya fue expuesta por B. Bolzano, y la idea de suma de vectores aparece implícitamente, en 1799, en trabajos de C. F. Gauss sobre la representación geométrica que hace de los imaginarios y la aplicación de ellos a la geometría elemental. Nombres como los de A. Cayley y H. Grassman aparecen asociados a las extensiones de las ideas que se expresan en espacios de 2 ó 3 coordenadas a ideas que se expresan en espacios de n coordenadas. Grassman, por ejemplo, introduce en estos últimos las nociones de independencia lineal o dimensión, y establece la relación fundamental $\dim(E)+\dim(F)=\dim(E+F)+\dim(E\cap F)$. Fue G. Peano, en 1888, quien apreció en todo su valor la obra de Grassman y dio la definición axiomática de los espacios vectoriales (de dimensión finita o no) sobre el cuerpo de los números reales, junto con la definición de aplicación lineal. Las técnicas de álgebra lineal asisten con éxito a cuestiones del análisis, y matemáticos de la talla de D. Hilbert las utilizaron con éxito desde un principio. En la década 1920-30, S. Banach tuvo la brillante idea de combinar técnicas de topología conjuntista con técnicas de álgebra lineal, obteniendo resultados tan potentes como los teoremas de Banach-Steinhaus, de Gráfica Cerrada y de Aplicación Abierta. Las ramas de topología, álgebra lineal y análisis funcional continúan beneficiándose mutuamente desde los principios del siglo XX.

1.3 Dualidad

En esta sección queremos introducir al lector en la riqueza que la interrelación de técnicas de topología, análisis y álgebra, presentando los espacios vectoriales topológicos como el marco adecuado para estudiar cuestiones relativas a la dualidad, que serán utilizadas con profusión en el contexto de los espacios de Banach en capítulos posteriores.

Funcionales subaditivos, seminormas y normas Sean E un espacio vectorial sobre el cuerpo $\mathbb{K}$ y $q: E \longrightarrow \mathbb{R}$ una función. Diremos que:

(i) q es *subaditiva* si $q(x+y) \leq q(x)+q(y)$, para todo $x,y \in E$.

(ii) q es *positivamente homogénea* si $q(\alpha x) = \alpha q(x)$, para todo x de E y $\alpha > 0$.

(iii) q es *sublineal* si es subaditiva y positivamente homogénea.

(iv) q es una *seminorma* si q es subaditiva y $q(\alpha x) = |\alpha| q(x)$, para todo $x \in E$ y $\alpha \in \mathbb{K}$.

(v) q es una *norma* si q es una seminorma y además la ecuación $q(x) = 0$ sólo tiene la solución $x = 0$.

Espacios normados y de Banach Un *espacio normado* es un espacio vectorial X dotado de una norma $\|\cdot\|$. Un espacio normado $(X, \|\cdot\|)$ se llama *espacio de Banach* si la distancia asociada a la norma, $d: X \times X \longrightarrow \mathbb{R}$, mediante la fórmula $d(x,y) = \|x-y\|$, es completa, es decir, si cada sucesión de Cauchy en (X,d) converge a un punto perteneciente a X.

La topología asociada a una norma es compatible con la estructura de espacio vectorial, en el sentido de que las operaciones *suma* y *producto por escalares* son continuas: esto se sigue directamente del hecho de que la norma satisface la desigualdad triangular y permite *sacar* escalares fuera en valor absoluto. En otras palabras, un espacio normado es un espacio vectorial topológico.

Espacios vectoriales topológicos Una *topología vectorial* en un espacio vectorial E es una topología $\mathfrak{T}$ para la cual las aplicaciones

$$s : (E,\mathfrak{T}) \times (E,\mathfrak{T}) \longrightarrow (E,\mathfrak{T}) \quad \text{y} \quad p : \mathbb{K} \times (E,\mathfrak{T}) \longrightarrow (E,\mathfrak{T}),$$

definidas por $s(x,y) = x+y$ y $p(\lambda,x) = \lambda x$ son continuas. Un *espacio vectorial topológico* es un espacio vectorial E dotado de una topología vectorial Hausdorff $\mathfrak{T}$. En lo que sigue, utilizaremos *e.v.t.* como abreviatura para espacio vectorial topológico, y escribiremos, indistintamente, $(E,\mathfrak{T})$ o $E[\mathfrak{T}]$ para denotarlo.

STOP Salvo que se especifique lo contrario, todos los *e.v.t.* se supondrán Hausdorff, aunque pueda ocurrir que algunas topologías vectoriales no lo sean.

Proposición 1.3.1. *Sea $E[\mathfrak{T}]$ un* e.v.t. *Entonces:*

(i) Para $a \in E$ y $\lambda \in \mathbb{K}$, $\lambda \neq 0$, las aplicaciones

$$s_a : E[\mathfrak{T}] \longrightarrow E[\mathfrak{T}] \quad y \quad p_\lambda : E[\mathfrak{T}] \longrightarrow E[\mathfrak{T}],$$

definidas por $s_a(x) = x+a$ y $p_\lambda(x) = \lambda x$ son homeomorfismos.

(ii) Si $\mathscr{U}$ es una base de entornos del origen en $E[\mathfrak{T}]$, $x \in E$ y $\alpha \neq 0$, entonces

$$x + \alpha\mathscr{U} := \{x + \alpha U : U \in \mathscr{U}\}$$

es una base de entornos de x.

(iii) Si $A \subset E$ es abierto y $B \subset E$ es un subconjunto cualquiera, entonces $A+B$ es abierto.

(iv) Si $A \subset E$ es compacto y $B \subset E$ es cerrado, entonces $A+B$ es cerrado.

(v) Si $F \subset E$ es un subespacio vectorial, su clausura $\overline{F}$ es un subespacio vectorial.

(vi) Sea $\mathscr{U}$ una base de entornos del origen de $E[\mathfrak{T}]$. Si A es un subconjunto de E, entonces se tiene que $\overline{A} = \bigcap_{U\in\mathscr{U}}(A+U)$.

(vii) Si $F[\tau]$ es otro e.v.t. *y $T : E \longrightarrow F$ es lineal, entonces T es continua si, y sólo si, T es continua en* 0. *Cuando $F = (\mathbb{K}, |\cdot|)$, T es continua si, y sólo si,* $\ker T$ *es cerrado en $E[\mathfrak{T}]$.*

Demostración. La demostración de las cinco primeras propiedades la dejamos como ejercicio. Para probar *(vi)* obsérvese que, por definición de clausura, se tiene que

$$\overline{A} \subset \bigcap_{U\in\mathscr{U}} (A+U).$$

Recíprocamente, si $x \in \bigcap_{U\in\mathscr{U}}(A+U)$, entonces, para cada $U \in \mathscr{U}$, existen $a_U \in A$ e $y_U \in U$ tales que $x = a_U + y_U$. Ordenando $\mathscr{U}$ por la relación binaria

$$U_1 \geq U_2 \text{ si, y sólo si, } U_1 \subset U_2,$$

para $U_1, U_2 \in \mathscr{U}$, es claro que la red $(y_U)_U \to 0$, y por lo tanto, $(a_U)_U$ converge a un cierto $a \in \overline{A}$, lo que implica que $x \in \overline{A}$.

La primera parte de la propiedad *(vii)* es sencilla de establecer. Sólo prestaremos atención a la prueba del caso $F = (\mathbb{K}, |\cdot|)$ para ver cómo la hipótesis $\ker T$ cerrado implica la continuidad de T. Procedamos por reducción al absurdo, y supongamos que $\ker T$ es cerrado y que T no es continua. Existe, por tanto, una red $(x_i)_{i\in D}$ en $E[\mathfrak{T}]$ tal que $x_i \to 0$ pero $T(x_i) \not\to 0$. La última condición significa que para algún $\varepsilon > 0$, el conjunto $J = \{i \in D : |T(x_i)| > \varepsilon\}$ es cofinal en $(D, \geq)$. Por lo tanto, la red $(x_j)_{j\in J}$ también converge a cero. Como $T \neq 0$, podemos tomar $a \in E$ tal que $T(a) = 1$, y así, si para cada $j \in J$ definimos

$$z_j = \frac{x_j}{T(x_j)} - a,$$

se tiene que $z_j \in \ker T$ y que $z_j \to (-a) \notin \ker T$. Por lo tanto, $\ker T$ no es cerrado, y hemos llegado a la contradicción que acaba la prueba. □

Conjunto absorbente Un conjunto $A \subset E$ se dice que es *absorbente* si, para cada $x \in E$, existe $\rho_0 > 0$ tal que $x \in \rho A$, para $|\rho| \geq \rho_0$.

Proposición 1.3.2. *Si $E[\mathfrak{T}]$ es un* e.v.t. *y $\mathscr{U}$ una base de entornos del origen para $\mathfrak{T}$, entonces:*

(i) Para cada $U \in \mathscr{U}$, existe $V \in \mathscr{U}$ tal que $V + V \subset U$.
(ii) Para cada $U \in \mathscr{U}$, existe $V \in \mathscr{U}$ tal que $\alpha V \subset U$, para cada $|\alpha| \leq 1$.
(iii) Cada $U \in \mathscr{U}$ es absorbente.

En particular,

$$\widetilde{\mathscr{U}} = \left\{\textstyle\bigcup_{|\alpha|\leq 1} \alpha U : U \in \mathscr{U}\right\} \quad y \quad \overline{\mathscr{U}} = \left\{\overline{U} : U \in \mathscr{U}\right\}$$

son bases de entornos del origen en $E[\mathfrak{T}]$. Así, cada e.v.t. *tiene una base de entornos del origen formada por conjuntos absorbentes, equilibrados y cerrados.*

Demostración. La propiedad *(i)* es consecuencia de la continuidad de la suma $s : E \times E \longrightarrow E$ en el punto $(0,0)$. Dado $x \in E$, *(iii)* se deduce de la continuidad de la aplicación $p_x : \mathbb{K} \longrightarrow E[\mathfrak{T}]$, dada por $p_x(\lambda) = \lambda x$, en $\lambda = 0$. Para demostrar *(ii)*, dado $U \in \mathscr{U}$, la continuidad del producto por escalares $p : \mathbb{K} \times E[\mathfrak{T}] \longrightarrow E[\mathfrak{T}]$ en $(0,0)$ nos asegura la existencia de una bola $B[0,\varepsilon] \subset \mathbb{K}$ y de un $W \in \mathscr{U}$ tales que $B[0,\varepsilon] \cdot W \subset U$. El conjunto $V := \varepsilon W$ satisface la propiedad requerida en *(ii)*.

La familia $\widetilde{\mathscr{U}}$ es una base de entornos del origen gracias a *(ii)*. La familia $\overline{\mathscr{U}}$ es una base de entornos del origen gracias a *(i)* y a la propiedad *(vi)* de la proposición 1.3.1. Es claro ahora que, tomando los cierres de los elementos de $\widetilde{\mathscr{U}}$, se consigue una base de entornos del origen formada por conjuntos equilibrados, cerrados y absorbentes en E. □

Proposición 1.3.3. *Sea E un espacio vectorial y sea $\mathscr{U}$ una base de filtro verificando:*

(i) *Cada $U \in \mathscr{U}$ es absorbente y equilibrado.*
(ii) $\bigcap\{U : U \in \mathscr{U}\} = \{0\}$.
(iii) *Para cada $U \in \mathscr{U}$, existe $V \in \mathscr{U}$ tal que $V + V \subset U$.*

Si para cada $x \in E$ consideramos $\mathscr{U}_x = \{x + U : U \in \mathscr{U}\}$, entonces existe una única topología vectorial $\mathfrak{T}$ tal que $\mathscr{U}_x$ es base de entornos de x, para cada $x \in E$.

Demostración. La familia $\{\mathscr{U}_x\}_{x\in E}$ satisface las propiedades (E_1), (E_2) y (E_3), y por tanto, véase la proposición 1.1.3, existe una única topología para la que $\{\mathscr{U}_x\}_{x\in E}$ es un sistema de entornos. Se comprueba que $\mathfrak{T}$ es una topología vectorial. □

Topología asociada a una familia de seminormas Sea $E[\mathfrak{T}]$ un *e.v.t.* Se dice que $\mathfrak{T}$ está *asociada a una familia de seminormas* $\mathscr{P}$ si la familia

$$\widetilde{\mathscr{U}} = \left\{ \bigcap_{i=1}^{n} \big\{x \in E : p_i(x) < \varepsilon\big\} : p_1, \dots, p_n \in \mathscr{P},\ \varepsilon > 0,\ n \in \mathbb{N} \right\} \tag{1.3}$$

es una base de entornos del origen para $\mathfrak{T}$; obsérvese que, en este caso, para la familia $\mathscr{P}$ se satisface $\bigcap_{p\in\mathscr{P}}\big\{x \in E : p(x) = 0\big\} = \{0\}$.

Es útil tener presente que si $\mathfrak{T}$ está asociada a $\mathscr{P}$, entonces la convergencia de una sucesión (o red) $x_i \xrightarrow{\mathfrak{T}} x$ es equivalente a la condición $p(x_i - x) \to 0$, para toda seminorma $p \in \mathscr{P}$.

Espacios localmente convexos Un *espacio localmente convexo* (brevemente *e.l.c.*) es un *e.v.t.* $E[\mathfrak{T}]$ cuya topología $\mathfrak{T}$ tiene una base de entornos del origen formada por conjuntos convexos. En este caso se dice también que $\mathfrak{T}$ es una topología localmente convexa.
Utilizando la proposición 1.3.2, es fácil convencerse de que todo *e.l.c.* tiene una base de entornos del origen formada por conjuntos absolutamente convexos y cerrados. Si $(X, \|\cdot\|)$ es un espacio normado y $\mathfrak{T}$ es la topología asociada a la norma $\|\cdot\|$, entonces $(X, \mathfrak{T})$ es un *e.l.c.* Más en general, si $(E, \mathfrak{T})$ es un *e.v.t.* cuya topología está asociada a una familia de seminormas, entonces $(E, \mathfrak{T})$ es un *e.l.c.*

Como veremos en las páginas siguientes, *e.l.c.* y *e.v.t.* cuyas topologías están asociadas a familias de seminormas son, en realidad, una misma cosa, teorema 1.3.7.

Proposición 1.3.4. *Sea $A \subset E$ un conjunto absorbente. Para cada $x \in E$, definimos*

$$p_A(x) := \inf\big\{t > 0 : x \in tA\big\}.$$

El funcional $p_A : E \longrightarrow \mathbb{R}$ es no negativo y positivamente homogéneo. Si A es convexo, entonces p_A es sublineal, y se tiene que

$$\big\{x \in E : p_A(x) < 1\big\} \subset A \subset \big\{x \in E : p_A(x) \leq 1\big\}. \tag{1.4}$$

Si además A es absolutamente convexo, entonces p_A es una seminorma.

Demostración. Para A absorbente, es claro que p_A está bien definido y es no negativo. Por otro lado, $0 \in A$, y así, $p_A(0) = 0$. Para $\sigma = 0$ se tiene que $p_A(\sigma x) = \sigma p_A(x)$. Para $\sigma > 0$ tenemos que

$$p_A(\sigma x) = \text{ínf}\{t > 0 : \sigma x \in tA\} = \text{ínf}\Big\{t > 0 : x \in \frac{t}{\sigma}A\Big\} = \sigma p_A(x),$$

y así, p_A es positivamente homogéneo.

Supongamos ahora que A es convexo y tomemos $x, y \in E$. Sean $\sigma > p_A(x)$ y $\rho > p_A(y)$. Elijamos σ' y ρ' de forma que $\sigma > \sigma' > p_A(x)$ y $\rho > \rho' > p_A(y)$, satisfaciendo $\frac{x}{\sigma'}, \frac{y}{\rho'} \in A$. Se tiene entonces que

$$\frac{x+y}{\sigma'+\rho'} = \frac{\sigma'}{\sigma'+\rho'}\frac{x}{\sigma'} + \frac{\rho'}{\sigma'+\rho'}\frac{y}{\rho'} \in A$$

y, en consecuencia, podemos concluir que

$$p_A(x+y) \leq p_A(x) + p_A(y).$$

Obsérvese que si A es equilibrado y $t > 0$, entonces la condición $\lambda x \in tA$ equivale a $|\lambda| x \in tA$, para cada $\lambda \in \mathbb{K}$. Así, cuando A es absolutamente convexo, tenemos que, para $\lambda \in \mathbb{K}$, $\lambda \neq 0$, la igualdad

$$p_A(\lambda x) = \text{ínf}\{t > 0 : x \in |\lambda|^{-1} tA\} = |\lambda|\, \text{ínf}\{|\lambda|^{-1} t : x \in |\lambda|^{-1} tA\} = |\lambda| p_A(x)$$

establece que p_A es una seminorma. Nos queda sólo demostrar las inclusiones (1.4). Es claro que $A \subset \{x \in E : p_A(x) \leq 1\}$. Por otro lado, si A es absorbente y convexo y $x \in \lambda A$ para $\lambda > 0$, entonces $x \in \rho A$ para $\rho > \lambda$, de donde se sigue que $\{x \in E : p_A(x) < 1\} \subset A$. $\square$

$\boxed{\textit{Funcional de Minkowski}}$ Si E es un espacio vectorial y $A \subset E$ es absorbente, p_A se denomina el *funcional de Minkowski* asociado a A. Si p es una seminorma y tomamos la bola unidad $A = \{x \in E : p(x) \leq 1\}$, entonces se tiene que $p_A = p$.

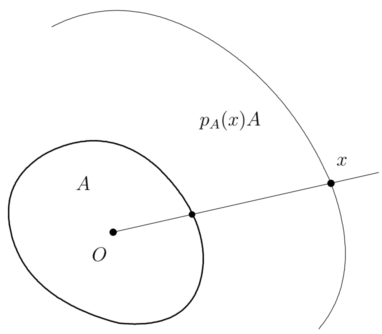


Figura 1.2: Funcional de Minkowski

Proposición 1.3.5. *Sean C un conjunto convexo y absorbente de un* e.v.t. $E[\mathfrak{T}]$ *y* p_C *su funcional de Minkowski asociado. Son equivalentes:*

(i) p_C *es continuo en* $E[\mathfrak{T}]$.

(ii) $0 \in \operatorname{int} C$ *(*i.e*, C es un entorno del origen).*

Además, en este caso se tiene que

$$\operatorname{int} C = \big\{x \in E : p_C(x) < 1\big\} \quad y \quad \overline{C} = \big\{x \in E : p_C(x) \leq 1\big\}. \tag{1.5}$$

Demostración. Como $0 \in \big\{x \in E : p_C(x) < 1\big\} \subset C$, si p_C es continuo, entonces $0 \in \operatorname{int} C$. Recíprocamente, si $0 \in \operatorname{int} C$, entonces C es un entorno del origen en $E[\mathfrak{T}]$ para el que se tiene la inclusión $\varepsilon C \subset \big\{x \in E : p_C(x) \leq \varepsilon\big\}$, lo que significa que p_C es continuo en el 0. El apartado *(ii)* de la proposición 1.3.1 permite obtener ahora que p_C es continuo en todos los puntos.

Supongamos ahora que p_C es continuo y demostremos la igualdad (1.5) correspondiente al interior. Es claro que $\big\{x \in E : p_C(x) < 1\big\} \subset \operatorname{int} C$. Para probar la inclusión contraria, tomemos $x \in \operatorname{int} C$ y fijemos $(t_n)_n$ una sucesión de números reales estrictamente mayores que 1 y convergente a 1. Entonces, $t_n x \to x$, y por lo tanto, existe $N \in \mathbb{N}$ tal que $t_N x \in \operatorname{int} C \subset C$. De aquí se sigue que $p_C(x) \leq \frac{1}{t_N} < 1$, que es lo que se quería demostrar. El resto de la prueba se deja como ejercicio. $\square$

Como consecuencia inmediata de la proposición anterior tenemos el siguiente resultado:

Corolario 1.3.6. *Sean* $E[\mathfrak{T}]$ *un* e.v.t. *y* $p : E \longrightarrow \mathbb{R}_+$ *una seminorma. Son equivalentes:*

(i) p *es continua.*

(ii) La bola $\big\{x \in E : p(x) < 1\big\}$ *es abierta.*

(iii) $0 \in \operatorname{int}\big\{x \in E : p(x) < 1\big\}$.

(iv) p *es continua en el* 0.

(v) Existe una seminorma continua $q : E[\mathfrak{T}] \longrightarrow \mathbb{R}_+$ *tal que* $p \leq q$.

La correspondencia existente entre entornos del origen absolutamente convexos y seminormas continuas permite demostrar que los *e.l.c.* son, exactamente, aquellos *e.v.t.* cuya topología está asociada a una familia de seminormas, y que en ellos, siempre hay formas lineales continuas que son no nulas.

Teorema 1.3.7. *Sea* $E[\mathfrak{T}]$ *un* e.v.t. *La topología* $\mathfrak{T}$ *está asociada a una familia de seminormas si, y sólo si,* $E[\mathfrak{T}]$ *es un espacio localmente convexo.*

Demostración. Si la topología $\mathfrak{T}$ está asociada a una familia de seminormas, entonces $\mathfrak{T}$ es localmente convexa. Recíprocamente, supongamos que $E[\mathfrak{T}]$ es un *e.l.c.* y sea $\mathscr{U}$ una base de entornos del origen para $\mathfrak{T}$ formada por conjuntos absolutamente convexos y cerrados. Después de la igualdad (1.5), para cada $U \in \mathscr{U}$ se tiene que $U = \big\{x \in E : p_U(x) \leq 1\big\}$, y así, $\mathfrak{T}$ es la topología asociada a la familia de seminormas $\{p_U : U \in \mathscr{U}\}$. $\square$

La continuidad de aplicaciones lineales entre *e.l.c.* se caracteriza de forma similar a como se caracteriza la continuidad de aplicaciones lineales entre espacios normados.

Proposición 1.3.8. *Sean $E[\mathfrak{T}]$ y $F[\tau]$ dos* e.l.c.*, y $T : E \longrightarrow F$ una aplicación lineal. Entonces, son equivalentes:*

(i) *T es continua.*

(ii) *T es continua en el 0.*

(iii) *Para cada seminorma continua q en $F[\tau]$, existe una seminorma continua p en $E[\mathfrak{T}]$ satisfaciendo $q\big(T(x)\big) \leq p(x)$, para cada $x \in E$.*

Demostración. Las implicaciones *(i)* $\Rightarrow$ *(ii)* y *(iii)* $\Rightarrow$ *(i)* son inmediatas. Veamos que *(ii)* $\Rightarrow$ *(iii)*. Sean q una seminorma continua en $F[\tau]$ y $U = \big\{x \in F : q(x) \leq 1\big\}$. U es un entorno del origen en F para el que podemos encontrar otro entorno del origen cerrado V en $E[\mathfrak{T}]$ tal que $T(V) \subset U$. Si p es el funcional de Minkowski asociado a V, entonces, para cada $x \in E$ y $\varepsilon > 0$, se tiene que

$$T\left(\frac{x}{p(x)+\varepsilon}\right) \in U,$$

y por lo tanto,

$$q\left(\frac{1}{p(x)+\varepsilon}T(x)\right) \leq 1,$$

lo que es equivalente a $q\big(T(x)\big) \leq p(x)+\varepsilon$, quedando así establecida la validez de *(iii)*. $\square$

$\boxed{\textit{Dual topológico}}$ Para un *e.v.t.* $E[\mathfrak{T}]$, llamamos *dual topológico* (o simplemente, dual, cuando no hay lugar a confusión) al conjunto de aplicaciones lineales de E en $\mathbb{K}$ (*i.e.*, formas lineales) que son continuas para $\mathfrak{T}$ y la topología natural de $\mathbb{K}$. El dual de $E[\mathfrak{T}]$ se denota por $(E[\mathfrak{T}])'$ o, simplemente, E', si la topología $\mathfrak{T}$ se da por supuesta.

STOP Queremos llamar la atención sobre lo siguiente: en el caso de un espacio normado $\big(X, \|\cdot\|\big)$, se utiliza la notación $\big(X, \|\cdot\|\big)^*$ o, simplemente, X^* para referirnos al dual topológico, que es la forma habitual en la que este dual es representado en los libros sobre espacios de Banach. Por esta razón, y a pesar de que la notación introducida en la definición anterior es distinta (es la habitual en los libros de *e.l.c.*), preferimos no utilizar la notación X' para el dual topológico de un espacio normado.

Espacios localmente convexos, 1920-1966. La teoría general de los espacios vectoriales topológicos se fundó en la década de 1920 a 1930, aunque sus orígenes son anteriores. D. Hilbert fue uno de los nombres más influyentes y activos que empujaron hacia la confluencia del análisis, el álgebra y la topología. En 1906, Hilbert, cuando investigaba sobre cuestiones desarrolladas por I. Fredholm, se dio cuenta de que la teoría de ecuaciones integrales podía ser presentada como un caso particular de la teoría de sistemas lineales con infinitas ecuaciones e incógnitas $(x_n)_n$: las únicas soluciones a considerar serían las que satisfacen $\sum_n x_n^2 < +\infty$. Desde ese momento estuvo claro que el *espacio de Hilbert* de las sucesiones de cuadrado convergente es esencial en toda la teoría, apareciendo de forma natural como *paso al límite* del espacio Euclídeo de dimensión infinita. Hilbert necesitó introducir en su espacio dos topologías, que correspondían a la topología fuerte y a lo que hoy en día llamamos topología débil; e incluso necesitó utilizar la compacidad débil de la bola unidad. En 1907, M. Fréchet, F. Riesz y E. Schmidt introdujeron el lenguaje de la geometría euclídea en el espacio de Hilbert, hablando de normas ($\|x\|$ con la notación actual), desigualdad triangular, etc. F. Riesz y E. Fisher demostraron, poco después, que el espacio de las funciones de cuadrado sumable (en el sentido que H. Lebesgue había definido en 1902) es isomorfo al espacio de Hilbert. La teoría de los espacios de Hilbert no fue presentada de forma axiomática hasta 1930, gracias a M. H. Stone y J. von Neumann. Antes, entre 1920 y 1922, S. Banach, H. Hahn y E. Helly dieron la definición de espacio normado general. Era sin embargo conocido, como Fréchet había notado en su tesis (en 1906), que existían nociones de convergencia clásicas que no correspondían a nociones de convergencia asociadas a una métrica: la topología de convergencia puntual en el espacio de las funciones reales acotadas. La definición general de espacio localmente convexo fue dada por von Neumann en 1935. Todos los conceptos anteriores son casos particulares de la noción de espacio vectorial topológico, que fue estudiada de forma sistemática hacia 1950 y, en particular, en los tratados de G. Köthe (1960) y N. Bourbaki (1966). Una estrella indiscutible del mundo de los espacios localmente convexos generales es el espacio de distribuciones de L. Schawrtz.

1.3.1 El teorema de Hahn-Banach: versión analítica

Extensiones de formas lineales Dados un *e.v.t.* E, un subespacio vectorial F de E y una aplicación lineal $f : F \longrightarrow \mathbb{R}$, siempre existe la posibilidad de obtener una prolongación lineal, $\tilde{f}$, de f a E. Basta, por ejemplo, considerar un complemento algebraico de F y definir $\tilde{f}$ como cero sobre él; o, de forma más general, tomar una base de Hamel en este complementario y extender f definiendo $\tilde{f}$ de manera arbitraria sobre los vectores de la base. De hecho, se obtienen así todas las posibles extensiones lineales de f a E. Si f es continua, no hay, a priori (incluso cuando E es un espacio normado de dimensión infinita), razón para suponer que alguna de tales extensiones también sea continua. En algunos espacios normados concretos es fácil construir, con procedimientos particulares, formas lineales continuas. La pregunta natural es, ¿existe un procedimiento general para construir funcionales lineales continuos en los espacios normados abstractos? El teorema de Hahn-Banach proporciona una respuesta afirmativa a dicha cuestión.

Teorema 1.3.9 (Hahn, 1927; Banach 1929)**.** *Sean E un espacio vectorial real y $p : E \longrightarrow \mathbb{R}$ un funcional subaditivo y positivamente homogéneo. Sean F un subespacio vectorial de E de codimensión 1 y $f : F \longrightarrow \mathbb{R}$ lineal de modo que $f(x) \leq p(x)$ para cada $x \in F$. Entonces, existe $\tilde{f} : E \longrightarrow \mathbb{R}$ lineal tal que $\tilde{f}$ restringida a F coincide con f y tal que $\tilde{f}(x) \leq p(x)$ para cada $x \in E$.*

Demostración. Por hipótesis, si $x_0 \in E \setminus F$, entonces $E = F \oplus \operatorname{span}\{x_0\}$. Así, para cada $x \in E$ se tiene que $x = y + ax_0$, con $y \in F$ y $a \in \mathbb{R}$, y por tanto, cualquier extensión lineal $\tilde{f}$ de f está dada por $\tilde{f}(x) = f(y) + a\alpha$, para cierto $\alpha \in \mathbb{R}$. Se trata ahora de determinar α para que se verifique la desigualdad

$$\tilde{f}(x) \leq p(x). \tag{1.6}$$

La existencia de un tal α exige condiciones que pasamos a analizar.

(a) Si $a > 0$, la desigualdad $\tilde{f}(ax_0 + y) = a\alpha + f(y) \leq p(ax_0 + y)$ es la misma que

$$\alpha + f(ya^{-1}) \leq p(x_0 + ya^{-1}),$$

lo que a su vez equivale a

$$\alpha \leq -f(z) + p(z + x_0), \qquad \text{para todo } z \in F.$$

(b) Si $a < 0$, la desigualdad $\tilde{f}(ax_0 + y) = a\alpha + f(y) \leq p(ax_0 + y)$, que equivale, después de dividir por $-a$, a que

$$-\alpha - f(ya^{-1}) \leq p(-x_0 - ya^{-1}),$$

es a su vez equivalente a

$$\alpha \geq f(w) - p(w - x_0), \qquad \text{para todo } w \in F.$$

Así pues, una condición necesaria para que exista $\tilde{f}$ cumpliendo (1.6), es que exista $\alpha \in \mathbb{R}$ satisfaciendo

$$f(w) - p(w - x_0) \leq \alpha \leq -f(z) + p(z + x_0), \qquad \text{para todo } z, w \in F. \tag{1.7}$$

Rehaciendo los cálculos anteriores en sentido inverso, es inmediato comprobar que dicha condición es también suficiente. Obsérvese ahora que, para que se cumpla la ecuación (1.7), basta con que se verifique

$$f(w) - p(w - x_0) \leq -f(z) + p(z + x_0)$$

para cada $z, w \in F$, lo cual es cierto debido a que

$$f(z) + f(w) = f(z + w) \leq p(z + w) = p(z + x_0 + w - x_0) \leq p(z + x_0) + p(w - x_0)$$

para cada $z, w \in F$. $\square$

Corolario 1.3.10 (Hahn, 1927; Banach 1929)**.** *Sean E un espacio vectorial real y $p : E \longrightarrow \mathbb{R}$ una seminorma. Sean F un subespacio vectorial de E de codimensión 1 y $f : F \longrightarrow \mathbb{R}$ lineal con $|f(x)| \leq p(x)$, para cada $x \in F$. Entonces, existe $\tilde{f} : E \longrightarrow \mathbb{R}$ lineal tal que $\tilde{f}$ restringida a F coincide con f y tal que $|\tilde{f}(x)| \leq p(x)$ para cada $x \in E$. En particular, si $\big(X, \|\cdot\|\big)$ es un espacio normado, $Y \subset X$ un subespacio y $f : Y \longrightarrow \mathbb{R}$ es lineal y continua, entonces existe $\tilde{f} : X \longrightarrow \mathbb{R}$ lineal y continua tal que $\tilde{f}$ restringida a Y coincide con f y $\|f\| = \|\tilde{f}\|$.*

Demostración. Para demostrar la primera parte del resultado basta utilizar el teorema 1.3.9, teniendo en cuenta que, dada una seminorma p, la acotación $f(x) \leq p(x)$, $x \in E$, equivale a la acotación $|f(x)| \leq p(x)$, para cada $x \in E$.

Supongamos ahora que X está dotado de una norma $\|\cdot\|$ y que $f : Y \longrightarrow \mathbb{R}$ es lineal y continua con norma $\|f\|$. Si definimos $p(x) := \|f\| \, \|x\|$, para $x \in X$, entonces p es una seminorma que cumple $|f(x)| \leq p(x)$, para cada $x \in Y$. Ahora, la segunda parte del corolario es consecuencia de lo demostrado en la primera parte del mismo. $\square$

Extensiones de Hahn-Banach en espacios normados separables Tal y como se pondrá de manifiesto en el teorema 1.3.12, la versión general del teorema de Hahn-Banach requiere del concurso del lema de Zorn. Sin embargo, el principio de Inducción en $\mathbb{N}$ y el corolario 1.3.10 son suficientes para probar el teorema de Hahn-Banach para espacios de Banach separables.

Teorema 1.3.11 (Hahn, 1927; Banach 1929)**.** *Sean $\big(X, \|\cdot\|\big)$ un espacio real normado separable, Y un subespacio vectorial de X, y $f : Y \longrightarrow \mathbb{R}$ una aplicación lineal y continua. Entonces, existe $\tilde{f} : X \longrightarrow \mathbb{R}$ lineal y continua, tal que $\tilde{f}$ restringida a Y coincide con f, satisfaciendo $\|f\| = \|\tilde{f}\|$.*

Demostración. Sea $\{x_n : n \in \mathbb{N}\}$ un conjunto numerable denso en X, y definamos

$$X_n := \operatorname{span}\big\{Y \cup \{x_1, x_2, \ldots, x_n\}\big\}, \qquad \text{para } n \in \mathbb{N}.$$

Entonces, o bien $X_n = X_{n+1}$, o bien X_n es un subespacio de X_{n+1} de codimensión 1. Por lo tanto, por inducción sobre n, y utilizando el corolario 1.3.10, podemos extender f a un funcional lineal g definido en el subespacio denso en X dado por $Z := \bigcup_{n=1}^{\infty} X_n$, el cual satisface $\|g\| = \|f\|$. Para cada $y \in X$, existe una sucesión $(y_n)_n$ en Z convergente a y. Como g es lineal y continua en Z, la sucesión $(g(y_n))_n$ es de Cauchy en $\mathbb{R}$, y así, podemos definir $\tilde{f}(y) := \lim_n g(y_n)$. Es inmediato comprobar que el valor $\tilde{f}(y)$ es independiente de la sucesión $(y_n)_n$ elegida en Z con tal de que ésta converja a y, que $\tilde{f}$ es lineal y que es continua, con $\|\tilde{f}\| = \|g\| (= \|f\|)$. $\square$

Lema de Kuratowski-Zorn y Teorema de Tychonoff Todas las demostraciones conocidas del teorema de Hahn-Banach, en su versión más general, se basan de alguna manera en el axioma de Elección. Siendo más específicos, el axioma de Elección se utiliza para probar el teorema clásico de Tychonoff sobre el producto de espacios compactos.

(AT) *El producto $\prod_{i \in I} K_i$ de una familia de espacios compactos $\{K_i\}_{i \in I}$ es compacto.*

Kelley demostró en [68] que el axioma de Elección y (AT) son equivalentes. Si por (AT2) denotamos el teorema de Tychonoff para espacios compactos separados, entonces se puede demostrar que (AT) no es equivalente a (AT2), que (AT2) implica el teorema de extensión de Hahn-Banach 1.3.12 y que este último no implica (AT), y por ende no implica el axioma de Elección (véase [97]). Para la demostración de 1.3.12, nosotros haremos uso del lema de Kuratowski-Zorn, a la postre equivalente al axioma de Elección, que ya ha sido utilizado para la justificación de la existencia de bases de Hamel en la página 20.

Teorema 1.3.12 (Hahn, 1927; Banach 1929). *Sean E un espacio vectorial real y $p : E \longrightarrow \mathbb{R}$ un funcional subaditivo y positivamente homogéneo. Sea F un subespacio vectorial de E y sea $f : F \longrightarrow \mathbb{R}$ lineal tal que $f(x) \leq p(x)$ para cada $x \in F$. Entonces, existe una aplicación lineal $\tilde{f} : E \longrightarrow \mathbb{R}$ tal que $\tilde{f}$ restringida a F coincide con f, verificando $\tilde{f}(x) \leq p(x)$ para cada $x \in E$.*

Demostración. Se obtiene aplicando el teorema 1.3.9 y el lema de Kuratowski-Zorn 1.2.1. Para aplicar dicho lema, se considera la colección $\mathscr{P}$ de todos los pares (Z, f_Z), donde Z es un subespacio vectorial de E con $F \subset Z$, y f_Z es una extensión lineal de f a Z verificando $f_Z(x) \leq p(x)$, para todo $x \in Z$. $\mathscr{P}$ se ordena mediante la relación

$$(Z, f_Z) \leq (W, f_W) \text{ si } Z \subset W \text{ y } f_W \text{ coincide con } f_Z \text{ sobre } Z.$$

Es inmediato que $(\mathscr{P}, \leq)$ es un conjunto no vacío, parcialmente ordenado, en el que cada cadena tiene supremo en $\mathscr{P}$. El lema de Kuratowski-Zorn garantiza que existen elementos maximales en $(\mathscr{P}, \leq)$. Si (Z, f_Z) es un elemento maximal, necesariamente $Z = E$, pues en caso contrario, tomando $x_0 \in E \setminus Z$, y de acuerdo con el teorema 1.3.9, f_Z se podría extender a $Z \oplus \operatorname{span}\{x_0\}$, lo cual contradice la maximalidad de Z. $\square$

El lema que sigue permite reducir el caso complejo al caso real.

Lema 1.3.13. *Sea E un espacio vectorial complejo.*

(i) *Si $f : E \longrightarrow \mathbb{C}$ es una forma $\mathbb{C}$-lineal, entonces su parte real, $\operatorname{Re} f$, es una forma $\mathbb{R}$-lineal, y $f(x) = \operatorname{Re} f(x) - i \operatorname{Re} f(ix)$.*

(ii) *Si $u : E \longrightarrow \mathbb{R}$ es una forma $\mathbb{R}$-lineal, entonces la forma $f(x) := u(x) - iu(ix)$ es $\mathbb{C}$-lineal, y $\operatorname{Re} f = u$.*

(iii) *Si $p : E \longrightarrow \mathbb{R}$ es una seminorma, entonces $|\operatorname{Re} f(x)| \leq p(x)$ para todo $x \in E$, si, y sólo si, $|f(x)| \leq p(x)$ para todo $x \in E$.*

Demostración. Los enunciados *(i)* y *(ii)* son de comprobación inmediata. Obsérvese que, para cada $x \in E$, existe un cierto escalar complejo λ, de módulo 1, tal que $|f(x)| = \lambda f(x)$, y por lo tanto se tiene que

$$|f(x)| = f(\lambda x) = \operatorname{Re} f(\lambda x) \leq p(\lambda x) = p(x),$$

lo que demuestra *(iii)*. $\square$

Teorema 1.3.14 (Hahn-Banach ($\mathbb{R}$), Sobczyk, 1939 ($\mathbb{C}$))**.** *Sean E un espacio vectorial, real o complejo, y $p : E \longrightarrow \mathbb{R}$ una seminorma. Sean F un subespacio vectorial de E y $f : F \longrightarrow \mathbb{K}$ una aplicación lineal tal que $|f(x)| \leq p(x)$, para todo $x \in F$. Entonces, existe una extensión lineal de f, $\tilde{f} : E \longrightarrow \mathbb{K}$, de modo que $|\tilde{f}(x)| \leq p(x)$, para todo $x \in E$.*

Demostración. Si el cuerpo es $\mathbb{R}$, como ya hemos hecho notar, la desigualdad $f(x) \leq p(x)$, para $x \in E$ (o $x \in F$), equivale a $|f(x)| \leq p(x)$, para $x \in E$ (o $x \in F$), y la conclusión se obtiene del teorema 1.3.12.

Cuando el cuerpo es $\mathbb{C}$, tomando $u := \operatorname{Re} f$ se obtiene una forma $\mathbb{R}$-lineal que verifica la desigualdad $|u(x)| \leq p(x)$. El caso anterior aplicado a u garantiza la existencia de una extensión $\mathbb{R}$-lineal $\tilde{u}$ de u a E conservando la acotación. Definiendo

$$\tilde{f}(x) := \tilde{u}(x) - i\tilde{u}(ix),$$

se obtiene del lema anterior que $|\tilde{f}(x)| \leq p(x)$, y que $\tilde{f}$ es una extensión $\mathbb{C}$-lineal de f. $\square$

Caracterización de e.v.t. *con dual no nulo* Con ayuda del teorema de Hahn-Banach 1.3.14, podemos caracterizar los *e.v.t.* cuyo dual es no nulo como se hace en el teorema 1.3.15. En particular, todo *e.l.c.* tiene dual no nulo (en el corolario 1.3.21 se probará que si E es un *e.l.c.*, entonces E' separa los puntos de E).

STOP Existen *e.v.t.* (como los espacios $L^p([0,1])$, para $0 < p < 1$) en los que la única forma lineal continua es la forma idénticamente cero, véase [71, p. 157-158].

Teorema 1.3.15. *Sea $E[\mathfrak{T}]$ un* e.v.t. *Las siguientes afirmaciones son equivalentes:*

(i) $(E[\mathfrak{T}])' \neq 0$.

(ii) Existe un entorno del origen en $E[\mathfrak{T}]$ que es convexo y distinto de E.

Demostración. Veamos cómo *(i)* $\Rightarrow$ *(ii)*. Si existe $f \in (E[\mathfrak{T}])'$ no idénticamente nula, entonces $U = \{x \in E : |f(x)| < 1\}$ es un abierto convexo que contiene al 0 y que no puede ser todo E. Efectivamente, si $U = E$, entonces, para cada $x \in E$, se tendría que $nx \in U$, para cada $n \in \mathbb{N}$, lo que conduce a que $|f(x)| < \frac{1}{n}$, para $n \in \mathbb{N}$; en consecuencia, $f = 0$, en contradicción con la hipótesis.

Demostremos que *(ii)* $\Rightarrow$ *(i)*. Supongamos que existe un entorno del origen convexo U distinto de E. Por la proposición 1.3.2, existe un entorno del origen equilibrado V tal que $V \subset U$. Para la envoltura convexa se tiene que $W := \operatorname{co}(V) \subset U$, y en consecuencia, W es distinto de E. W es un entorno del origen absolutamente convexo y, por tanto, su funcional de Minkowski p es una seminorma continua y distinta de cero en algún vector $x_0 \in E$. La forma lineal $g : \operatorname{span}\{x_0\} \longrightarrow \mathbb{K}$, dada por $g(\lambda x_0) = \lambda p(x_0)$, satisface $|g(y)| \leq p(y)$ para cada $y \in \operatorname{span}\{x_0\}$. El teorema de Hahn-Banach 1.3.14 nos garantiza la existencia de una forma lineal $f : E \longrightarrow \mathbb{K}$ que extiende a g, por tanto no nula, y que satisface $|f(x)| \leq p(x)$ para cada $x \in E$. La forma f es continua después de la proposición 1.3.8, y así acaba la demostración. $\square$

Versión analítica del teorema de Hahn-Banach, 1912-1923. El conocido teorema de Hahn-Banach es una herramienta fundamental en Análisis Funcional y, por tanto, un resultado de *obligada lectura* en este campo. El verdadero padre del teorema de Hahn-Banach es E. Helly, quien lo establece, en 1912, en el espacio $C([a,b])$, simplificando resultados anteriores de F. Riesz de 1911. A partir de los trabajos de Riesz y Helly, era natural generalizar sus resultados a espacios arbitrarios, lo que fue hecho por H. Hahn en 1922 y por S. Banach en 1923, restringiéndose, en primera instancia, a espacios normados completos. En la monografía de Banach [5] ya no aparece la restricción de completitud sobre los espacios involucrados. Toda la teoría de dualidad en espacios de Banach (o espacios localmente convexos) se basa en el teorema de Hahn-Banach. A partir del teorema de Hahn-Banach se puede deducir el teorema de Krein-Milman, calcular el dual de $C([a,b])$, probar la existencia de límites generalizados, obtener la caracterización de las mejores aproximaciones (uniformes) de funciones continuas a espacios de polinomios (Čebyšev), etc.

1.3.2 Ejemplos de *e.l.c.*

Topología de convergencia puntual Sean Z un conjunto y $E = \mathbb{K}^Z$ el conjunto de todas las aplicaciones de Z en $\mathbb{K}$. Para cada $z \in Z$, definimos $p_z(f) = |f(z)|$, con $f \in E$. La familia de seminormas $\{p_z : z \in Z\}$ satisface

$$\bigcap_{z\in Z}\big\{f \in E : p_z(f) = 0\big\} = \{0\},$$

y así, existe una única topología localmente convexa Hausdorff τ_p, para la cual, una base de entornos del punto $f \in E$ viene dada por la familia de conjuntos

$$\big\{V(f,z_1,z_2,\dots,z_n,\varepsilon) : z_i \in Z,\ 1 \le i \le n,\ n \in \mathbb{N},\ \varepsilon > 0\big\},$$

donde

$$V(f,z_1,z_2,\dots,z_n,\varepsilon) = \big\{g \in E : |f(z_i) - g(z_i)| < \varepsilon,\ i = 1,\dots,n\big\}.$$

τ_p es la topología en $E = \mathbb{K}^Z$ de la *convergencia puntual* sobre Z.

*Topologías débiles y débiles** Sean $E[\mathfrak{T}]$ un *e.l.c.* y E' su dual topológico. Entonces se define la *topología débil*, $\sigma(E,E')$, de E (respectivamente, *débil**, $\sigma(E',E)$, de E'), como la topología asociada a la familia de seminormas $\{p_{x'} : x' \in E'\}$ (respectivamente, $\{p_x : x \in E\}$) dadas por $p_{x'}(x) = |x'(x)|$ (respectivamente, $p_x(x') = |x'(x)|$).

Puesto que $\sigma(E',E)$ puede considerarse como la topología inducida por $(\mathbb{K}^E, \tau_p)$ en E', la topología débil* es una topología localmente convexa Hausdorff. Una base de entornos del origen para $\sigma(E',E)$ viene dada por

$$\big\{V(0,x_1,\dots,x_n,\varepsilon) : x_1,\dots,x_n \in E,\ n \in \mathbb{N},\ \varepsilon > 0\big\},$$

donde
$$V(0,x_1,\dots,x_n,\varepsilon)=\{x'\in E':|x'(x_i)|<\varepsilon,\, i=1,\dots,n\}.$$
Dado que la topología débil de E está generada por la familia de seminormas $\{p_{x'}:x'\in E'\}$, donde $p_{x'}(x)=|x'(x)|$ para $x\in E$, y además se satisface que
$$\bigcap_{x'\in E'}\{x:p_{x'}(x)=0\}=\{0\},$$
véase el apartado *(ii)* del corolario 1.3.21, se tiene que $\sigma(E,E')$ es una topología localmente convexa Hausdorff, cuya base de entornos del origen viene dada por
$$\{V(0,x'_1,\dots,x'_n,\varepsilon):x'_i\in E',\, 1\le i\le n,\, n\in\mathbb{N},\, \varepsilon>0\},$$
donde
$$V(0,x'_1,\dots,x'_n,\varepsilon)=\{x\in E:|x'_i(x)|<\varepsilon,\, i=1,\dots,n\}.$$
Cuando $(X,\|\cdot\|)$ es un espacio normado, es un sencillo ejercicio comprobar que la topología débil $\sigma(X,X^*)$ (respectivamente, débil* $\sigma(X^*,X)$) es más gruesa que la topología asociada a la norma (respectivamente, asociada a la norma dual).

Topologías en espacios de funciones continuas Sea S un espacio topológico completamente regular y denotemos por $C(S)$ (respectivamente, $C_b(S)$) el espacio de las funciones escalares y continuas (respectivamente, funciones escalares, continuas y acotadas) definidas en S. $C(S)\subset\mathbb{K}^S$, luego en $C(S)$ se puede considerar la topología inducida por τ_p, que será una topología localmente convexa Hausdorff en $C(S)$, para la cual, una base de entornos de cada punto f viene dada por
$$\{V(f,x_1,\dots,x_n,\varepsilon):x_i\in S,\, 1\le i\le n,\, n\in\mathbb{N},\, \varepsilon>0\},$$
donde
$$V(f,x_1,\dots,x_n,\varepsilon)=\{g\in C(S):|f(x_i)-g(x_i)|<\varepsilon,\, i=1,\dots,n\}.$$
Si denotamos por $\mathscr{K}$ la familia de los compactos de S, y para cada $K\in\mathscr{K}$ definimos $p_K(f):=\sup_{x\in K}|f(x)|$, entonces $\{p_K:K\in\mathscr{K}\}$ es una familia de seminormas en $C(S)$ satisfaciendo la condición
$$\bigcap_{K\in\mathscr{K}}\{f:p_K(f)=0\}=\{0\}.$$
Por tanto, existe una topología localmente convexa Hausdorff, τ_K, en $C(S)$ para la cual la base de entornos de cada punto f viene dada por
$$\{V(f,K,\varepsilon):K\in\mathscr{K},\, \varepsilon>0\},$$
donde
$$V(f,K,\varepsilon)=\{g\in C(S):|f(x)-g(x)|<\varepsilon,\ \text{para cada } x\in K\}.$$

Cuando $S = \Omega \subset \mathbb{K}^k$ es un abierto, podemos tomar $K_n \subset \Omega$, $n \in \mathbb{N}$, una sucesión de compactos verificando:

(i) $\Omega = \bigcup_{n=1}^{\infty} K_n$;
(ii) $K_n \subset \operatorname{int} K_{n+1}$, $n \in \mathbb{N}$.

Se deduce, de las propiedades *(i)* y *(ii)* anteriores, que para cada compacto $K \subset \Omega$, existe $m \in \mathbb{N}$ tal que $K \subset K_m$. Una sucesión del tipo $(K_n)_n$ se denomina sucesión exhaustiva de compactos de Ω. Así, τ_K en $C(\Omega)$ es la topología asociada a la sucesión de seminormas $\{p_{K_n} : n \in \mathbb{N}\}$. Razonamientos estándar permiten demostrar que τ_K es la topología asociada a la métrica

$$d(f,g) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{p_{K_n}(f-g)}{1+p_{K_n}(f-g)},$$

véase el teorema 1.3.16. El espacio $(C(\Omega), \tau_K)$ es un *e.l.c.* metrizable y completo.

Topologías en espacios de funciones holomorfas Si $\Omega \subset \mathbb{C}$ es un abierto, podemos considerar el espacio de funciones holomorfas $\mathscr{H}(\Omega)$ en el abierto Ω como subespacio de $(C(\Omega), \tau_K)$ con la topología inducida. $(\mathscr{H}(\Omega), \tau_K)$ es un *e.l.c.* metrizable y completo, debido a que el teorema de Weierstrass, [35, Theorem 2.1], garantiza que el límite de una sucesión de funciones holomorfas uniformemente convergente sobre compactos es, a su vez, una función holomorfa.

Topologías en espacios de funciones diferenciables Sea $\Omega \subset \mathbb{R}^k$ abierto. Consideramos los siguientes espacios de funciones:

$\mathscr{E}^m(\Omega) = \big\{ f : \Omega \longrightarrow \mathbb{R} : m\text{-veces diferenciables con continuidad en } \Omega \big\}$,

$\mathscr{E}(\Omega) = \big\{ f : \Omega \longrightarrow \mathbb{R} : \text{ infinitamente diferenciables con continuidad en } \Omega \big\}$.

Para $f \in \mathbb{K}^\Omega$, definimos

$$\operatorname{sop}(f) := \overline{\{x \in \Omega : f(x) \neq 0\}}.$$

Sean,

$\mathscr{D}_K^m(\Omega) = \big\{ f \in \mathscr{E}^m(\Omega) : \operatorname{sop}(f) \subset K \big\}$, para $K \subset \Omega$, compacto.

$\mathscr{D}_K(\Omega) = \big\{ f \in \mathscr{E}(\Omega) : \operatorname{sop}(f) \subset K \big\}$, para $K \subset \Omega$, compacto.

$\mathscr{D}(\Omega) = \bigcup_{K \in \mathscr{K}} \mathscr{D}_K(\Omega)$, donde $\mathscr{K}$ es la familia de los subconjuntos compactos de Ω.

En el espacio $\mathscr{E}^m(\Omega)$ se considera la topología τ_K^m de convergencia uniforme sobre compactos de las funciones y sus derivadas hasta el grado m. Una familia de seminormas definiendo dicha topología es $\{p_K : K \in \mathscr{K}\}$, donde

$$p_K(f) := \sup_{|\alpha| \leq m} \sup_{x \in K} |D^\alpha f(x)|,$$

siendo $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k) \in (\mathbb{N} \cup \{0\})^k$, $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_k$ y

$$D^\alpha f(x) := \frac{\partial^{|\alpha|} f(x)}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_k^{\alpha_k}}.$$

En el espacio $\mathscr{E}(\Omega)$ se considera la topología de la convergencia uniforme sobre compactos de las funciones y todas sus derivadas. La familia de seminormas $\{p_{K,m} : K \in \mathscr{K}, m \in \mathbb{N}\}$, donde

$$p_{K,m}(f) := \sup_{|\alpha| \le m} \sup_{x \in K} |D^\alpha f(x)|,$$

define la topología de $\mathscr{E}(\Omega)$. Utilizando que Ω se puede poner como unión de una sucesión exhaustiva de compactos, es fácil probar que los *e.l.c.* $\mathscr{E}^m(\Omega)$ y $\mathscr{E}(\Omega)$ son metrizables. En $\mathscr{D}_K^m(\Omega)$ y $\mathscr{D}_K(\Omega)$ se consideran las topologías inducidas, respectivamente, por $\mathscr{E}^m(\Omega)$ y $\mathscr{E}(\Omega)$. El espacio $\mathscr{D}(\Omega)$ es un espacio no nulo al que se da el nombre de espacio *base de distribuciones*, cuando se le dota de la topología localmente convexa más fina que hace continuas las inmersiones $\mathscr{D}_K(\Omega) \hookrightarrow \mathscr{D}(\Omega)$, véase [36, Definition 5.19].

1.3.3 Espacios localmente convexos metrizables y normables

Metrizabilidad y Primer Axioma de Numerabilidad Por definición, un espacio topológico satisface el *Primer Axioma de Numerabilidad* si todos los puntos tienen una base de entornos numerable. Los espacios métricos satisfacen el Primer Axioma de Numerabilidad, aunque no todos los espacios topológicos que satisfacen el Primer Axioma de Numerabilidad son metrizables (el intervalo de ordinales $[0, \omega_1)$ es numerablemente compacto, no compacto, satisface el Primer Axioma de Numerabilidad y no es metrizable, [99, p. 151]). Sin embargo, para *e.v.t.*, metrizabilidad y Primer Axioma de Numerabilidad son equivalentes, véase [71, §15.11.1]: esto es así porque todo *e.v.t.* es un espacio uniforme, y cuando satisface el Primer Axioma de Numerabilidad, entonces su uniformidad tiene una base numerable; los espacios uniformes con una base de la uniformidad numerable son metrizables, [69, Teorema 13, Capítulo 6].

La equivalencia entre metrizabilidad y Primer Axioma de Numerabilidad para *e.l.c.* se puede probar como sigue:

Teorema 1.3.16. *Sea $E[\mathfrak{T}]$ un* e.l.c. *Entonces, las siguientes afirmaciones son equivalentes:*

(i) *$E[\mathfrak{T}]$ es metrizable.*
(ii) *Existe una base de entornos del origen numerable,* i.e.*, $E[\mathfrak{T}]$ satisface el Primer Axioma de Numerabilidad.*
(iii) *Existe una familia numerable de seminormas continuas generando la topología $\mathfrak{T}$.*

Demostración. La afirmación *(i)* claramente implica *(ii)*. Supongamos que 0 tiene una base de entornos del origen numerable $\{V_n\}_n$. Podemos suponer que cada V_n es absolutamente convexo y que se satisface $V_n \supset V_{n+1}$, $n \in \mathbb{N}$. Sea p_n el funcional de Minkowski asociado a V_n. Por la proposición 1.3.4, cada p_n es una seminorma. Se tiene además que $p_n \le p_{n+1}$ para cada $n \in \mathbb{N}$, y es claro que $\mathfrak{T}$ está asociada a la familia $\{p_n : n \in \mathbb{N}\}$. Así acaba la prueba de *(ii)* $\Rightarrow$ *(iii)*.

Demostremos, para terminar, que *(iii)* $\Rightarrow$ *(i)*. Supongamos que $\mathfrak{T}$ está asociada a una familia de seminormas $\{p_n : n \in \mathbb{N}\}$, para la que podemos suponer, sin pérdida de generalidad, que se cumple $p_n \leq p_{n+1}$, $n \in \mathbb{N}$ (bastaría cambiar la familia $\{p_n : n \in \mathbb{N}\}$ por la familia $\{q_n : n \in \mathbb{N}\}$, donde $q_n(x) := \sup_{1\leq k \leq n} p_k(x)$, para $x \in E$). La fórmula

$$d(x,y) := \sum_{n=1}^{\infty} \frac{1}{2^n} \operatorname{mín}\{p_n(x-y), 1\}$$

define una distancia en E cuya topología asociada es $\mathfrak{T}$. Dejamos como ejercicio el comprobar que d es una distancia. Para ver que $\mathfrak{T}$ es la topología asociada a d, es suficiente demostrar que la familia $\big\{\{y \in E : p_n(y-x) < \varepsilon\} : n \in \mathbb{N}, 0 < \varepsilon < 1\big\}$ forma una base de d-entornos del punto x, para cada $x \in E$. Efectivamente, dados $x \in E$ y $0 < \varepsilon < 1$, consideremos $B_d(x, \varepsilon)$. Si tomamos m tal que $\sum_{n=m+1}^{\infty} \frac{1}{2^n} < \frac{\varepsilon}{2}$ y

$$V := \Big\{y \in E : p_m(x-y) < \frac{\varepsilon}{2m}\Big\},$$

entonces $V \subset B_d(x, \varepsilon)$. Recíprocamente, si $0 < \varepsilon < 1$ y $V = \big\{y \in E : p_n(x-y) < \varepsilon\big\}$, entonces $B_d\big(x, \frac{\varepsilon}{2^n}\big) \subset V$, y así acaba la prueba. $\square$

$\boxed{\textit{Espacio de Fréchet}}$ Un *espacio de Fréchet* es un *e.l.c.* metrizable y completo. Los espacios de funciones $(C(\Omega), \tau_K)$, $(\mathscr{E}^m(\Omega), \tau_K^m)$ y $(\mathscr{H}(\Omega), \tau_K)$, analizados en la sección 1.3.2, son espacios de Fréchet.

Maurice René Fréchet, 1878-1974, Francia. Fréchet estudió en su tesis ejemplos concretos de espacios localmente convexos, metrizables y completos (el espacio de todas las sucesiones de $\mathbb{R}^n$ y el espacio de las funciones holomorfas en el disco unidad). Fréchet hizo énfasis en la posibilidad de definir la topología de estos espacios por una distancia completa. S. Banach, primero, abstrajo las ideas de Fréchet y consideró lo que se llaman los *espacios de tipo (F)*; S. Mazur y W. Orlicz, después, sistematizaron las ideas previas en lo que llamaron *espacios de tipo (B_0)*, que corresponden, exactamente, a lo que hoy se conoce con el nombre de espacios de Fréchet, véanse [41, p. 215-216] y las referencias allí dadas.

$\boxed{\textit{Conjunto acotado}}$ Sea $E[\mathfrak{T}]$ un *e.l.c.* Un conjunto $A \subset E$ se dice *acotado para* $\mathfrak{T}$ si, para cada entorno del origen U, existe $\rho > 0$ tal que $A \subset \rho U$. El conjunto $A \subset E$ es acotado para $\mathfrak{T}$ si, y sólo si, para cada seminorma $\mathfrak{T}$-continua $p : E \longrightarrow \mathbb{R}_+$ se tiene $\sup\{p(x) : x \in A\} < \infty$. El concepto de acotado en un *e.l.c.* extiende el concepto de acotado en espacios normados.

Teorema 1.3.17 (Kolmogoroff, 1934). *Si $E[\mathfrak{T}]$ es un* e.l.c.*, entonces las siguientes afirmaciones son equivalentes:*

(i) $E[\mathfrak{T}]$ *es normable,* i.e.*, $\mathfrak{T}$ es la topología asociada a una norma.*

(ii) $E[\mathfrak{T}]$ *tiene un entorno del origen acotado.*

Demostración. La implicación *(i)* $\Rightarrow$ *(ii)* es clara, dado que la bola unidad para una norma es siempre un entorno del origen acotado. Recíprocamente, demostremos *(ii)* $\Rightarrow$ *(i)*. Sea U un entorno abierto del origen absolutamente convexo y acotado, y sean p_U su funcional de Minkowski y $\mathfrak{T}_{p_U}$ la topología asociada a p_U en E. Se tiene claramente que $\mathfrak{T} \geq \mathfrak{T}_{p_U}$, dado que U es entorno del origen. Si vemos que $\mathfrak{T} \leq \mathfrak{T}_{p_U}$, entonces $\mathfrak{T} = \mathfrak{T}_{p_U}$, y la prueba estará terminada. Gracias a la proposición 1.3.5 se tiene la igualdad $U = \big\{x \in E : p_U(x) < 1\big\}$ y, en consecuencia, tenemos que $\big\{\varepsilon U = \{x \in E : p_U(x) < \varepsilon\} : \varepsilon > 0\big\}$ es una $\mathfrak{T}_{p_U}$-base de entornos del origen en E. Para probar que $\mathfrak{T} \leq \mathfrak{T}_{p_U}$ es suficiente demostrar que, dado un $\mathfrak{T}$-entorno del origen V, existe $\varepsilon > 0$ tal que $\varepsilon U \subset V$, lo que se obtiene de la definición de acotación para U. $\square$

Proposición 1.3.18. *Sea $\Omega \subset \mathbb{C}$ abierto. Entonces, $(\mathscr{H}(\Omega), \tau_K)$ es un espacio de Fréchet que no es un espacio de Banach.*

Demostración. Si $(\mathscr{H}(\Omega), \tau_K)$ fuese normable, existiría un τ_K-entorno del origen acotado, $\mathscr{V}$, en $\mathscr{H}(\Omega)$. Así, para cada compacto $K \subset \Omega$, se tendría

$$\sup_{f \in \mathscr{V}} \Big\{\sup_{z \in K} |f(z)|\Big\} < \infty.$$

De esta forma, $\mathscr{V}$ es una familia normal de funciones holomorfas, véase [96, Theorem 14.6], y por lo tanto, $\overline{\mathscr{V}}^{\tau_K}$ es un entorno del origen compacto en el espacio normable $(\mathscr{H}(\Omega), \tau_K)$. Un bien conocido resultado de Riesz establece que los espacios normados de dimensión finita son aquéllos que tienen un entorno del origen compacto, [23, Teorema 1.2.8], y por tanto, $\mathscr{H}(\Omega)$ es de dimensión finita, lo cual es absurdo; en consecuencia, $(\mathscr{H}(\Omega), \tau_K)$ no es normable. $\square$

STOP Como existen *e.l.c.* metrizables y completos que no son normables y *e.v.t.* que no son *e.l.c.* (por ejemplo, $L^p([0,1])$ para $0 < p < 1$), se tienen las siguientes inclusiones estrictas:
$\{$espacios Hilbert$\} \subsetneq \{$espacios Banach$\} \subsetneq \{$espacios Fréchet$\} \subsetneq \{$*e.v.t.* metrizables completos$\}$

Espacios de Fréchet-Montel Un *e.l.c.* en el que los subconjuntos acotados son relativamente compactos se dice que es un espacio de *Montel*. A los espacios de Fréchet que son de Montel se les llama espacios de *Fréchet-Montel*. Puede probarse que todo espacio de Fréchet-Montel es un espacio separable, véase [71, §27.2.(5)]. En particular, para cada abierto $\Omega \subset \mathbb{C}$, el espacio $(\mathscr{H}(\Omega), \tau_K)$ es un espacio separable.

1.3.4 Teoremas de separación de conjuntos convexos

Estableceremos aquí los teoremas de separación de conjuntos convexos, aunque empezaremos obteniendo algunas consecuencias analíticas del teorema de Hahn-Banach 1.3.14.

Teorema 1.3.19. *Sean $E[\mathfrak{T}]$ un* e.l.c. *y $F \subset E$ un subespacio. Si $u : F \longrightarrow \mathbb{K}$ es lineal y continua, entonces existe $f : E \longrightarrow \mathbb{K}$, lineal y continua, tal que $f|_F = u$.*

Demostración. Si u es continua en F, $V := \{x \in F : |u(x)| \leq 1\}$ es un entorno del origen en F. Existe un entorno del origen absolutamente convexo U en E tal que $U \cap F \subset V$. El funcional de Minkowski p de U es una seminorma en E que satisface $|u(x)| \leq p(x)$, para cada $x \in F$, dado que $U \cap F \subset V$. Por el teorema 1.3.14, existe una extensión lineal de u, $f : E \longrightarrow \mathbb{K}$, tal que

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

para cada $x \in E$. Como p es una seminorma $\mathfrak{T}$-continua, proposición 1.3.5, se tiene que f es continua después de la proposición 1.3.8. $\square$

El lema que sigue será utilizado en el corolario demostrado posteriormente.

Lema 1.3.20. *Sean $E[\mathfrak{T}]$ un* e.v.t. *y $F \subset E$ un subespacio de dimensión finita n. Entonces, todo isomorfismo algebraico $f : \mathbb{K}^n \longrightarrow F$ es un isomorfismo topológico, y F es cerrado en $E[\mathfrak{T}]$.*

Demostración. Haremos la demostración por inducción. Para $n = 1$, el resultado es cierto. Supongamos ahora que es cierto para $n-1$, y demostrémoslo para n. Si $f : \mathbb{K}^n \longrightarrow F$ es un isomorfismo algebraico y $\{e_i : i = 1, \ldots, n\}$ es la base canónica de $\mathbb{K}^n$, entonces

$$f\big((\alpha_1, \alpha_2, \ldots, \alpha_n)\big) = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n,$$

donde $u_i := f(e_i)$, $i = 1, \ldots, n$. De aquí se sigue que f es continua, dado que las operaciones de espacio vectorial son continuas para $\mathfrak{T}$. Por otro lado, dado $x \in E$, existen unos únicos $f_i(x) \in \mathbb{K}$, $i = 1, \ldots, n$, tales que

$$x = f_1(x)u_1 + f_2(x)u_2 + \ldots + f_n(x)u_n.$$

Para cada $i = 1, \ldots, n$, la forma lineal $f_i : F \longrightarrow \mathbb{K}$ es no nula. $\ker f_i$ es un subespacio de dimensión $n-1$ de F que, por hipótesis de inducción, es cerrado. Esto implica que cada f_i es continua por el apartado *(vii)* de la proposición 1.3.1, y así, $f^{-1} = (f_1, f_2, \ldots, f_n)$ es continua, con lo que la demostración queda terminada. $\square$

Corolario 1.3.21. *Sea $E[\mathfrak{T}]$ un* e.l.c.

(i) Si $F \subset E$ es un subespacio, entonces la aplicación restricción

$$(E[\mathfrak{T}])' \longrightarrow (F[\mathfrak{T}|_F])'$$

que a cada $x' \in (E[\mathfrak{T}])'$ le hace corresponder su restricción $x'|_F$, es sobreyectiva.

(ii) Si $x_0 \in E$, $x_0 \neq 0$, entonces existe $f \in E'$ tal que $f(x_0) \neq 0$.

(iii) Si $\{x_1, \ldots, x_n\}$ son vectores linealmente independientes en E, entonces existen $f_1, \ldots, f_n \in E'$ tales que

$$f_i(x_j) = \delta_{ij}, \quad i, j = 1, \ldots, n.$$

Demostración. La afirmación *(i)* es una consecuencia inmediata del teorema 1.3.19. Para probar *(ii)* observemos que si $x_0 \neq 0$, podemos tomar una seminorma $p : E \longrightarrow \mathbb{R}_+$, continua para $\mathfrak{T}$, tal que $p(x_0) \neq 0$. La forma lineal $u : \operatorname{span}\{x_0\} \longrightarrow \mathbb{K}$ dada por $u(\lambda x_0) = \lambda p(x_0)$ satisface $|u(y)| \leq p(y)$, para cada $y \in \operatorname{span}\{x_0\}$. Por el teorema 1.3.14, existe $f : E \longrightarrow \mathbb{K}$ lineal, extendiendo u, y tal que $|f(x)| \leq p(x)$, para cada $x \in E$; f es $\mathfrak{T}$-continua y $f(x_0) = p(x_0) \neq 0$. Para demostrar *(iii)* utilizaremos el lema anterior. Si $\{x_1, x_2, \ldots, x_n\}$ es un conjunto linealmente independiente en E, $F := \operatorname{span}\{x_1, x_2, \ldots, x_n\}$ y $f_i : F \longrightarrow \mathbb{K}$, $i = 1, \ldots, n$, son las únicas aplicaciones lineales satisfaciendo $f_i(x_j) = \delta_{ij}$, $i, j = 1, \ldots, n$, entonces cada f_i es continua en F para la topología inducida por $E[\mathfrak{T}]$ (dado que el lema anterior nos dice que, en F, existe una única topología localmente convexa que se obtiene por isomorfismo con $\mathbb{K}^n$). La prueba acaba considerando extensiones continuas de f_i a $E[\mathfrak{T}]$, garantizadas por el teorema 1.3.19. □

Variedades afines e hiperplanos Se llama *variedad afín* en el espacio vectorial E a cualquier conjunto de la forma $x_0 + F$, donde F es un subespacio vectorial de E y $x_0 \in E$. Un subespacio propio maximal de E es un *hiperplano*, y toda variedad afín correspondiente a un hiperplano se denomina *hiperplano afín*.

Si $E[\mathfrak{T}]$ es un *e.v.t.*, entonces $M \subset E$ es un hiperplano afín si, y sólo si, existen $f : E \longrightarrow \mathbb{K}$ lineal y $a \in \mathbb{K}$ tales que $M = \{x \in E : f(x) = a\}$. Las propiedades *(i)* y *(vii)* de la proposición 1.3.1 aseguran que M es cerrado si, y sólo, si $f \in E'$.

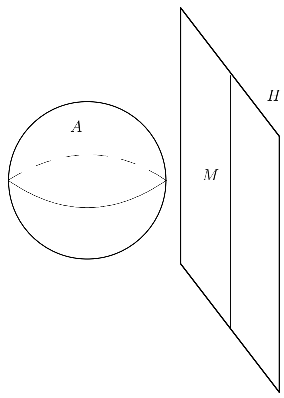


Figura 1.3: Teorema de Mazur

Teorema 1.3.22 (Mazur)**.** *Sean $E[\mathfrak{T}]$ un* e.v.t.*, $M \subset E$ una variedad afín y $A \subset E$ un conjunto abierto y convexo no vacío. Si $A \cap M = \emptyset$, entonces existe un hiperplano afín cerrado H en $E[\mathfrak{T}]$ tal que $A \cap H = \emptyset$ y $M \subset H$.*

Demostración. No es restrictivo suponer que $0 \in A$ (bastaría efectuar una traslación si fuese necesario); suponemos así que A es un entorno abierto y convexo del origen. La demostración se hace distinguiendo los casos $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$.

Caso $\mathbb{K}=\mathbb{R}$. $M=x_0+F$ para cierto $x_0\in E$ y para cierto subespacio vectorial real $F\subset E$. Si p es el funcional de Minkowski asociado a A, entonces $A=\{x\in E: p(x)<1\}$ por la proposición 1.3.5. Como $A\cap M=\emptyset$, se tiene que, para cada $y\in F$, $p(x_0+y)\geq 1$. Si definimos $u: F\oplus \operatorname{span}\{x_0\}\longrightarrow \mathbb{R}$ mediante $u(y+\lambda x_0):=\lambda$, entonces u es lineal y se satisface que $u(y+\lambda x_0)\leq p(y+\lambda x_0)$ para todo $\lambda\in\mathbb{R}$ y todo $y\in F$. Efectivamente:

(i) Si $\lambda\leq 0$, entonces $u(y+\lambda x_0)=\lambda\leq 0\leq p(y+\lambda x_0)$.
(ii) Si $\lambda>0$, entonces $u(y+\lambda x_0)=\lambda\cdot 1\leq \lambda p\big(\frac{y}{\lambda}+x_0\big)\leq p(y+\lambda x_0)$.

Por el teorema 1.3.12, existe $f:E\longrightarrow\mathbb{R}$ lineal tal que $f|_{F\oplus\operatorname{span}\{x_0\}}=u$, y verificando que $f(x)\leq p(x)$, para cada $x\in E$. Si $U\subset A$ es un entorno del origen absolutamente convexo y q es su funcional de Minkowski, entonces q es una seminorma $\mathfrak{T}$-continua, proposición 1.3.5, y se tiene $p\leq q$. En particular, $|f(x)|\leq q(x)$ para cada $x\in E$ y, en consecuencia, f es $\mathfrak{T}$-continua. Si consideramos el hiperplano cerrado $H:=\{x\in E: f(x)=1\}$, es claro que $M\subset H$ y que $H\cap A=\emptyset$, dado que si $x\in H$, se tiene $f(x)=1\leq p(x)$, lo que significa que $x\notin A$, quedando terminada la prueba.

Caso $\mathbb{K}=\mathbb{C}$. $M=x_0+F$ para cierto $x_0\in E$ y cierto subespacio vectorial complejo $F\subset E$. Por el apartado anterior, existe $f:E\longrightarrow\mathbb{R}$ lineal y continua con $H:=\{x\in E: f(x)=1\}\supset M$, siendo $f(x_0)=1$ y $H\cap A=\emptyset$. Es claro que, para $y\in F$, se tiene $f(y)=0$. Como F es un subespacio vectorial complejo, se verifica que $iy\in F$ siempre que $y\in F$ y, en consecuencia, $f(iy)=0$ para cada $y\in F$. Si definimos $g(x):=f(x)-if(ix)$, entonces el hiperplano complejo $H_0:=\big\{x\in E: g(x)=1-if(ix_0)\big\}$ es cerrado y satisface las propiedades requeridas.

La demostración está ahora completa. □

$\boxed{\textit{Semiespacios determinados por hiperplanos}}$ Sean $E[\mathfrak{T}]$ un *e.v.t.*, $f\in E'$ y $\alpha\in\mathbb{R}$. Los semiespacios abiertos (respectivamente, cerrados) en $E[\mathfrak{T}]$ determinados por f y α son los conjuntos

$$G_\alpha=\big\{x\in E: f(x)<\alpha\big\},\qquad G^\alpha=\big\{x\in E: f(x)>\alpha\big\}$$

$$\text{(respectivamente, } F_\alpha=\big\{x\in E: f(x)\leq\alpha\big\},\qquad F^\alpha=\big\{x\in E; f(x)\geq\alpha\big\}\text{)}.$$

$\boxed{\textit{Separación de conjuntos}}$ Si A y B son dos subconjuntos de $E[\mathfrak{T}]$, se dice que:

(i) A y B están *estrictamente separados* si existen $f\in E'$ y $\alpha\in\mathbb{R}$ de tal manera que $A\subset G_\alpha$ y $B\subset G^\alpha$; en este caso, se dice que el hiperplano $\big\{x\in E: f(x)=\alpha\big\}$ separa estrictamente A y B.
(ii) A y B *están separados* si existen f y α como en el apartado anterior con $A\subset F_\alpha$ y $B\subset F^\alpha$; en este caso, se dice que el hiperplano $\big\{x\in E: f(x)=\alpha\big\}$ separa A y B.

Lema 1.3.23. *Toda forma lineal no nula f definida en un* e.v.t. *es abierta.*

Demostración. Sean $E[\mathfrak{T}]$ un *e.v.t.* y $f:E\longrightarrow\mathbb{K}$ lineal tal que $f(x_0)\neq 0$, para algún $x_0\in E$. Probaremos que si G es un abierto de $E[\mathfrak{T}]$, entonces $f(G)$ es abierto en $\mathbb{K}$, *i.e.*, demostraremos que, para cada $y\in G$, el conjunto $f(G)$ es un entorno de $f(y)$. Distinguimos dos casos:

Cuando $f(y) \neq 0$. Como G es abierto y la operación multiplicación por escalares es continua, existe $\delta > 0$ tal que $ay \in G$ cuando $|a-1| < \delta$. Si $\lambda \in \mathbb{K}$ es tal que $|\lambda - f(y)| < \delta|f(y)|$, entonces $|\lambda/f(y) - 1| < \delta$, y por tanto $\lambda = af(y) = f(ay) \in f(G)$.

Cuando $f(y) = 0$. Como G es abierto, existe un entorno V de 0 tal que $y + V \subset G$. Además, dado que $0_{\mathbb{K}} \cdot x_0 = 0_E$ y la operación multiplicación por escalares es continua, existe $\rho > 0$ tal que $\lambda x_0 \in V$ si $|\lambda| < \rho$. Tenemos entonces que

$$\{\lambda : |\lambda| < \rho|f(x_0)|\} = \{f(y + \lambda x_0) : |\lambda| < \rho\} \subset f(y+V) \subset f(G),$$

y la demostración está terminada. □

Corolario 1.3.24 (Primer teorema de separación)**.** *Sean $E[\mathfrak{T}]$ un* e.v.t. *y A, B subconjuntos convexos no vacíos de E, con $A \cap B = \emptyset$. Si A es abierto, existen $f : E \longrightarrow \mathbb{R}$ lineal y $\mathfrak{T}$-continua, y $\alpha \in \mathbb{R}$, tales que*

$$f(a) < \alpha \leq f(b), \quad \textit{para cada } a \in A,\ b \in B.$$

Si ambos, A y B, son abiertos, f puede tomarse satisfaciendo

$$f(a) < \alpha < f(b), \quad \textit{para cada } a \in A,\ b \in B;$$

es decir, A y B están separados estrictamente por un hiperplano real cerrado.

Demostración. El conjunto $A - B$ es un abierto no vacío y $0 \notin A - B$. El teorema 1.3.22 nos asegura la existencia de $f : E \longrightarrow \mathbb{R}$, lineal y continua, y $\beta \in \mathbb{R}$, tales que $0 \in H := \{x \in E : f(x) = \beta\}$ y $H \cap (A - B) = \emptyset$. Obsérvese que $\beta = 0$, y así, se tiene que $f(A - B)$ es un convexo de $\mathbb{R}$ (un intervalo) tal que $0 \notin f(A - B)$. En consecuencia, o bien $f(a - b) > 0$, para cada $a \in A$ y $b \in B$, o bien $f(a - b) < 0$, para cada $a \in A$ y $b \in B$. Por ejemplo, en el segundo caso, $f(a) < f(b)$ para cada $a \in A$ y $b \in B$, y por lo tanto, podemos asegurar la existencia de $\alpha \in \mathbb{R}$ satisfaciendo

$$\sup\{f(a) : a \in A\} \leq \alpha \leq \operatorname{ínf}\{f(b) : b \in B\}.$$

Si A es abierto, entonces $f(A)$ es un intervalo abierto de $\mathbb{R}$ después del lema 1.3.23, y se concluye que $f(a) < \alpha$, para cada $a \in A$. En el caso de que ambos A y B sean abiertos, se tendrá también que $\alpha < f(b)$, para cada $b \in B$s y, en consecuencia, A y B estarán estrictamente separados por un hiperplano real cerrado. □

Para el segundo teorema de separación se necesita el siguiente lema, que queda como ejercicio.

Lema 1.3.25. *Sean $E[\mathfrak{T}]$ un* e.v.t.*, $K \subset E$ un conjunto compacto y $F \subset E$ un conjunto cerrado con $K \cap F = \emptyset$. Entonces, existe un entorno abierto del origen W tal que*

$$(K + W) \cap (F + W) = \emptyset.$$

Corolario 1.3.26 (Segundo teorema de separación)**.** *Sean* $E[\mathfrak{T}]$ *un* e.l.c. *y* K*,* F *subconjuntos convexos disjuntos de* E*, con* K *compacto y* F *cerrado. Entonces, existe un hiperplano real cerrado que separa estrictamente* K *y* F*. Más aún, existen* $f : E \longrightarrow \mathbb{R}$ *lineal y continua,* $\alpha \in \mathbb{R}$ *y* $\varepsilon > 0$*, tales que, para todo* $y \in K$ *y todo* $z \in F$*,*

$$f(y) \leq \alpha - \varepsilon < \alpha < f(z).$$

Demostración. Por el lema anterior, existe un entorno del origen abierto y absolutamente convexo W tal que $(K+W)\cap(F+W) = \emptyset$. Los conjuntos $K+W$ y $F+W$ son abiertos y convexos, véase el apartado *(iii)* de la proposición 1.3.1. Ahora, el corolario 1.3.24 nos garantiza la existencia de $f : E \longrightarrow \mathbb{R}$, lineal y continua, y $\alpha \in \mathbb{R}$, tales que

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

para todo $y \in K$ y todo $z \in F$. Como toda función real continua en un compacto alcanza su máximo, se tiene que $\max_{y\in K} f(y) < \alpha$ y, en consecuencia, para cierto $\varepsilon > 0$ se tendrá que

$$f(y) \leq \alpha - \varepsilon < \alpha < f(z),$$

para todo $y \in K$ y todo $z \in F$. $\square$

STOP Para conjuntos cerrados y convexos, el teorema anterior no es cierto. Basta tomar en $\mathbb{R}^2$ los cerrados convexos $K = \{(x,0) \in \mathbb{R}^2 : x \geq 0\}$ y $F = \{(x,y) \in \mathbb{R}_+ \times \mathbb{R}_+ : xy \geq 1\}$. El teorema anterior nos permite separar estrictamente, en cualquier *e.l.c.*, un conjunto convexo de cerrado de un punto que no pertenece a él. En el caso finito-dimensional $X = \mathbb{R}^n$, se puede separar (no necesariamente de forma estricta) cualquier convexo de un punto que no está en el convexo, véase [98, Theorem 1.3.4].

Corolario 1.3.27. *Sean* K *y* F *subconjuntos disjuntos de un* e.l.c. $E[\mathfrak{T}]$*, tales que* K *es compacto y* F *es cerrado. Si* K *se supone absolutamente convexo y* F *convexo, entonces existe* $u \in E'$ *tal que*

$$\sup_{x\in K} |u(x)| < \inf_{y\in F} |u(y)|.$$

Demostración. Demostramos el caso complejo. Tomamos $f : E \longrightarrow \mathbb{R}$, lineal y continua, y $\alpha \in \mathbb{R}$ como en el corolario anterior, y definimos $u(x) := f(x) - if(ix)$, para $x \in E$. Entonces u es lineal y satisface que $\mathrm{Re}\, u(x) = f(x)$ para cada $x \in E$. De esta forma tenemos que

(i) para cada $x \in K$, existe un cierto $\theta \in \mathbb{R}$ tal que $|u(x)| = e^{i\theta}u(x)$, y por lo tanto se tiene

$$|u(x)| = e^{i\theta}u(x) = u(e^{i\theta}x) = \mathrm{Re}\big(u(e^{i\theta}x)\big) \leq \alpha - \varepsilon;$$

(ii) $\alpha < \mathrm{Re}\, u(y) \leq |u(y)|$, para cada $y \in F$.

Así queda terminada la prueba. $\square$

Corolario 1.3.28. *Sean $E[\mathfrak{T}]$ un* e.l.c. *y A un subconjunto de E. La envoltura convexa y cerrada de A es la intersección de todos los semiespacios que contienen a A determinados por hiperplanos afines cerrados en E,* i.e.,

$$\overline{\mathrm{co}(A)} = \bigcap \{F_\alpha : A \subset F_\alpha,\ F_\alpha \text{ semiespacio cerrado}\}\,. \tag{1.8}$$

En particular, si A es convexo y cerrado, entonces

$$A = \bigcap \{F_\alpha : A \subset F_\alpha,\ F_\alpha \text{ semiespacio cerrado}\}\,.$$

Demostración. Claramente se tiene $\overline{\mathrm{co}(A)} \subset \bigcap \{F_\alpha : A \subset F_\alpha,\ F_\alpha$ semiespacio cerrado$\}$. Por otro lado, si $x \notin \overline{\mathrm{co}(A)}$, entonces existen una forma $f : E \longrightarrow \mathbb{R}$ lineal y continua y $\alpha \in \mathbb{R}$, tales que $f(x) > \alpha > f(y)$, para cada $y \in \overline{\mathrm{co}(A)}$. Esto significa que, para el semiespacio determinado por f y α, se tiene que $x \notin F_\alpha$, mientras que $A \subset F_\alpha$, lo que nos da la otra inclusión que necesitamos para concluir la igualdad (1.8). □

El siguiente corolario es una consecuencia inmediata de lo que acabamos de demostrar.

Corolario 1.3.29. *Sean E un espacio vectorial y $\mathfrak{T}$ y τ dos topologías localmente convexas separadas en E, tales que $(E[\mathfrak{T}])' = (E[\tau])'$. Entonces, los conjuntos convexos $\mathfrak{T}$-cerrados y τ-cerrados son los mismos.*

Si $E[\mathfrak{T}]$ es un *e.l.c.*, entonces $(E[\mathfrak{T}])' = (E[\sigma(E,E')])'$, véase la proposición 1.3.33 y, en consecuencia, los subconjuntos convexos $\mathfrak{T}$-cerrados y $\sigma(E,E')$-cerrados son los mismos.

Corolario 1.3.30. *Sean $E[\mathfrak{T}]$ un* e.l.c. *y $F \subset E$ un subespacio vectorial. Entonces*

$$\overline{F} = \big\{x \in E : f(x) = 0,\ \text{si } f \in E' \text{ y } f|_F = 0\big\}.$$

Demostración. Se sigue inmediatamente del hecho de que una forma lineal acotada en un subespacio vectorial se anula en este subespacio, y del corolario 1.3.28. □

Corolario 1.3.31. *Sean $E[\mathfrak{T}]$ un* e.l.c. *y S un subconjunto de E. Son equivalentes:*

(i) *S es total en $E[\mathfrak{T}]$,* i.e., $\overline{\mathrm{span}(S)}^{\mathfrak{T}} = E$.
(ii) *Si $f \in E'$ y $f|_S = 0$, entonces $f = 0$ en E.*

Stanislaw Mazur, 1905-1981. En la monografía de S. Banach *Théorie des opérations linéaires*, 1932, las propiedades de los conjuntos convexos eran prácticamente ignoradas. Sin embargo, de forma casi simultánea, 1933, S. Mazur había desarrollado ya una versión geométrica del teorema de Hahn-Banach, generalizando la teoría de Minkowski al demostrar que, si K es un subconjunto convexo de un espacio normado E, entonces, para cada punto en la frontera de K, existe un hiperplano cerrado que es soporte de K. Los teoremas de separación de conjuntos convexos han encontrado numerosas aplicaciones en geometría, análisis, optimización, etc. Mazur fue co-fundador con Banach de lo que se conoce como la *Escuela Polaca de Análisis Funcional*. Mazur fue estudiante de Banach, y uno de los artífices del *Scottish book*. S. Ulam, escribe:

> *Para aquéllos que no conocen el "Scottish book", debo empezar diciendo que es una colección informal de problemas en matemáticas. Fue empezado en Lwow, Polonia, en 1935. Realmente, algunos de los problemas se habían originado antes, quizás 6 ó 7 años antes.*

El *Scottish book* recoge los problemas que fueron discutidos por Banach, Mazur, Ulam y otros, cuando se reunían en el *Scottish Cafe* en Lwow. Para cada problema propuesto se ofrecía un premio por su solución. Hay una anécdota sobre uno de los más famosos problemas del *Scottish book*. Éste es el problema 153, propuesto por Mazur, cuyo premio consistía en un *ganso vivo*. El problema 153 estuvo originalmente formulado en términos de aproximaciones de funciones continuas en dos variables por combinaciones lineales de sus secciones. A. Grothendieck demostró en su tesis, en 1955, que el problema 153 es equivalente al problema de la aproximación, que en su tiempo fue considerado uno de los problemas centrales en el análisis funcional. El problema de la aproximación es el siguiente: ¿Es cada operador lineal compacto de un espacio de Banach X en otro espacio Y el límite en norma de una sucesión de operadores de rango finito? El espacio de Banach Y se dice que tiene la propiedad de la aproximación si la respuesta a la cuestión anterior es positiva para cada espacio de Banach X. Espacios de Hilbert y espacios de Banach con base de Schauder tienen la propiedad de la aproximación. La respuesta al problema de la aproximación (por tanto, al problema 153, y al problema de la base) es negativa. Esto fue probado en 1972 por P. Enflo. El ganso fue entregado por Mazur a Enflo un año después, cuando este último estaba dando una charla en Varsovia presentando la solución al problema de la aproximación.

1.3.5 Pares duales. Polares. El teorema del bipolar

Par dual

Un *par dual* $\langle F,G\rangle$ es dos espacios vectoriales F y G junto con una aplicación bilineal $\langle\cdot,\cdot\rangle : F\times G \longrightarrow \mathbb{K}$ satisfaciendo las propiedades:

(i) Si $\langle x,y\rangle = 0$ para cada $x\in F$, entonces $y=0$.

(ii) Si $\langle x,y\rangle = 0$ para cada $y\in G$, entonces $x=0$.

Dados $x_0\in F$ e $y_0\in G$, consideramos las aplicaciones lineales

$$f_{x_0} : G \longrightarrow \mathbb{K} \quad \text{y} \quad f^{y_0} : F \longrightarrow \mathbb{K}$$

dadas por $f_{x_0}(y) := \langle x_0, y\rangle$ y $f^{y_0}(x) := \langle x, y_0\rangle$, para $x \in F$ e $y \in G$. La aplicación $x \to f_x$ (respectivamente, $y \to f^y$) es lineal e inyectiva de F en un subespacio del dual algebraico de G, $G^{\#}$ (respectivamente, del dual algebraico de F, $F^{\#}$). Algunas veces, consideraremos F (respectivamente, G) ya identificado como subespacio de $G^{\#}$ (respectivamente, $F^{\#}$) a través de la correspondiente inyección anterior.

Topologías débiles Si $\langle F, G\rangle$ es un par dual, llamamos *topología débil* $\sigma(F,G)$ de F inducida por G, a la topología más gruesa en F para la cual las aplicaciones $f^y : F \longrightarrow \mathbb{K}$, para cada $y \in G$, son continuas. De forma dual, se define la topología débil $\sigma(G,F)$ de G como la topología más gruesa en G para la cual son continuas las aplicaciones $f_x : G \longrightarrow \mathbb{K}$, para cada $x \in F$. Obsérvese que $\sigma(F,G)$ es una topología localmente convexa separada, generada por la familia de seminormas

$$\mathscr{P} := \{p_y : y \in G\},$$

donde $p_y(x) = |\langle x, y\rangle|$, para cada $x \in F$.

Si $E[\mathfrak{T}]$ es un *e.l.c.* y consideramos la dualidad $\langle E, E'\rangle$, entonces las nociones de topologías débiles en el par dual coinciden con las nociones de topologías débiles y débiles* introducidas en la página 34.

Ejemplo 1.3.32.

(i) Si E es un espacio vectorial y $E^{\#}$ es su dual algebraico, $\langle E, E^{\#}\rangle$ es un par dual con la aplicación bilineal natural $\langle x, x^*\rangle = x^*(x)$ para cada $(x, x^*) \in E \times E^{\#}$.

(ii) Si $E[\mathfrak{T}]$ es un *e.l.c.* y E' es su dual topológico, entonces $\langle E, E'\rangle$ es un par dual con la aplicación bilineal $\langle x, x'\rangle = x'(x)$, gracias al apartado *(ii)* del corolario 1.3.21.

(iii) Si I es un conjunto de índices y $\mathbb{K}_i = \mathbb{K}$, para cada $i \in I$, entonces $\left\langle \prod_{i\in I} \mathbb{K}_i, \bigoplus_{i\in I} \mathbb{K}_i \right\rangle$ es un par dual con la aplicación bilineal $\left\langle (\lambda_i)_{i\in I}, (\xi_i)_{i\in I} \right\rangle = \sum_{i\in I} \lambda_i \xi_i$.

(iv) Sea $E = (C(K), \|\cdot\|_\infty)$, con K compacto. Para cada $x \in K$, consideremos $\delta_x : C(K) \longrightarrow \mathbb{K}$ el funcional lineal continuo $\delta_x(f) := f(x)$. Definamos $\widehat{K} = \{\delta_x : x \in K\} \subset (C(K), \|\cdot\|_\infty)^*$, y sea $F = \operatorname{span}(\widehat{K})$. Entonces, $\langle E, F\rangle$ es un par dual con la aplicación bilineal inducida por la aplicación bilineal natural del par dual $\left\langle C(K), (C(K), \|\cdot\|_\infty)^* \right\rangle$. Obsérvese que, si consideramos el par dual $\langle E, F\rangle$, entonces la topología débil $\sigma(E,F) = \sigma\big(C(K), \operatorname{span}(\widehat{K})\big) = \tau_p$ es la topología de la convergencia puntual inducida por $\mathbb{K}^K$. $\square$

Proposición 1.3.33. *Si $\langle F, G\rangle$ es un par dual y $f : F \longrightarrow \mathbb{K}$ es una forma lineal, entonces f es $\sigma(F,G)$-continua si, y sólo si, existe $y \in G$ (necesariamente único) tal que $f = f^y$.*

Demostración. Es claro que cada f^y es $\sigma(F,G)$-continua, por definición de $\sigma(F,G)$. Recíprocamente, sea $f : F \longrightarrow \mathbb{K}$ lineal $\sigma(F,G)$-continua. Caso de existir $y \in G$ satisfaciendo $f = f^y$, este y debe ser único por la propiedad *(i)* de la definición de par dual dada en la página 46. La existencia de y se razona de la siguiente forma: dado que f es $\sigma(F,G)$-continua, de acuerdo con la proposición 1.3.8, existe una seminorma $\sigma(F,G)$-continua p definida en F tal que $|f(x)| \le p(x)$, para todo $x \in F$. Puesto que $\{x \in F : p(x) < 1\}$ es un entorno del origen en $\sigma(F,G)$, existen vectores $y_1, y_2, \dots, y_n \in G$ de modo que

$$\{x : |\langle x, y_i\rangle| < 1,\ i = 1, \dots, n\} \subset \{x : p(x) < 1\}.$$

Esta inclusión implica que $p(x) \leq \sup_{1\leq i\leq n} |f^{y_i}(x)|$, para cada $x \in F$, y por tanto,

$$|f(x)| \leq \sup_{1\leq i\leq n} |f^{y_i}(x)|,$$

para cada $x \in F$. De la desigualdad anterior se obtiene que $\bigcap_{i=1}^{n} \ker f^{y_i} \subset \ker f$. De aquí, se sigue la existencia de escalares $\lambda_1, \lambda_2, \dots, \lambda_n$ tales que $f = \sum_{i=1}^{n} \lambda_i f^{y_i}$, véase el lema 1.2.4, lo que significa que si $y = \sum_{i=1}^{n} \lambda_i y_i$, entonces $f = f^y$, y así acaba la demostración. $\square$

Cuando $E[\mathfrak{T}]$ es un *e.l.c.* se tiene que $E[\sigma(E,E')]' = E'$ y, después de la oportuna identificación, también se tiene que $E'[\sigma(E',E)]' = E$.

Corolario 1.3.34. *Sean $\langle F,G\rangle$ un par dual y $G_1 \subset G$ un subespacio vectorial. La aplicación bilineal $\langle\cdot,\cdot\rangle$ induce un par dual $\langle F, G_1\rangle$ si, y sólo si, $\overline{G_1}^{\sigma(G,F)} = G$.*

Demostración. Después de la proposición anterior tenemos que $G = F[\sigma(F,G)]'$. Para $G_1 \subset G$, el corolario 1.3.31 nos dice que $\overline{G_1}^{\sigma(G,F)} = G$ si, y sólo si, cada forma $f_x \in G[\sigma(G,F)]'$ satisfaciendo $f_x|_{G_1} = 0$ es idénticamente nula. Esta condición significa que, si $x \in F$ y $\langle x,y\rangle = 0$ para cada $y \in G_1$, entonces $x = 0$, es decir, $\langle F, G_1\rangle$ es un par dual. $\square$

Obsérvese que, como consecuencia de lo demostrado en el corolario anterior, se obtiene que si $E[\mathfrak{T}]$ es un *e.l.c.*, entonces el dual topológico E' es $\sigma(E^{\#},E)$-denso en el dual algebraico $E^{\#}$, *i.e.*, las formas lineales se aproximan por formas lineales y continuas.

Polares Sean $\langle F,G\rangle$ un par dual y A un subconjunto de F. Llamamos *polar* (absoluta) de A al subconjunto A° de G dado por

$$A^{\circ} = \{y \in G : |\langle x,y\rangle| \leq 1, \text{ para todo } x \in A\}.$$

Llamaremos bipolar de A al subconjunto $A^{\circ\circ} = (A^{\circ})^{\circ}$ de F.

Ejemplo 1.3.35.

(i) Sea $(X, \|\cdot\|)$ un espacio normado, consideremos $\langle X, X^*\rangle$ y sea $A = B_X$ la bola unidad cerrada. Entonces,

$$B_X^{\circ} = \{x^* \in X^* : |x^*(x)| \leq 1,\ x \in B_X\} = B_{X^*},$$

$$B_X^{\circ\circ} = \{x \in X : |x^*(x)| \leq 1,\ x^* \in B_{X^*}\} = B_X.$$

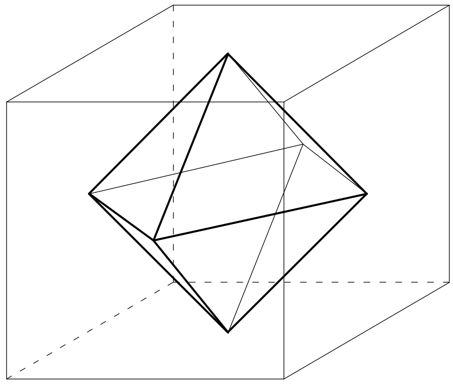


Figura 1.4: Polar de $B_{(\mathbb{R}^3,\|\cdot\|_\infty)}$

(ii) Sean $\langle F,G\rangle$ un par dual y $M \subset F$ un subespacio vectorial. Entonces,

$$M^\circ = \{y \in G : |\langle x,y\rangle| \leq 1, \text{ para todo } x \in M\}$$
$$= \{y \in G : \langle x,y\rangle = 0, \text{ para todo } x \in M\} =: M^\perp. \qquad \square$$

Proposición 1.3.36. *Sean $\langle F,G\rangle$ un par dual y A un subconjunto de F. Entonces:*

(i) A° *es un conjunto absolutamente convexo y* $\sigma(G,F)$*-cerrado.*
(ii) Si $A_1 \subset A$*, entonces* $A^\circ \subset A_1^\circ$.
(iii) Si $\alpha \in \mathbb{K}$, $\alpha \neq 0$*, entonces* $(\alpha A)^\circ = \alpha^{-1}A^\circ$.
(iv) $A \subset A^{\circ\circ}$.
(v) $A^\circ = A^{\circ\circ\circ}$.
(vi) $\left(\bigcup_{i\in I} A_i\right)^\circ = \bigcap_{i\in I} A_i^\circ$.

Demostración. Estas propiedades son de comprobación inmediata y se dejan como ejercicio. $\square$

Teorema 1.3.37 (Teorema del bipolar)**.** *Sean $\langle F,G\rangle$ un par dual y $A \subset F$ un subconjunto. Entonces, $A^{\circ\circ}$ es la envoltura absolutamente convexa $\sigma(F,G)$-cerrada de A.*

Demostración. Sea $A_1 := \overline{\Gamma(A)}$ la envoltura absolutamente convexa y cerrada de A. Como $A^{\circ\circ}$ es absolutamente convexo y cerrado, se tiene que $A_1 \subset A^{\circ\circ}$. Para probar la otra inclusión demostraremos que si $x_0 \in F \setminus A_1$, entonces $x_0 \in F \setminus A^{\circ\circ}$. Como $F[\sigma(F,G)]$ es un *e.l.c.*, A_1 es cerrado y absolutamente convexo y $\{x_0\}$ es convexo compacto, disjunto de A_1, podemos aplicar el corolario 1.3.27 para obtener $u \in G$ y $\alpha \in \mathbb{R}$ satisfaciendo

$$\sup_{x\in A_1} \left|\langle x,u\rangle\right| < \alpha < \alpha + \varepsilon < \left|\langle x_0,u\rangle\right|.$$

Puesto que $0 \in A_1$, se tiene que $\alpha > 0$. Reemplazando u por $u' = \alpha^{-1}u$ y denotando por $\varepsilon' = \alpha^{-1}\varepsilon$, se tiene que

$$\sup_{x \in A_1} \left|\langle x, u'\rangle\right| < 1 < 1 + \varepsilon' < \left|\langle x_0, u'\rangle\right|.$$

De aquí se deduce que $x_0 \in F \setminus A^{\circ\circ}$, con lo que acaba la demostración. □

Conjuntos equicontinuos Sea $E[\mathfrak{T}]$ un *e.l.c.* Un conjunto $M \subset E'$ se dice que es $\mathfrak{T}$*-equicontinuo* si, para cada $\varepsilon > 0$, existe un $\mathfrak{T}$-entorno del origen U en E tal que $|x'(x)| < \varepsilon$, para cada $x \in U$ y cada $x' \in M$. Una familia de $\mathfrak{T}$-equicontinuos $\mathscr{E}$ en E' se dice *fundamental* si, para cada conjunto $\mathfrak{T}$-equicontinuo $M \subset E'$, existe $N \in \mathscr{E}$ tal que $M \subset N$.

Corolario 1.3.38. *Sean $E[\mathfrak{T}]$ un* e.l.c. *y $\langle E, E'\rangle$ el par dual canónico asociado.*

(i) Si $U \subset E$ es un $\mathfrak{T}$-entorno del origen, entonces U° es $\mathfrak{T}$-equicontinuo. Recíprocamente, si $M \subset E'$ es $\mathfrak{T}$-equicontinuo, entonces M° es un entorno del origen.

(ii) Si $\mathscr{U}$ es una base de entornos del origen de $E[\mathfrak{T}]$, entonces las polares $\{U^\circ\}_{U \in \mathscr{U}}$ forman una familia fundamental de equicontinuos de E'.

(iii) Si $\mathscr{E}$ es una familia fundamental de $\mathfrak{T}$-equicontinuos de E', entonces la familia de las polares $\{M^\circ\}_{M \in \mathscr{E}}$ forma una base de $\mathfrak{T}$-entornos del origen en E.

Demostración. Demostremos *(i)*. Si U es un $\mathfrak{T}$-entorno del origen, entonces U° es equicontinuo, ya que dado $\varepsilon > 0$, para cada $x \in \varepsilon U$ y cada $x' \in U^\circ$, se tiene $|x'(x)| \leq \varepsilon$. Recíprocamente, si $M \subset E'$ es $\mathfrak{T}$-equicontinuo, entonces existe un $\mathfrak{T}$-entorno del origen $U \subset E$ tal que $|x'(x)| < 1$, para cada $x \in U$ y cada $x' \in M$. Por definición de polar se tiene que $U \subset M^\circ$.

Los dos apartados *(ii)* y *(iii)* se prueban de forma similar. Demostraremos *(iii)*, dejando *(ii)* como ejercicio. Si $\mathscr{E}$ es una familia fundamental de equicontinuos, entonces M° es $\mathfrak{T}$-entorno del origen para cada $M \in \mathscr{E}$. Para terminar la prueba será suficiente comprobar que, para cualquier $\mathfrak{T}$-entorno del origen $U \subset E$ absolutamente convexo y $\mathfrak{T}$-cerrado ($\sigma(E, E')$-cerrado después del corolario 1.3.28), existe $M \in \mathscr{E}$ tal que $M^\circ \subset U$. Como U es $\mathfrak{T}$-entorno del origen, su polar U° es $\mathfrak{T}$-equicontinuo y, por lo tanto, existe $M \in \mathscr{E}$ tal que $U^\circ \subset M$. El teorema del bipolar y la proposición 1.3.36 nos aseguran ahora que $M^\circ \subset U^{\circ\circ} = U$, y así acaba la demostración. □

Corolario 1.3.39. *Si $E[\mathfrak{T}]$ es un* e.l.c.*, entonces $\mathfrak{T}$ es la topología de convergencia uniforme sobre los $\mathfrak{T}$-equicontinuos de E'.*

Corolario 1.3.40. *Sea $\langle F, G\rangle$ un par dual y sea $\{A_i : i \in I\}$ una familia de subconjuntos de F $\sigma(F, G)$-cerrados absolutamente convexos. Entonces:*

$$\Big(\bigcap_{i \in I} A_i\Big)^\circ = \overline{\Gamma\Big(\bigcup_{i \in I} A_i^\circ\Big)}^{\sigma(F,G)}.$$

Demostración. El teorema del bipolar nos asegura que $A_i^{\circ\circ} = A_i$. Se tiene por tanto que

$$\Big(\bigcup_{i\in I} A_i^{\circ}\Big)^{\circ} = \bigcap_{i\in I} A_i^{\circ\circ} = \bigcap_{i\in I} A_i.$$

Tomando ahora polares en la igualdad anterior y haciendo uso, otra vez, del teorema del bipolar se obtiene el resultado deseado. □

Teorema 1.3.41 (Alaoglu-Bourbaki)**.** *Sea* $E[\mathfrak{T}]$ *un* e.l.c. *Todo subconjunto* $\mathfrak{T}$*-equicontinuo* M *de* E' *es* $\sigma(E',E)$*-relativamente compacto.*

Demostración. Sea $M \subset E'$ un subconjunto $\mathfrak{T}$-equicontinuo. Considérese el espacio $E'[\sigma(E',E)]$ como subespacio de $(\mathbb{K}^E, \tau_p(E))$. Para ver que M es $\sigma(E',E)$-relativamente compacto es suficiente, después del teorema de Tychonoff, probar que $\overline{M}^{\sigma(E',E)} = \overline{M}^{\tau_p(E)} \subset E'$, y que M es puntualmente acotado. Ahora bien, $\overline{M}^{\tau_p(E)}$ está formado por aplicaciones lineales, ya que sus elementos son límites puntuales de redes equicontinuas. Para ver que M es acotado razonamos como sigue. Sea U un $\mathfrak{T}$-entorno del origen tal que $|x'(y)| < 1$, para cada $x' \in M$ e $y \in U$. Dado $x \in E$, sea $\rho > 0$ tal que $\rho x \in U$. De aquí se tiene que el conjunto $\{|x'(x)| : x' \in M\}$ está acotado por $1/\rho$. □

Corolario 1.3.42. *Sea* $(X, \|\cdot\|)$ *un espacio normado. Entonces, la bola unidad cerrada* B_{X^*} *es* $\sigma(X^*,X)$*-compacta en* X^*.

1.3.6 Topologías débiles en espacios de Banach. Reflexividad

Si X es un espacio de Banach, entonces su dimensión algebraica es finita o es no numerable. Para convencerse de que esto es así, se demuestra que X no puede tener dimensión algebraica numerable: Supongamos que la tiene, y sea $\{e_n\}_n$ una base algebraica. Entonces $X = \bigcup_{n=1}^{\infty} X_n$, con $X_n = \operatorname{span}\{e_i : 1 \le i \le n\}$. Como los X_n son subespacios propios, no tienen puntos interiores, siendo además cerrados debido al lema 1.3.20. Por tanto, X es de primera categoría en sí mismo, lo que contradice el teorema de Baire 1.1.15, y llegamos así a una contradicción que termina la prueba.

Proposición 1.3.43. *Sea* $(X, \|\cdot\|)$ *un espacio normado. Se tienen las siguientes propiedades:*

(i) *La topología débil* $\sigma(X,X^*)$ *(respectivamente, débil** $\sigma(X^*,X)$*) es más gruesa que la topología asociada a la norma en* X *(respectivamente, la norma de* X^**).*

(ii) *La topología débil* $\sigma(X,X^*)$ *es metrizable si, y sólo si,* X *es de dimensión finita. En este caso,* $\sigma(X,X^*)$ *coincide con la topología asociada a la norma.*

(iii) *Un subconjunto convexo de* X *es* $\sigma(X,X^*)$*-cerrado si, y sólo si, es cerrado en la topología asociada a la norma.*

(iv) *La aplicación* $\widehat{\ }: X \longrightarrow X^{**}$ *dada por* $\widehat{x}(x^*) = x^*(x)$, $x^* \in X^*$, *es un isomorfismo isométrico sobre su imagen. Sumergido* X *en* X^{**} *vía* $\widehat{\ }$, X *es* $\sigma(X^{**},X^*)$*-denso en* X^{**}.

Demostración. La propiedad *(i)* es inmediata de comprobar. Demostremos *(ii)*. Si X es de dimensión finita, entonces el lema 1.3.20 nos dice que la topología débil y la asociada a la norma coinciden, y por lo tanto, $\sigma(X,X^*)$ es metrizable. Recíprocamente, supongamos que $\sigma(X,X^*)$ es metrizable, y sean $x_1^*, x_2^*, \ldots, x_n^*, \ldots$ una sucesión en X^* y $(\varepsilon_n)_n$ una sucesión de números reales positivos tales que $\{V(0,x_1^*,\ldots,x_n^*,\varepsilon_n)\}_n$ es una base de entornos del origen para $\sigma(X,X^*)$. Dado $x^* \in X^*$, el entorno $V(0,x^*,1)$ contendrá cierto entorno $V(0,x_1^*,\ldots,x_n^*,\varepsilon_n)$ y, en consecuencia, $\bigcap_{i=1}^n \ker x_i^* \subset V(0,x^*,1)$. De aquí se obtiene que x^* está acotada en el subespacio $\bigcap_{i=1}^n \ker x_i^*$, y por tanto, x^* debe anularse en él, lo que significa que $\bigcap_{i=1}^n \ker x_i^* \subset \ker x^*$. Si utilizamos el lema 1.2.4, obtenemos que x^* es una combinación lineal de $x_1^*, x_2^*, \ldots, x_n^*$, y de esta forma concluimos que X^* no es de dimensión no numerable, siendo, en consecuencia, de dimensión finita después de la observación previa a esta proposición. Si el espacio dual X^* es de dimensión finita, entonces el espacio X también es de dimensión finita, gracias al apartado (iii) del corolario 1.3.21; la prueba de esta parte está terminada.

La afirmación recogida en *(iii)* es un caso particular del corolario 1.3.29. La primera parte de *(iv)* es consecuencia del teorema de Hahn-Banach. La segunda parte se obtiene del corolario 1.3.34, dado que el par dual $\langle X^*, X^{**}\rangle$ induce, de forma natural, el par dual $\langle X^*, X\rangle$. □

$\boxed{\textit{Reflexividad}}$ Un espacio de Banach $(X, \|\cdot\|)$ se dice *reflexivo* si la aplicación $\widehat{\ }: X \longrightarrow X^{**}$ dada por $\widehat{x}(x^*) = x^*(x)$, $x^* \in X^*$, es sobreyectiva. En otras palabras, tras la correspondiente identificación de X con $\widehat{X}$, X es reflexivo si, y sólo si, $X = X^{**}$.

STOP Obsérvese que la noción de espacio reflexivo ha de estar, necesariamente, ligada a espacios de Banach, ya que el bidual X^{**} de un espacio normado es un espacio de Banach.

Ejemplo 1.3.44.

(i) Si $1 < p < \infty$, entonces $(\ell^p, \|\cdot\|_p)$ es reflexivo, dado que, para $1/p + 1/q = 1$,

$$(\ell^p, \|\cdot\|_p)^{**} = (\ell^q, \|\cdot\|_q)^* = (\ell^p, \|\cdot\|_p).$$

(ii) Más en general, sean $1 < p < \infty$ y (Ω, Σ, μ) un espacio de medida. Entonces, el espacio $(L^p(\mu), \|\cdot\|_p)$ es reflexivo, dado que

$$(L^p(\mu), \|\cdot\|_p)^{**} = (L^q(\mu), \|\cdot\|_q)^* = (L^p(\mu), \|\cdot\|_p),$$

donde $1/p + 1/q = 1$, véase el corolario 1.3.53.

(iii) El espacio $(c_0, \|\cdot\|_\infty)$ es no reflexivo, dado que se tienen las igualdades

$$(c_0, \|\cdot\|_\infty)^{**} = (\ell^1, \|\cdot\|_1)^* = (\ell^\infty, \|\cdot\|_\infty),$$

siendo $c_0 \neq \ell^\infty$. □

El siguiente resultado, que es una mejora de la segunda parte de *(iv)* en la proposición 1.3.43, es fundamental para caracterizar los espacios de Banach reflexivos.

Teorema 1.3.45 (Goldstine)**.** *Sean $(X, \|\cdot\|)$ un espacio normado, B_X su bola unidad y $B_{X^{**}}$ la bola unidad de X^{**}. Si consideramos X como un subespacio de X^{**} (vía $\widehat{\ }$), entonces B_X es un subconjunto $\sigma(X^{**},X^*)$-denso de $B_{X^{**}}$.*

Demostración. Claramente $B_X \subset B_{X^{**}}$. El teorema de Alaoglu-Bourbaki 1.3.41 nos asegura que $B_{X^{**}}$ es $\sigma(X^{**},X^*)$-compacto, y por lo tanto se tiene que $\overline{B_X}^{\sigma(X^{**},X^*)} \subset B_{X^{**}}$. Si suponemos que la inclusión es estricta, entonces podemos tomar $x^{**} \in B_{X^{**}} \setminus \overline{B_X}^{\sigma(X^{**},X^*)}$. Haciendo uso ahora del corolario 1.3.27, existe $x^* \in X^*$ de manera que se verifican las siguientes desigualdades,

$$\|x^*\| = \sup_{x \in B_X} \left|x^*(x)\right| < \left|x^{**}(x^*)\right| \leq \|x^{**}\| \, \|x^*\| \leq \|x^*\|,$$

con lo que se llega a una contradicción que termina la demostración. □

Teorema 1.3.46 (Caracterización de la reflexividad)**.** *Sea $(X, \|\cdot\|)$ un espacio de Banach. Las siguientes afirmaciones son equivalentes:*

(i) X es reflexivo.
(ii) La bola unidad B_X es $\sigma(X,X^*)$-compacta.

Demostración. Veamos cómo *(i)* ⇒ *(ii)*. Si X es reflexivo, entonces $X = X^{**}$ y, en consecuencia, $B_X = B_{X^{**}}$ es $\sigma(X,X^*)$-compacto por el teorema de Alaoglu-Bourbaki 1.3.41.

Para la implicación *(ii)* ⇒ *(i)* razonamos como sigue. Si B_X es $\sigma(X,X^*)$-compacto, entonces B_X debe ser igual a $B_{X^{**}}$, dado que B_X es $\sigma(X^{**},X^*)$-cerrado y $\sigma(X^{**},X^*)$-denso en $B_{X^{**}}$ después del teorema de Goldstine. Si $B_X = B_{X^{**}}$, entonces $X = X^{**}$, y la prueba termina. □

Corolario 1.3.47. *Sea $(X, \|\cdot\|)$ un espacio de Banach.*

(i) Si X es reflexivo e $Y \subset X$ es un subespacio cerrado, entonces Y es reflexivo.
(ii) X es reflexivo si, y sólo si, $(X^*, \|\cdot\|)$ es reflexivo.
(iii) Si X es reflexivo, entonces X es separable si, y sólo si, X^* es separable.

Demostración. La propiedad *(i)* se sigue inmediatamente de la caracterización dada en el teorema anterior, teniendo en cuenta que Y es $\sigma(X,X^*)$-cerrado en X, véase *(iii)* en la proposición 1.3.43, y que gracias al teorema de Hahn-Banach 1.3.19 se tiene que $\sigma(X,X^*)|_Y = \sigma(Y,Y^*)$. Demostremos *(ii)*: si X es reflexivo, entonces X^* es, evidentemente, reflexivo. Recíprocamente, si X^* es reflexivo, entonces X^{**} es reflexivo, y su subespacio cerrado X también lo es por lo probado en el apartado *(i)*. Para la propiedad *(iii)* utilizamos la lema 1.3.48 que se demuestra a continuación: si X^* es separable, entonces X es separable. Por otro lado, cuando X es reflexivo, si $X = (X^*)^*$ es separable, entonces X^* es separable. □

Lema 1.3.48. *Si $(X, \|\cdot\|)$ es un espacio de Banach con dual separable, entonces X es separable.*

Demostración. Como X^* es separable, encontramos $x_n^* \in S_{X^*}$, $n \in \mathbb{N}$, tales que $\overline{\{x_n^*\}_{n=1}^\infty}^{\|\cdot\|} = S_{X^*}$. Sean entonces $x_n \in S_X$, $n \in \mathbb{N}$, tales que $x_n^*(x_n) \geq 1/2$. Vamos a probar que $\overline{\operatorname{span}\{x_n : n \in \mathbb{N}\}}^{\|\cdot\|} = X$. Para ello, supongamos que $X \setminus \overline{\operatorname{span}\{x_n : n \in \mathbb{N}\}}^{\|\cdot\|} \neq \emptyset$. Entonces, en virtud del teorema de Hahn-Banach, existe $x^* \in S_{X^*}$ tal que $x^*(x_n) = 0$, para todo $n = 1, 2, \ldots$ En consecuencia,

$$\|x_n^* - x^*\| \geq (x_n^* - x^*)(x_n) = x_n^*(x_n) \geq 1/2, \quad n = 1, 2, \ldots$$

lo que nos da la una contradicción que concluye la prueba. □

Con los resultados anteriores podemos completar la discusión sobre qué espacios de Banach clásicos son reflexivos y cuáles no.

Ejemplo 1.3.49.

(i) Todo espacio de dimensión finita es reflexivo.
(ii) Ni ℓ^1 ni ℓ^∞ son reflexivos, ya que ℓ^1 es separable y $\ell^\infty = (\ell^1, \|\cdot\|_1)^*$ no es separable.
(iii) Ni $C([a,b])$ ni su dual, $\big(C([a,b]), \|\cdot\|_\infty\big)^*$, son reflexivos, ya que $C([a,b])$ es separable y su dual no.
(iv) Ni $L^1([a,b])$ ni su dual, $L^\infty([a,b]) = \big(L^1([a,b]), \|\cdot\|_1\big)^*$, son reflexivos, ya que $L^1([a,b])$ es separable y $L^\infty([a,b])$ no lo es. □

Norma estrictamente convexa Una norma $\|\cdot\|$ en un espacio vectorial X se dice *estrictamente convexa* si $\left\|\frac{1}{2}(x+y)\right\| < 1$, siempre que $\|x\| = \|y\| = 1$ y $x \neq y$.

Norma uniformemente convexa Una norma $\|\cdot\|$ en X se dice *uniformemente convexa* si se cumple la siguiente condición: para cada $\varepsilon > 0$, existe un $\delta(\varepsilon) > 0$ tal que, si

$$x, y \in X, \quad \text{con } \|x\| \leq 1, \ \|y\| \leq 1, \ \text{y } \|x - y\| \geq \varepsilon,$$

entonces

$$\left\|\frac{1}{2}(x+y)\right\| \leq 1 - \delta(\varepsilon).$$

Si una norma es uniformemente convexa, entonces también es estrictamente convexa. Por otro lado, si $\big(H, \langle\cdot,\cdot\rangle\big)$ es un espacio prehilbertiano, la norma asociada es uniformemente convexa, gracias a que se satisface la Ley del Paralelogramo.

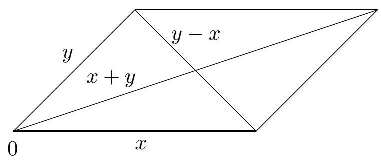


Figura 1.5: Ley del Paralelogramo: $\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$.

Efectivamente, si $x,y \in H$ son tales que $\|x\| \leq 1$ y $\|y\| \leq 1$, con $\|x-y\| \geq \varepsilon$, entonces

$$\left\|\frac{1}{2}(x+y)\right\|^2 = \frac{1}{4}\left(2\|x\|^2 + 2\|y\|^2 - \|x-y\|^2\right) \leq 1 - \frac{\varepsilon^2}{4}.$$

Proposición 1.3.50. *Sea $\left(X, \|\cdot\|\right)$ un espacio normado. Son equivalentes:*

(i) La norma $\|\cdot\|$ es uniformemente convexa.

(ii) Para cada par de sucesiones $(x_n)_n$ e $(y_n)_n$ en X, con $\|x_n\| \leq 1$ y $\|y_n\| \leq 1$, $n \in \mathbb{N}$, si

$$\lim_n \left\|\frac{1}{2}(x_n + y_n)\right\| = 1,$$

entonces $\lim_n \|x_n - y_n\| = 0$.

Demostración. Veamos *(i)* $\Rightarrow$ *(ii)*. Supongamos que *(ii)* no se da: entonces existen $x_n, y_n \in X$, con $\|x_n\| \leq 1$, y $\|y_n\| \leq 1$, $n \in \mathbb{N}$, tales que

$$\lim_n \left\|\frac{1}{2}(x_n + y_n)\right\| = 1,$$

sin ser cero el límite $\lim_n \|x_n - y_n\|$. Para cierto $\varepsilon > 0$ y para subsucesiones adecuadas, se tendrá que $\|x_{n_k} - y_{n_k}\| \geq \varepsilon$. Como también se tiene $\lim_k \left\|\frac{1}{2}(x_{n_k} + y_{n_k})\right\| = 1$, llegamos a que *(i)* no es cierta, y esta parte de la prueba está acabada. La implicación *(ii)* $\Rightarrow$ *(i)* la dejamos como ejercicio. □

Proposición 1.3.51. *Si (Ω, Σ, μ) es un espacio de medida y $2 \leq p < \infty$, $\left(L^p(\Omega, \Sigma, \mu), \|\cdot\|_p\right)$ es un espacio uniformemente convexo.*

Demostración. El resultado es consecuencia de la desigualdad

$$|\alpha + \beta|^p + |\alpha - \beta|^p \leq 2^{p-1}\left(|\alpha|^p + |\beta|^p\right), \text{ para todo } \alpha, \beta \in \mathbb{C}. \tag{1.9}$$

Supuesto que la desigualdad anterior ha sido demostrada, si $f, g \in L^p(\Omega, \Sigma, \mu)$, entonces

$$\|f+g\|_p^p + \|f-g\|_p^p \leq 2^{p-1}\left(\|f\|_p^p + \|g\|_p^p\right). \tag{1.10}$$

Si $(f_n)_n$ y $(g_n)_n$ son sucesiones en $L^p(\Omega, \Sigma, \mu)$, con $\|f_n\|_p \leq 1$, $\|g_n\|_p \leq 1$ y $\lim_n \left\|\frac{1}{2}(f_n + g_n)\right\|_p = 1$, la desigualdad (1.10) implica entonces que

$$2^p + \limsup_n \|f_n - g_n\|_p^p \leq 2^{p-1} 2,$$

es decir, $\lim_n \|f_n - g_n\|_p = 0$. La proposición 1.3.50 nos dice ahora que $\|\cdot\|_p$ es uniformemente convexa.

Para terminar la prueba, demostramos la desigualdad (1.9). Si $p \geq 2$, entonces

$$\left(|\alpha+\beta|^p + |\alpha-\beta|^p\right)^{1/p} \leq \left(|\alpha+\beta|^2 + |\alpha-\beta|^2\right)^{1/2} \leq \sqrt{2}\left(|\alpha|^2 + |\beta|^2\right)^{1/2}.$$

Usando la desigualdad de Hölder para $\frac{2}{p}+\frac{p-2}{p}=1$ (véase [96, Teorema 3.5]), obtenemos que

$$|\alpha|^2+|\beta|^2 \leq \big(|\alpha|^p+|\beta|^p\big)^{2/p}(1+1)^{(p-2)/2} \leq \big(|\alpha|^p+|\beta|^p\big)^{2/p}2^{(p-2)/2}.$$

Combinando las dos últimas desigualdades se obtiene (1.9), y con ello concluimos la prueba. $\square$

La proposición anterior también es cierta para $1<p<2$, pero su demostración requiere de otros cálculos distintos a los que hemos realizado: el lector interesado puede encontrar una prueba para este caso en [71, p. 358] (véase también [50]).

Diferenciabilidad Fréchet de la norma La noción de convexidad uniforme está ligada, vía dualidad, a la noción de diferenciabilidad. De forma más precisa, se verifica el siguiente teorema de Šmulian, véanse [71, p. 365] y [50, p. 290]: una norma $\|\cdot\|$ en un espacio de Banach X es uniformemente convexa si, y sólo si, su norma dual es *uniformemente diferenciable Fréchet*, es decir, si $\lim_{t\to 0} t^{-1}\big(\|x^*+th^*\|-\|x^*\|\big)$ existe uniformemente para x^*, h^* en la esfera unidad de X^*.

Teorema 1.3.52 (Milman, 1938)**.** *Sea $(X,\|\cdot\|)$ un espacio de Banach con una norma uniformemente convexa. Entonces X es reflexivo.*

Demostración. Demostraremos directamente que se tiene $(X,\|\cdot\|)^{**}=X$. Para ello es suficiente ver que, si $z\in X^{**}$ y $\|z\|=1$, entonces $z\in X$. Por el teorema de Goldstine 1.3.45, existe una red $(x_i)_{i\in D}$ en X tal que $\lim_i x_i = z$ en la topología débil* $\sigma(X^{**},X^*)$. Si vemos que la red $(x_i)_{i\in D}$ es de Cauchy en $(X,\|\cdot\|)$, se tendrá que $z\in X$, y nuestra prueba habrá terminado. Efectivamente, como $\|\cdot\|$ es uniformemente convexa, dado $\varepsilon>0$, existe $\delta(\varepsilon)>0$ tal que, si

$$x,y\in X, \quad \text{con } \|x\|\leq 1,\ \|y\|\leq 1,\ \text{ y } \|x-y\|\geq\varepsilon,$$

entonces

$$\left\|\frac{1}{2}(x+y)\right\| \leq 1-\delta(\varepsilon).$$

Tomemos $f\in X^*$ tal que $f(z)>1-\delta(\varepsilon)$ y $\|f\|=1$. Por la convergencia débil* de la red $(x_i)_{i\in D}$, existe $k\in D$ tal que, si $i\geq k$, entonces $f(x_i)>1-\delta(\varepsilon)$. Por lo tanto, si $i,j\geq k$, tenemos que

$$f\left(\frac{1}{2}(x_i+x_j)\right) > 1-\delta(\varepsilon)$$

y, consecuentemente,

$$\left\|\frac{1}{2}(x_i+x_j)\right\| > 1-\delta(\varepsilon).$$

La última desigualdad implica que $\|x_i-x_j\|\leq\varepsilon$ si $i,j\geq k$, y así, $(x_i)_{i\in D}$ es una red de Cauchy, como queríamos demostrar. $\square$

$\boxed{\textit{Espacios superreflexivos}}$ Un espacio de Banach X se dice *superreflexivo* si, para cada $\varepsilon > 0$, existe $N(\varepsilon) \in \mathbb{N}$ tal que cada ε-árbol diádico contenido en la bola unidad B_X tiene longitud $N \leq N(\varepsilon)$. Dado $\varepsilon > 0$, un ε-árbol diádico con raíz $x \in X$, de longitud $N \in \mathbb{N} \cup \{\infty\}$, es una familia $\{x(s)\}_s$ de elementos de X indicada en $s \in \{-1,1\}^{<N+1}$, tal que

$$\begin{aligned} &x = x(\emptyset), \\ &x(s) = \tfrac{1}{2}\big(x(s,-1) + x(s,1)\big), \quad \text{y} \\ &\|x(s,-1) - x(s,1)\| \geq \varepsilon, \end{aligned}$$

para cada $s \in \{-1,1\}^{<N}$. Un profundo resultado de Enflo, véase [46], establece que un espacio de Banach X admite una norma equivalente uniformemente convexa si, y sólo si, es superreflexivo.

Corolario 1.3.53. *Sea (Ω, Σ, μ) un espacio de medida. Si $1 < p < \infty$, entonces $L^p(\Omega, \Sigma, \mu)$ es reflexivo. Más en concreto, si $\frac{1}{p} + \frac{1}{q} = 1$, entonces la aplicación Φ que hace corresponder, a cada $g \in L^q(\mu)$, el elemento de $\big(L^p(\mu), \|\cdot\|_p\big)^*$ que lleva cada $f \in L^p(\mu)$ a $\int_\Omega f g\, d\mu$, es un isomorfismo isométrico.*

Demostración. La aplicación Φ está bien definida, es lineal y, gracias a la desigualdad de Hölder, satisface

$$\big|\Phi(g)(f)\big| = \left|\int_\Omega f g\, d\mu\right| \leq \|f\|_p \|g\|_q.$$

Por tanto, Φ es continua y $\|\Phi(g)\| \leq \|g\|_q$. Veamos a continuación que la desigualdad anterior es, de hecho, una igualdad. Si $g = 0$, la igualdad es obvia. Si $g \neq 0$, definimos la siguiente función medible:

$$h(x) = \begin{cases} \frac{g(x)}{|g(x)|} & \text{si } g(x) \neq 0, \\ 0 & \text{si } g(x) = 0. \end{cases}$$

Construimos ahora la función

$$f = \|g\|_q^{1-q}\, \overline{h}\, |g|^{q-1}.$$

La función f así definida es una función medible que, además, está en $L^p(\Omega, \Sigma, \mu)$ y $\|f\|_p \leq 1$. Se tiene que

$$\big\|\Phi(g)\big\| \geq \int_\Omega \|g\|_q^{1-q}\, \overline{h}\, |g|^{q-1} g\, d\mu = \|g\|_q.$$

Concluimos entonces que Φ es una isometría.

Veamos ahora que $L^p(\Omega, \Sigma, \mu)$ es reflexivo para $1 < p < \infty$. Obsérvese que, combinando la proposición 1.3.51 y el Teorema de Milman 1.3.52, se obtiene la reflexividad del espacio $L^p(\Omega, \Sigma, \mu)$ para $2 \leq p < \infty$. Por otro lado, si $1 < p < 2$, entonces $L^q(\Omega, \Sigma, \mu)$ es reflexivo y, en consecuencia, su dual $L^q(\Omega, \Sigma, \mu)^*$ también lo es. Así, el subespacio cerrado $L^p(\Omega, \Sigma, \mu) \subset L^q(\Omega, \Sigma, \mu)^*$ es reflexivo, después del corolario 1.3.47.

Nos resta ver que Φ es sobreyectiva. Será suficiente probar que el subespacio $\Phi\big(L^q(\Omega,\Sigma,\mu)\big)$ es denso en $\big(L^p(\Omega,\Sigma,\mu),\|\cdot\|_p\big)^*$. Para ello, utilizaremos el corolario 1.3.31. Si tomamos un elemento $f \in \big(L^p(\Omega,\Sigma,\mu),\|\cdot\|_p\big)^{**} = \big(L^p(\Omega,\Sigma,\mu),\|\cdot\|_p\big)$ y suponemos que $f|_{\Phi(L^q(\Omega,\Sigma,\mu))} = 0$, *i.e.*,

$$\Phi(g)(f) = \int_\Omega fg\,d\mu = 0 \text{ para todo } g \in L^q(\Omega,\Sigma,\mu), \tag{1.11}$$

hemos de ver que $f = 0$. Pero como

$$\|f\|_p = \sup\left\{\left|\int_\Omega fg\,d\mu\right| : \|g\|_q \leq 1,\ g \in L^q\right\},$$

de la igualdad (1.11) se concluye que $\|f\|_p = 0$, y así termina la prueba. □

1.3.7 El teorema de completitud de Grothendieck

$\boxed{\text{Compleción de un espacio normado}}$ Si $\big(X,\|\cdot\|\big)$ es un espacio normado, un *modelo para su compleción* puede construirse de la siguiente forma: tómese el bidual X^{**} de X, que es un espacio de Banach, y considérese la inmersión $\widehat{\ }: X \longrightarrow X^{**}$. El cierre de X en $\big(X^{**},\|\cdot\|\big)$ es un modelo para la compleción de X.

En esta sección nos encargaremos, entre otras cosas, de describir de forma explícita cómo se calcula el cierre de X en X^{**}.

Lema 1.3.54 (Lema de aproximación)**.** *Sean $E[\mathfrak{T}]$ un* e.l.c.*, A un subconjunto absolutamente convexo $\mathfrak{T}$-cerrado de E y $x^* : E \longrightarrow \mathbb{K}$ una forma lineal. Entonces, las siguientes afirmaciones son equivalentes:*

(i) $x^*|_A$ es $\mathfrak{T}$-continua.

(ii) Para cada $\varepsilon > 0$, existe $x' \in E'$ tal que $|x'(x) - x^*(x)| < \varepsilon$, para todo $x \in A$.

Demostración. La implicación *(ii)* $\Rightarrow$ *(i)* es consecuencia de que el límite uniforme de redes de funciones continuas es continuo.

Probemos la implicación *(i)* $\Rightarrow$ *(ii)*. Como $x^*|_A$ es continua, dado $\varepsilon > 0$, existe un $\mathfrak{T}$-entorno del origen absolutamente convexo y cerrado U tal que $|x^*(x)| < \varepsilon$, para cada $x \in U \cap A$. Esta condición significa que $x^* \in \varepsilon(U\cap A)^\circ$, donde la polar está calculada respecto al par dual $\langle E, E^\#\rangle$. Como U y A son $\sigma(E,E^\#)$-cerrados, el corolario 1.3.40 nos asegura que $(U\cap A)^\circ \subset \overline{U^\circ + A^\circ}^{\sigma(E^\#,E)}$. Por otro lado, U° es $\sigma(E^\#,E)$-compacto, después del teorema de Alaoglu-Bourbaki 1.3.41, y por lo tanto, tenemos que $U^\circ + A^\circ$ es $\sigma(E^\#,E)$-cerrado; en consecuencia, $\varepsilon(U\cap A)^\circ \subset \varepsilon(U^\circ + A^\circ)$. Entonces, $x^* \in \varepsilon U^\circ + \varepsilon A^\circ$ y, dado que $U^\circ \subset E'$, podemos concluir la existencia de $x' \in E'$ tal que $x^* - x' \in \varepsilon A^\circ$, lo que significa que

$$\big|x^*(x) - x'(x)\big| < \varepsilon, \text{ para cada } x \in A,$$

quedando así terminada la demostración. □

Teorema 1.3.55 (Teorema de completitud de Grothendieck, 1950)**.** *Sean $E[\mathfrak{T}]$ un* e.l.c. *y $\mathscr{E}$ la familia de los $\mathfrak{T}$-equicontinuos de E'. Sea*

$$\widehat{E} = \left\{x^* \in (E')^\# : x^*|_M \text{ es } \sigma(E',E)\text{-continuo, para cada } M \in \mathscr{E}\right\}.$$

El espacio $\widehat{E}$ dotado de la topología $\widehat{\mathfrak{T}}$ de convergencia uniforme sobre $\mathscr{E}$, es un modelo para la complección de $E[\mathfrak{T}]$, i.e.*, $E[\mathfrak{T}]$ es un subespacio denso de $\widehat{E}[\widehat{\mathfrak{T}}]$ y $\widehat{E}[\widehat{\mathfrak{T}}]$ es completo.*

Demostración. Si consideramos la inclusión natural $\widehat{\ }: E \longrightarrow \widehat{E}$ que hace corresponder, a cada x de E, la aplicación $\widehat{x} : E' \longrightarrow \mathbb{K}$ dada por $\widehat{x}(x') := x'(x)$, es claro que la topología inducida por $\widehat{E}[\widehat{\mathfrak{T}}]$ en E es $\mathfrak{T}$, después del corolario 1.3.39. La densidad de E en $\widehat{E}[\widehat{\mathfrak{T}}]$ es el lema 1.3.54, aplicado a $E'[\sigma(E',E)]$ y a cada elemento $x^* \in \widehat{E}$ restringido a equicontinuos absolutamente convexos $\sigma(E',E)$-cerrados. La completitud de $\widehat{E}[\widehat{\mathfrak{T}}]$ es, simplemente, el hecho de que los límites de redes de formas lineales son lineales, junto con el hecho de que el límite uniforme de redes de funciones continuas es continuo. □

Corolario 1.3.56. *Para un* e.l.c. *$E[\mathfrak{T}]$, las siguientes afirmaciones son equivalentes:*

(i) $E[\mathfrak{T}]$ es completo.
(ii) Toda forma lineal $x^* : E' \longrightarrow \mathbb{K}$ que es $\sigma(E',E)$-continua sobre los equicontinuos de E', es $\sigma(E',E)$-continua en E', es decir, es un elemento de E.

Corolario 1.3.57. *Para un espacio normado $\left(X,\|\cdot\|\right)$, son equivalentes:*

(i) $\left(X,\|\cdot\|\right)$ es un espacio de Banach.
(ii) Toda forma lineal $x^* : X^* \longrightarrow \mathbb{K}$ que es $\sigma(X^*,X)$-continua en B_{X^*}, es $\sigma(X^*,X)$-continua en X^*, i.e.*, es un elemento de X.*

Corolario 1.3.58. *Sean $\left(X,\|\cdot\|\right)$ un espacio normado, y $K = \left(B_{X^*},\sigma(X^*,X)\right)$, y sea*

$$i : \left(X,\|\cdot\|\right) \longrightarrow \left(C(K),\|\cdot\|_\infty\right)$$

la aplicación que hace corresponder, a cada $x \in X$, la función $\widehat{x} : K \longrightarrow \mathbb{K}$ dada por $\widehat{x}(x^*) = x^*(x)$. Entonces:

(i) i es un isomorfismo isométrico sobre su imagen que permite identificar el espacio $\left(X,\|\cdot\|\right)$ con un subespacio de $C(K)$. i es un isomorfismo de $\left(X,\sigma(X,X^*)\right)$ sobre su imagen en $\left(C(K),\tau_p(K)\right)$.
(ii) Si $\left(X,\|\cdot\|\right)$ es además completo, entonces $\left(X,\sigma(X,X^*)\right)$ se identifica con un subespacio cerrado de $\left(C(K),\tau_p(K)\right)$.

Demostración. La primera parte en la propiedad *(i)* es consecuencia del teorema de Hahn-Banach; la segunda parte se sigue de las definiciones de topología débil y topología de convergencia puntual. La propiedad *(ii)* se obtiene inmediatamente del corolario anterior. □

Alexander Grothendieck, 1928. A. Grothendieck fue uno de los miembros de Bourbaki que trabajó en numerosos volúmenes en los cuales se recogieron, de forma enciclopédica, muchas ramas de las matemáticas durante varias décadas. Grothendieck empezó su carrera en matemáticas dedicado al Análisis Funcional, donde influyó de forma decisiva en cuestiones concernientes a: bases, teoremas de gráfica cerrada, dualidad, caracterizaciones de compacidad débil, etc. Su dedicación al análisis ocupó, sin embargo, sólo una pequeña parte de su carrera científica, pues muy pronto cambió sus intereses hacia otras ramas de las matemáticas donde su influencia todavía sobrevive: cohomología, funtores derivados, teoría de categorías, geometría algebraica, etc. En 1966 recibió la medalla Field y, según se cuenta, hacia esa época, Grothendieck dio grandes muestras de generosidad y dedicación, al compartir numerosas ideas con estudiantes y colaboradores que eran cuidadosamente desarrolladas y presentadas posteriormente.

PARA SABER MÁS

- **Topología**. Los libros de topología [48] y [69] son nuestras referencias básicas. En el libro [69] encontramos una exposición detallada de las nociones de red, subred, etc., y de cómo éstas son utilizadas para caracterizar la compacidad, etc. El libro [48] abarca más cuestiones que el anterior, y además, contiene unas excelentes notas históricas que ayudan a ubicar en contexto muchos de los resultados expuestos. Ambos libros cuentan con una muy buena colección de ejercicios y problemas. Junto a las referencias anteriores, el libro [32] es de lectura casi imprescindible para saborear técnicas de topología aplicadas al estudio del Análisis Funcional.
- **Análisis Funcional**. El texto [23] es un curso introductorio al Análisis Funcional que cubre parte de la teoría de espacios de Hilbert, teoría espectral para operadores compactos autoadjuntos y espacios de Banach, prestando atención a los teoremas de Hahn-Banach, Gráfica Cerrada y Acotación Uniforme. El texto se complementa con diversas aplicaciones, así como con una pequeña colección de ejercicios propuestos y resueltos. El libro [50] puede servir como referencia actual, y muy completa, sobre cuestiones clásicas o recientes en Análisis Funcional y, especialmente, en la teoría de los espacios de Banach. Una referencia por excelencia para cuestiones de dualidad es [71]. El libro [95] es otra referencia clásica de Análisis Funcional con incursiones en las álgebras de Banach, distribuciones, etc. Los libros [108, 109] son referencias perfectas para encontrar aplicaciones de las técnicas de Análisis Funcional a parcelas como optimización o ecuaciones en derivadas parciales.

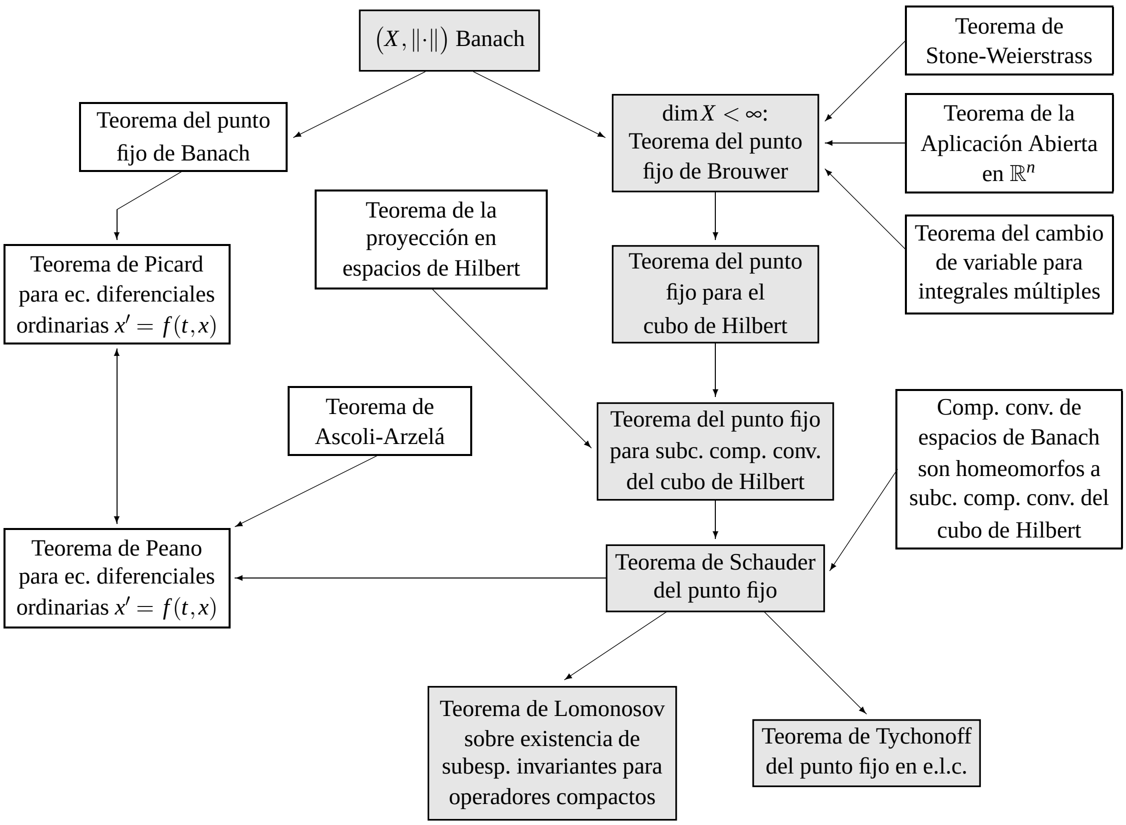
Capítulo 2

El teorema del punto fijo

«OBJETIVOS»

- Estudiar el teorema de Stone-Weierstrass sobre aproximación de funciones continuas y obtener, en particular, aproximaciones uniformes de funciones continuas en subconjuntos compactos de $\mathbb{R}^n$ por polinomios.
- Demostrar el teorema del punto fijo de Banach y recordar cómo se aplica al problema de Cauchy que asegura existencia y unicidad de determinadas ecuaciones diferenciales ordinarias –teorema de Picard-Lindelöf.
- Demostrar el teorema del punto fijo de Brouwer sobre existencia de puntos fijos para funciones continuas en subconjuntos compactos convexos de $\mathbb{R}^n$.
- Utilizar el teorema de Brouwer para demostrar versiones infinito-dimensionales de los teoremas del punto fijo: teorema de Schauder y teorema de Tychonoff. Deducir el teorema de Peano sobre existencia de soluciones de algunas ecuaciones diferenciales a partir del teorema de Schauder.
- Demostrar, como aplicación del teorema de Schauder, que todo operador acotado en un espacio de Banach que conmuta con un operador compacto no trivial tiene un subespacio invariante no trivial.

EL primer y más sencillo teorema del punto fijo es el teorema de Bolzano, 1817, de los valores intermedios de funciones continuas reales, véase la figura 2.4. Cauchy, en un artículo publicado en 1835, utilizó un *método de aproximaciones sucesivas* para dar un teorema de existencia para algunos tipos generales de ecuaciones diferenciales, para los que no se tenían soluciones explícitas. Picard, en 1890, utilizó métodos de aproximaciones sucesivas para garantizar la existencia de soluciones de algunas ecuaciones diferenciales con ciertas condiciones frontera. En 1922, Banach demostró, en su tesis doctoral, el teorema del punto fijo que hoy lleva su nombre, y que también se conoce como Principio de la Aplicación Contractiva. Algunos de los resultados sobre existencia de soluciones de ecuaciones diferenciales, soluciones de sistemas con infinitas ecuaciones e incógnitas, existencia de funciones implícitas, métodos numéricos, existencia de fractales, medidas invariantes, existencia de subespacios invariantes, etc., pueden ser obtenidos como consecuencia de teoremas del punto fijo. En este capítulo exponemos los teoremas del punto fijo de Banach 2.2.1, Brouwer 2.3.6, Schauder 2.4.9 y Tychonoff 2.4.15, y damos, como aplicaciones, los teoremas de Picard 2.2.5, Peano 2.4.10 y Lomonosov 2.5.8. La demostración que presentamos del teorema de Brouwer sigue a [93]. El teorema de Schauder se obtiene fácilmente del teorema de Brouwer pasando por el cubo de Hilbert. La prueba que damos del teorema de Tychonoff está tomada de [43], y la demostración del teorema de Lomonosov de [36]. En las notas históricas de la página 100 incluimos la prueba original de Lomonosov según I. Namioka.



Cuadro 2.1: Esquema del capítulo *El teorema del punto fijo*

2.1 El teorema de Stone-Weierstrass

El teorema de Stone-Weierstrass, teorema 2.1.4, es un resultado de aproximación en el que se establece que, para un espacio compacto K, toda subfamilia $\mathscr{A}$ de $C(K)$ que es suficientemente rica (separa y distingue puntos) y que es cerrada bajo ciertas operaciones (es un álgebra) puede ser usada para aproximar uniformemente cada función continua por miembros de $\mathscr{A}$. El teorema de Stone-Weierstrass juega un papel central en muchas cuestiones del Análisis, como por ejemplo, para concluir que determinadas sucesiones de polinomios ortogonales y el sistema trigonométrico son bases hilbertianas de $L^2([0,2\pi])$, [72, p. 194-195], para la obtención de fórmulas de cuadratura Gaussiana, [72, p. 76], para obtener teoremas de representación de álgebras abstractas, [95, Teorema de Gel'fand-Naimark, p. 285], etc.

$\boxed{\textit{Álgebra}}$ Un *álgebra* es un espacio vectorial X dotado de una tercera operación interna (llamada usualmente multiplicación) la cual asigna, a cada dos vectores $x,y \in X$, otro elemento $xy \in X$ que satisface las propiedades:

- $x(yz) = (xy)z$ (asociativa),
- $x(y+z) = xz + yz$, o $(x+y)z = xz + yz$ (distributiva),
- $\alpha(xy) = (\alpha x)y = x(\alpha y)$,

para cualesquiera x,y,z de X y para cada escalar α.
Si además X es un espacio de Banach respecto de una norma que satisface la desigualdad multiplicativa

- $\|xy\| \le \|x\|\,\|y\|$,

entonces X se llama *álgebra de Banach*. Si K es un espacio compacto, el espacio de funciones continuas $C(K)$ sobre K tiene estructura natural de álgebra cuando se dota de la multiplicación que consiste en multiplicar puntualmente las funciones. Además, $C(K)$ dotado de la norma usual

$$\|f\|_\infty := \sup\big\{|f(x)| : x \in K\big\},$$

es también un álgebra de Banach.

$\boxed{\textit{Retículo}}$ En $C(K,\mathbb{R})$ consideramos la relación de orden parcial definida por

$$f \le g \ \text{ si } f(x) \le g(x), \text{ para cada } x \in K,$$

$f,g \in C(K,\mathbb{R})$. Es bien conocido que si $f,g \in C(K,\mathbb{R})$, entonces las funciones $\text{máx}\{f,g\}$ y $\text{mín}\{f,g\}$ pertenecen también a $C(K,\mathbb{R})$, *i.e.*, $C(K,\mathbb{R})$ es un *retículo*. En general, una parte $\mathscr{A} \subset C(K,\mathbb{R})$ se dice *reticulada* si $\text{máx}\{f,g\}, \text{mín}\{f,g\} \in \mathscr{A}$, cada vez que $f,g \in \mathscr{A}$.

Lema 2.1.1. *Sean K un espacio compacto y $\mathscr{A} \subset C(K,\mathbb{R})$ una parte reticulada. Para una función $f \in C(K,\mathbb{R})$, las siguientes afirmaciones son equivalentes:*

(i) $f \in \overline{\mathscr{A}}^{\|\cdot\|_\infty}$.
(ii) Para cada $\varepsilon > 0$ y para cada $x,y \in K$, existe $g_{xy} \in \mathscr{A}$ tal que

$$|f(x) - g_{xy}(x)| < \varepsilon \quad y \quad |f(y) - g_{xy}(y)| < \varepsilon. \tag{2.1}$$

Demostración. La implicación *(i)* $\Rightarrow$ *(ii)* es evidente. Para probar *(ii)* $\Rightarrow$ *(i)* tenemos que demostrar que, para cada $\varepsilon > 0$, existe una función $g \in \mathscr{A}$ tal que $\|f-g\|_\infty < \varepsilon$. Si hemos fijado $\varepsilon > 0$, la condición *(ii)* nos garantiza, para cualesquiera $x, y \in K$, la existencia de $g_{xy} \in \mathscr{A}$ satisfaciendo (2.1). Dados $x, y \in K$, definimos

$$G_{xy} := \{z \in K : g_{xy}(z) < f(z) + \varepsilon\}.$$

Como las funciones involucradas son continuas G_{xy} es abierto. Para cada $x \in K$ fijo, la familia $\{G_{xy}\}_{y \in K}$ cubre K, gracias a que $y \in G_{xy}$ para cada $y \in K$. Entonces, por la compacidad de K, existen $y_1, y_2, \ldots, y_n \in K$ tales que $K = \bigcup_{i=1}^n G_{xy_i}$. Para cada $x \in K$, definimos ahora la función

$$g_x := \operatorname{mín}\{g_{xy_i} : i = 1, \ldots, n\}.$$

Por las hipótesis hechas sobre $\mathscr{A}$, se tiene que esta función $g_x \in \mathscr{A}$, verificándose además las siguientes desigualdades:

$$g_x(z) < f(z) + \varepsilon, \text{ para cada } z \in K,$$

y

$$g_x(x) > f(x) - \varepsilon.$$

Si escribimos

$$O_x := \{z \in K : g_x(z) > f(z) - \varepsilon\},$$

entonces O_x es abierto y $x \in O_x$. Así, $\{O_x\}_{x \in K}$ cubre K y, de nuevo por la compacidad de K, existen $x_1, x_2, \ldots, x_m \in K$ tales que $K = \bigcup_{i=1}^m O_{x_i}$. Definiendo ahora

$$g := \operatorname{máx}\{g_{x_i} : i = 1, \ldots, m\},$$

se tiene que $g \in \mathscr{A}$ y que, para cada $z \in K$, se satisfacen simultáneamente las desigualdades

$$g(z) < f(z) + \varepsilon \quad \text{y} \quad g(z) > f(z) - \varepsilon,$$

lo que significa que $\|f - g\|_\infty < \varepsilon$; esto concluye la prueba. $\square$

El lema anterior encierra ya una versión del teorema de Stone-Weierstrass.

Teorema 2.1.2 (Stone-Weierstrass (versión reticular)). *Sea $\mathscr{A}$ un subespacio vectorial reticulado de $C(K, \mathbb{R})$ satisfaciendo:*

(i) $\mathscr{A}$ contiene a la función $\mathbf{1}$, constantemente igual a 1 en K.
(ii) Para cada $x, y \in K$, existe $f \in \mathscr{A}$ tal que $f(x) \neq f(y)$ ($\mathscr{A}$ separa los puntos de K).

Entonces, $\mathscr{A}$ es denso en $(C(K, \mathbb{R}), \|\cdot\|_\infty)$.

Demostración. Es suficiente ver que se satisface la condición *(ii)* del lema 2.1.1 y, para ello, simplemente probamos que:

(a) Si $x = y \in K$ y $\alpha \in \mathbb{R}$, entonces existe $g \in \mathscr{A}$ tal que $g(x) = \alpha$.
(b) Si $x \neq y \in K$ y $\alpha, \beta \in \mathbb{R}$, entonces existe $g \in \mathscr{A}$ tal que $g(x) = \alpha$ y $g(y) = \beta$.

Obsérvese primero que, como las funciones constantes pertenecen a $\mathscr{A}$, la condición (a) es claramente satisfecha. Veamos (b). Dados $x, y \in K$ con $x \neq y$, existe, por hipótesis, $f \in \mathscr{A}$ tal que $f(x) \neq f(y)$. Para alguna combinación de la forma $g = \lambda f + \mu \mathbf{1}$, se tiene que $g(x) = \alpha$ y $g(y) = \beta$, dado que el sistema lineal

$$\left.\begin{array}{l} \lambda f(x) + \mu = \alpha \\ \lambda f(y) + \mu = \beta \end{array}\right\},$$

en las incógnitas λ y μ, tiene determinante no nulo. $\square$

El siguiente lema permite demostrar la versión del teorema de Stone-Weierstrass para álgebras.

Lema 2.1.3. *Si $\mathscr{A}$ es una subálgebra de $C(K,\mathbb{R})$, entonces su cierre, $\overline{\mathscr{A}}^{\|\cdot\|_\infty}$, es una subálgebra reticulada de $C(K,\mathbb{R})$.*

Demostración. Cálculos elementales conducen a que $\overline{\mathscr{A}}^{\|\cdot\|_\infty}$ es, de nuevo, un álgebra. Así, el lema estará demostrado si vemos que toda subálgebra cerrada de $C(K,\mathbb{R})$ es reticulada.

Supongamos pues que $\mathscr{A}$ es cerrada en $\big(C(K,\mathbb{R}), \|\cdot\|_\infty\big)$, y probemos que si $f, g \in \mathscr{A}$, entonces $\max\{f,g\}, \min\{f,g\} \in \mathscr{A}$. Obsérvese que, para dos funciones $f, g \in C(K,\mathbb{R})$, se tiene que

$$\max\{f,g\} = \frac{1}{2}\Big(f + g + |f - g|\Big) \quad \text{y} \quad \min\{f,g\} = \frac{1}{2}\Big(f + g - |f - g|\Big).$$

Por tanto, para demostrar que $\mathscr{A}$ es reticulada es suficiente probar que si $f \in \mathscr{A}$, entonces $|f| \in \mathscr{A}$. Veamos esto último. Sea $f \in \mathscr{A}$, tomemos $M > |f(x)|$, para cada $x \in K$, y consideremos la función $g := (1/M)f$. Claramente, $g(x) \in [-1,1]$, para cada $x \in K$; si demostramos que $|g| \in \mathscr{A}$, nuestra prueba habrá terminado. Para demostrar esto último, tengamos presente la igualdad

$$|y| = \big(1 + (y^2 - 1)\big)^{1/2} = \sum_{n=0}^{\infty} \binom{1/2}{n} (y^2 - 1)^n = \lim_m \sum_{n=0}^{m} \binom{1/2}{n} (y^2 - 1)^n =: \lim_m P_m(y),$$

con convergencia uniforme en $|y| \leq \sqrt{2}$. Obsérvese que cada P_m es un polinomio, y si definimos $Q_m = P_m - P_m(0)$, entonces cada Q_m es un polinomio sin término independiente, verificándose de nuevo que

$$|y| = \lim_m Q_m(y),$$

con convergencia uniforme en $|y| \leq \sqrt{2}$. En particular, tenemos que

$$\big|g(x)\big| = \lim_m Q_m\big(g(x)\big)$$

uniformemente en $x \in K$. Esto significa que la sucesión $\big(Q_m(g)\big)_m$, que está en $\mathscr{A}$, converge hacia g en $\|\cdot\|_\infty$, y de esta manera hemos demostrado que $g \in \mathscr{A}$, lo que termina la prueba. $\square$

Teorema 2.1.4 (Stone-Weierstrass (versión para álgebras reales))**.** *Sea $\mathscr{A}$ una subálgebra de $C(K,\mathbb{R})$ que satisface las siguientes condiciones:*

(i) Para cada $x \in K$, existe $f \in \mathscr{A}$ tal que $f(x) \neq 0$ ($\mathscr{A}$ distingue los puntos de K).

(ii) Para cada $x,y \in K$, existe $f \in \mathscr{A}$ tal que $f(x) \neq f(y)$ ($\mathscr{A}$ separa los puntos de K).

Entonces, $\mathscr{A}$ es densa en $\big(C(K,\mathbb{R}),\|\cdot\|_\infty\big)$.

Demostración. La demostración es análoga a la del teorema 2.1.2, por lo que simplemente demostraremos que se satisfacen las condiciones (a) y (b) de dicha prueba. La condición de que $\mathscr{A}$ distingue los puntos de K, junto con el hecho de que $\mathscr{A}$ es un álgebra, implican que (a) se satisface. Veamos (b). Dados $x,y \in K$ con $x \neq y$, existe, por hipótesis, $f \in \mathscr{A}$ tal que $f(x) \neq f(y)$. Distinguimos dos casos:

CASO 1: $f(x) = 0$ ó $f(y) = 0$.

Supongamos que $f(x) = 0$ y tomemos, utilizando que $\mathscr{A}$ distingue los puntos de K, una función $h \in \mathscr{A}$ tal que $h(x) \neq 0$. Entonces, para $\alpha,\beta \in \mathbb{R}$, una combinación de la forma $g = \lambda f + \mu h$ satisface $g(x) = \alpha$ y $g(y) = \beta$, dado que el sistema lineal

$$\left.\begin{aligned} \lambda f(x) + \mu h(x) &= \alpha \\ \lambda f(y) + \mu h(y) &= \beta \end{aligned}\right\},$$

en las incógnitas λ y μ, tiene determinante no nulo.

CASO 2**:** $f(x) \neq 0$ y $f(y) \neq 0$.

Dados $\alpha,\beta \in \mathbb{R}$, una combinación de la forma $g = \lambda f + \mu f^2$ satisface $g(x) = \alpha$ y $g(y) = \beta$, dado que el sistema lineal

$$\left.\begin{aligned} \lambda f(x) + \mu f^2(x) &= \alpha \\ \lambda f(y) + \mu f^2(y) &= \beta \end{aligned}\right\},$$

en las incógnitas λ y μ, tiene determinante no nulo. □

Corolario 2.1.5. *Sean K un subconjunto compacto de $\mathbb{R}^n$ y $f \in C(K,\mathbb{R})$. Entonces, para cada $\varepsilon > 0$, existe un polinomio P en n variables tal que $\|f - P\|_\infty < \varepsilon$.*

Demostración. Basta utilizar que la familia $\mathscr{A}$ de los polinomios en n variables restringidos a K satisface las hipótesis del teorema 2.1.4. □

En particular, si $K = [a,b]$ es un intervalo de la recta real, los polinomios en una variable restringidos a K son una subálgebra densa de $C\big([a,b],\mathbb{R}\big)$ (éste fue el resultado original demostrado por Weierstrass, 1885). Comentamos que, de forma alternativa, se puede probar este teorema de Weierstrass utilizando el teorema de Korovkin que enunciamos a continuación, cuya demostración puede encontrarse en [72, p. 19-21] y [23, p. 61-63].

Teorema 2.1.6 (Korovkin, 1953)**.** *Consideremos las funciones f_0, f_1 y f_2 definidas en $[a,b]$ por*

$$f_0(x) = 1, \quad f_1(x) = x \quad y \quad f_2(x) = x^2,$$

donde $x \in [a,b]$. Para cada $n = 1,2,\ldots$, sea $B_n : C\big([a,b]\big) \longrightarrow C\big([a,b]\big)$ lineal, y supongamos que:

(i) Cada B_n es positivo, i.e.*, $B_n(f) \geq 0$, para $n = 1,2,\ldots$ y cada $f \geq 0$ en $C\big([a,b]\big)$.*

(ii) Para $m = 0,1,2$, se tiene que $\lim_{n\to\infty}\big\|B_n(f_m) - f_m\big\|_\infty = 0$.

Entonces,

$$\lim_{n\to\infty}\big\|B_n(f) - f\big\|_\infty = 0 \tag{2.2}$$

para cada $f \in C\big([a,b]\big)$.

Dando por supuesto el teorema de Korovkin, la demostración del teorema de Weierstrass en $C\big([0,1],\mathbb{R}\big)$ se puede hacer considerando los operadores asociados a los polinomios de Bernstein, $B_n : C\big([0,1]\big) \longrightarrow C\big([0,1]\big)$, $n = 1,2,\ldots$, dados por

$$B_n(f)(x) = \sum_{k=0}^{n} f(k/n)\binom{n}{k}x^k(1-x)^{n-k}, \tag{2.3}$$

para $f \in C\big([0,1]\big)$ y $t \in [0,1]$. La sucesión $(B_n)_n$ satisface las hipótesis del teorema 2.1.6 debido a que cada B_n es lineal y positivo y se tiene que

$$\begin{aligned} B_n(f_0) &= f_0, \\ B_n(f_1) &= f_1, \\ B_n(f_2) &= (1-1/n)f_2 + (1/n)f_1. \end{aligned}$$

Por tanto obtenemos que, para cada función $f \in C\big([0,1],\mathbb{R}\big)$, la sucesión de sus polinomios de Bernstein asociada, $\big(B_n(f)\big)_n$, converge uniformemente hacia f en $[0,1]$. Es claro que, establecido el teorema de Weierstrass en $[0,1]$, éste queda demostrado en cualquier intervalo $[a,b]$.

Las figuras 2.1, 2.2 y 2.3 ilustran resultados obtenidos con el ordenador en distintos problemas concretos de aproximación. El primero de los gráficos corresponde a la convergencia de los polinomios de Bernstein. El segundo, a la aproximación proporcionada por polinomios de interpolación utilizando distintos nodos. El tercero corresponde a la aproximación mín − máx de grado 1 para una función dada: el teorema de Čebyšev, [60, p. 147-149], nos asegura que, para cada función continua $f \in C\big([a,b]\big)$ y cada $n \in \mathbb{N}$, existe un único polinomio p_n de grado, a lo más n, tal que

$$\|f - p_n\| = \text{mín}\{\|f-q\| : q \text{ polinomio, a lo más, de grado } n\}.$$

El algoritmo de Remes nos proporciona un método para calcular p_n (p_n es el único polinomio de grado menor o igual que n que equioscila respecto de f en $n+2$ puntos distintos de $[a,b]$), véase [60, p. 144-157]. Conociendo a priori la validez del teorema de Stone-Weierstrass, la sucesión de polinomios $(p_n)_n$ así obtenida converge uniformemente hacia f.

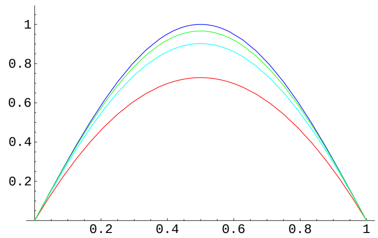


Figura 2.1: Polinomios de Bernstein $\sum_{k=0}^{n} \operatorname{sen}\left(\pi \frac{k}{n}\right)\binom{n}{k} t^k (1-t)^{n-k}$ de grados 4, 12 y 36, para la función $\operatorname{sen}(\pi x)$ en $[0,1]$

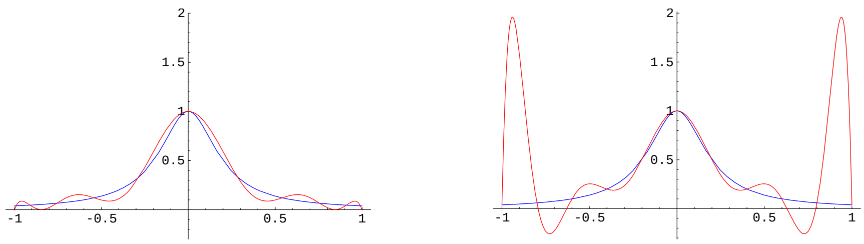


Figura 2.2: Polinomios de interpolación sobre nodos de Čebyšev y efecto Runge para la función $\dfrac{1}{1+25x^2}$: 10 nodos

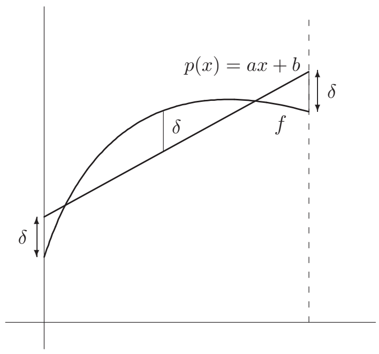


Figura 2.3: Mejor aproximación sobre espacios de polinomios: Čebyšev

Acabamos esta sección con la versión compleja del teorema de Stone-Weierstrass.

Teorema 2.1.7 (Stone-Weierstrass (versión para álgebras complejas))**.** *Sea $\mathscr{A}$ una subálgebra de $C(K,\mathbb{C})$ que satisface las siguientes condiciones:*

(i) Para cada $x \in K$, existe $f \in \mathscr{A}$ tal que $f(x) \neq 0$ ($\mathscr{A}$ distingue los puntos de K).

(ii) Para cada $x, y \in K$, existe $f \in \mathscr{A}$ tal que $f(x) \neq f(y)$ ($\mathscr{A}$ separa los puntos de K).

(iii) $\mathscr{A}$ es auto-conjugada, i.e., *para cada $f \in \mathscr{A}$, la función conjugada $\overline{f} \in \mathscr{A}$.*

Entonces, $\mathscr{A}$ es densa en $\big(C(K,\mathbb{C}), \|\cdot\|_\infty\big)$.

Demostración. La subálgebra (sobre $\mathbb{R}$)

$$\mathscr{A}_r := \big\{f \in \mathscr{A} : f(K) \subset \mathbb{R}\big\}$$

separa y distingue puntos de K. Así, el teorema 2.1.4 nos permite concluir que $\overline{\mathscr{A}_r}^{\|\cdot\|_\infty} = C(K,\mathbb{R})$, lo cual implica claramente que

$$C(K,\mathbb{C}) = \overline{\mathscr{A}_r + i\mathscr{A}_r}^{\|\cdot\|_\infty} \subset \overline{\mathscr{A}}^{\|\cdot\|_\infty} \subset C(K,\mathbb{C}),$$

y la prueba acaba. □

Debemos hacer notar que la condición *(iii)* sobre la auto-conjugación de $\mathscr{A}$ en el caso complejo es necesaria, tal y como pone de manifiesto el siguiente sencillo ejemplo: sea $K := \big\{z \in \mathbb{C} : |z| \leq 1\big\}$, y sea $\mathscr{A}$ el conjunto de los polinomios complejos (en la variable z) restringidos a K. Es claro que $\mathscr{A}$ es un álgebra que satisface *(i)* y *(ii)* en el teorema 2.1.7, pero que no satisface *(iii)*. Por otro lado, $\mathscr{A}$ no es densa en $\big(C(K,\mathbb{C}), \|\cdot\|_\infty\big)$, dado que existen funciones continuas en K que no son holomorfas en su interior.

El teorema de Stone-Weierstrass, 1885-1937. La noción de convergencia de una sucesión de funciones numéricas se venía empleando de forma más o menos consciente desde los comienzos del Cálculo Infinitesimal, aunque hubo que esperar hasta B. Bolzano (1781-1848) y A. L. Cauchy (1789-1857) para que los conceptos de función continua y serie convergente fuesen definidos de forma precisa. Incluso para Cauchy, la diferencia entre convergencia *puntual* y *uniforme* de una serie de funciones no estaba totalmente clara, pues creyó que podía demostrar que toda serie convergente de funciones continuas tenía suma continua. El estudio de la noción de convergencia uniforme se desarrolló durante el último tercio del siglo XIX bajo la influencia de K. Weierstrass y B. Riemann; nombres involucrados en este estudio son los de U. Dini y C. Arzelá: el primero precisó las condiciones para que el límite de una sucesión de funciones continuas fuese continuo, mientras que el segundo introdujo la noción de *equicontinuidad*, a través de la cual se caracterizan los subconjuntos compactos de espacios de funciones continuas; éstos desempeñan un papel fundamental, por ejemplo, a la hora de demostrar el teorema de Montel concerniente a familias normales de funciones holomorfas. En 1885, Weierstrass descubrió la posibilidad de aproximar uniformemente, mediante polinomios, una función real continua de una o varias variables reales, en un conjunto compacto. La contribución moderna dentro de estas cuestiones ha consistido, fundamentalmente, en darles todo el alcance del que son susceptibles, abordándolas para funciones cuyos dominios no se restringen a espacios de dimensión finita. En 1937, M. H. Stone generalizó el teorema de Weierstrass y simplificó su prueba, hasta darle el alcance y aspecto con el que se presenta actualmente en los libros de texto (véase [11, p. 282-283] para estos comentarios históricos).

2.2 El teorema del punto fijo de Banach

EL teorema del punto fijo, conocido generalmente como el *Principio de Contracción de Banach*, apareció de forma explícita en la tesis de Banach en 1922, donde se utilizaba para establecer la existencia de una solución en una ecuación integral. Desde entonces, debido a su simplicidad y utilidad, se ha convertido en una herramienta muy común para la resolución de problemas de existencia en muchas ramas del Análisis Matemático. En este capítulo demostramos el principio de contracción de Banach, discutimos algunas de sus variantes más útiles, y presentamos unos pocos y diversos ejemplos de sus aplicaciones.

Sea M una espacio métrico con función distancia (métrica) d. Una aplicación $T : M \longrightarrow M$ se dice que es *lipschitziana* si existe un $k \geq 0$ tal que, para todos $x, y \in M$,

$$d(Tx, Ty) \leq kd(x, y). \tag{2.4}$$

El menor número k para el cual se satisface (2.4), se denomina la *constante de Lipschitz* de T. Con frecuencia, representaremos por $k(T)$ y $k(S)$ las respectivas constantes de Lipschitz de diferentes aplicaciones T y S; y cuando sea relevante, $K_d(T)$ se utilizará para denotar la constante de Lipschitz de T respecto a la métrica d.

Para dos aplicaciones Lipschitzianas $S, T : M \longrightarrow M$, se verifica que

$$k(T \circ S) \leq k(T)k(S)$$

y, en particular,

$$k(T^n) \leq k(T)^n, \qquad n = 1, 2, \ldots$$

Además, si M es un espacio lineal cuya métrica está generada por una norma, entonces se tiene que $k(T + S) \leq k(T) + k(S)$ y, para $\alpha \geq 0$, $k(\alpha T) = \alpha k(T)$.

Una aplicación $T : M \longrightarrow M$ se dice que es una *contracción* si $k(T) < 1$; de manera más precisa, T es una k-contracción con respecto a d si $k_d(T) \leq k < 1$.

Teorema 2.2.1 (Principio de Contracción de Banach)**.** *Sean (M,d) un espacio métrico completo y $T : M \longrightarrow M$ una contracción. Entonces, T tiene un único punto fijo en M, y para cada $x_0 \in M$, la sucesión de iteradas $(T^n x_0)_n$ converge a dicho punto fijo.*

Vamos a dar dos demostraciones del teorema 2.2.1. La primera, que es una variación de la prueba original, no sólo establece la existencia de un punto fijo, sino que también, al igual que en la demostración de Banach, proporciona un método para su aproximación. Finalmente, damos la prueba original (y más comúnmente conocida).

Demostración 1. Definimos la función $\varphi(x) := (1-k)^{-1} d(x,Tx)$, para $x \in M$ (donde $k = k_d(T)$). Entonces,

$$d(x,Tx) - kd(x,Tx) \leq d(x,Tx) - d(Tx,T^2x)$$

y, en consecuencia,

$$d(x,Tx) \leq \varphi(x) - \varphi(Tx), \qquad x \in M. \tag{2.5}$$

Así, para $x_0 \in M$ y $n,m \in \mathbb{N}$ con $n < m$, se tiene que

$$d(T^n x_0, T^{m+1} x_0) \leq \sum_{i=n}^{m} d(T^i x_0, T^{i+1} x_0) \leq \varphi(T^n x_0) - \varphi(T^{m+1} x_0). \tag{2.6}$$

En particular,

$$\sum_{i=0}^{\infty} d(T^i x_0, T^{i+1} x_0) < \infty.$$

Por tanto, $(T^n x_0)_n$ es una sucesión de Cauchy y, debido a que T es continua, converge a un punto fijo x de T. La velocidad de esta convergencia puede obtenerse de (2.6) haciendo tender $m \to \infty$:

$$d(T^n x_0, x) \leq \varphi(T^n x_0) = (1-k)^{-1} d(T^n x_0, T^{n+1} x_0) \leq \frac{k^n}{1-k} d(x_0, Tx_0). \qquad \square$$

Observación 2.2.2. La prueba anterior muestra que *cualquier* aplicación continua que satisface (2.5), para una aplicación $\varphi : M \longrightarrow \mathbb{R}^+$ arbitraria, debe tener un punto fijo. De hecho, puede demostrarse por otros métodos que, si φ es inferiormente semicontinua, entonces una aplicación $T : M \longrightarrow M$ *arbitraria* verificando (2.5) debe tener un punto fijo. Este hecho, conocido comúnmente como el teorema de Caristi, se presentará con detalle más adelante. Éste es equivalente al Principio de Minimización de Ekeland (suponiendo el axioma de elección), [44], y tiene numerosas aplicaciones en Análisis (véase, por ejemplo, [16] para una discusión minuciosa). El punto fijo, en los dos casos anteriores, no tiene por qué ser necesariamente único, e incluso, en el segundo de ellos, la sucesión $(T^n x_0)_n$ puede no converger a un punto fijo de T.

Demostración 2. Seleccionamos un punto $x_0 \in M$, y definimos la sucesión iterativa $(x_n)_n$ dada por $x_{n+1} = Tx_n$ (o, equivalentemente, $x_n = T^n x_0$), para $n = 0,1,2,\ldots$ Obsérvese que, para índices cualesquiera $n, p \in \mathbb{N}$,

$$\begin{aligned} d(x_n, x_{n+p}) &= d(T^n x_0, T^{n+p} x_0) = d(T^n x_0, T^n \circ T^p x_0) \leq k(T^n) d(x_0, T^p x_0) \\ &\leq k^n \left[d(x_0, Tx_0) + d(Tx_0, T^2 x_0) + \ldots + d(T^{p-1} x_0, T^p x_0) \right] \\ &\leq k^n (1 + k + \ldots + k^{p-1}) d(x_0, Tx_0) \\ &\leq k^n \left(\frac{1-k^p}{1-k} \right) d(x_0, Tx_0). \end{aligned} \tag{2.7}$$

Esto demuestra que $(x_n)_n$ es una sucesión de Cauchy y, debido a que M es completa, existe un $x \in M$ tal que $\lim_{n\to\infty} x_n = x$. Para ver que x es el único punto fijo de T, obsérvese que

$$x = \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} Tx_n = Tx;$$

y, más aún, $x = Tx$ e $y = Ty$ implican

$$d(x,y) = d(Tx, Ty) \leq k d(x,y),$$

de donde se deduce que $d(x,y) = 0$.

Como en la primera demostración, haciendo tender $p \to \infty$ en (2.7), se tiene finalmente que

$$d(x_n, x) = d(T^n x_0, x) \leq \frac{k^n}{1-k} d(x_0, Tx_0). \qquad \square$$

La simplicidad y utilidad del principio de contracción de Banach ha inspirado a muchos autores para analizarlo e intentar llegar más lejos. Estos estudios han conducido a una serie de generalizaciones y modificaciones de dicho principio, siendo una de las más profundas aquélla a la que aludíamos en la observación 2.2.2.

Teorema 2.2.3 (Caristi)**.** *Sean (M,d) un espacio métrico completo y $\varphi : M \longrightarrow \mathbb{R}$ una función inferiormente semicontinua que está acotada inferiormente. Supongamos que $T : M \longrightarrow M$ es una aplicación arbitraria que satisface*

$$d(x, Tx) \leq \varphi(x) - \varphi(Tx), \quad x \in M.$$

Entonces T tiene un punto fijo.

La relación del resultado anterior con el principio de contracción de Banach ha sido remarcada en la observación 2.2.2 anterior. Sin embargo, el teorema de Caristi es, realmente, sólo un pariente lejano del principio de Banach. La demostración original de Caristi, [18], y el refinamiento de Wong (1976) de ésta, utilizan inducción transfinita. Se conocen varias demostraciones elegantes que utilizan el lema de Zorn (véanse, por ejemplo, [70] o [83]), así como pruebas constructivas algo intrincadas (por ejemplo, [100]).

Aquí, deduciremos el teorema de Caristi a partir de la siguiente reformulación de un resultado de Brézis y Browder, [16]. En este teorema, X representa un conjunto (preordenado) parcialmente ordenado y, para cada $x \in X$, $S(x) = \{y \in X : y \geq x\}$. Una sucesión $(x_n)_n$ en X se dice que es creciente si $x_n \leq x_{n+1}$ para todo n.

Teorema 2.2.4. *Sea $\psi : X \longrightarrow \mathbb{R}$ una función que satisface las siguientes condiciones:*

(i) $x \leq y$, *y* $x \neq y$ *implica* $\psi(x) < \psi(y)$.

(ii) Para cualquier sucesión creciente $(x_n)_n$ en X tal que $\psi(x_n) \leq C < \infty$, para todo $n \in \mathbb{N}$, existe algún $y \in X$ de forma que $x_n \leq y$ para todo n.

(iii) Para cada $x \in X$, $\psi\big(S(x)\big)$ está acotada superiormente.

Entonces, para cada $x \in X$, existe $x' \in S(x)$ tal que x' es maximal, i.e., $\{x'\} = S(x')$.

Demostración. Para $a \in X$, sea $\rho(a) := \sup\big\{\psi(b) : b \in S(a)\big\}$. Supongamos que la tesis del teorema falla para algún $x \in X$, y definamos, por inducción, una sucesión $(x_n)_n$ tal que $x_1 = x$ y $x_{n+1} \in S(x_n)$ satisface $\rho(x_n) \leq \psi(x_{n+1}) + (1/n)$, para todo $n \in \mathbb{N}$. Dado que $\psi(x_{n+1}) \leq \rho(x) < \infty$, se deduce de *(ii)* que existe algún $y \in X$ tal que $x_n \leq y$ para todo n. Por hipótesis, y no es maximal en $S(x)$, y por tanto, existe $u \in X$ tal que $y \leq u$ y $\psi(y) < \psi(u)$. Ya que $x_n \leq u$, $\psi(u) \leq \rho(x_n)$ para todo n. Además, $x_{n+1} \leq y$, y así, $\psi(x_{n+1}) \leq \psi(y)$. En consecuencia,

$$\psi(u) \leq \rho(x_n) \leq \psi(x_{n+1}) + (1/n) \leq \psi(y) + (1/n)$$

para todo n, de donde se deduce que $\psi(u) \leq \psi(y)$, lo cual es una contradicción. □

Demostración del teorema de Caristi. Sea $\psi = -\varphi$. Dados $x, y \in M$, decimos que

$$x \leq y \quad \text{si} \quad d(x,y) \leq \varphi(x) - \varphi(y).$$

Obsérvese que, por hipótesis, $x \leq Tx$ para todo $x \in M$. Debemos verificar las condiciones *(i)*, *(ii)* y *(iii)* del teorema 2.2.4. *(i)* es evidente y, para ver que *(ii)* se satisface, basta darse cuenta de que si $(x_n)_n$ es una sucesión creciente cualquiera, entonces $\big(\varphi(x_n)\big)_n$ es decreciente y acotada inferiormente; por tanto, $\big(\varphi(x_n)\big)_n$ converge, digamos a un $r \in \mathbb{R}$, lo cual implica que $(x_n)_n$ es una sucesión de Cauchy. En consecuencia, $(x_n)_n$ converge a un punto $y \in M$, y dado que φ es inferiormente semicontinua, se deduce que

$$d(x_n, y) \leq \varphi(x_n) - r \leq \varphi(x_n) - \varphi(x).$$

Por lo tanto, $x_n \leq y$ para todo $n \in \mathbb{N}$, lo que demuestra que *(ii)* se satisface. Ya que *(iii)* se sigue del hecho de que φ está acotada inferiormente, podemos concluir que, para cada $x \in M$, existe un $x' \geq x$ tal que $Tx' = x$. □

Como aplicación del teorema del punto fijo de Banach incluimos el clásico problema de Cauchy sobre la existencia y unicidad de la solución de una ecuación diferencial verificando una condición inicial dada.

Ejemplo 2.2.5 (Teorema de Picard). Sea $f(t,x)$ una función real continua definida para t en el intervalo $[0,b]$ y para x en $\mathbb{R}$. Un problema de condiciones iniciales de Cauchy es el problema de encontrar una función x continuamente diferenciable en $[0,b]$, verificando la ecuación diferencial

$$\begin{cases} x'(t) = f\big(t,x(t)\big), & t \in [0,b], \\ x(0) = \xi. \end{cases} \tag{2.8}$$

Un resultado clásico establece que si f es lipschitziana con respecto a x, *i.e.*, si existe un $L > 0$ tal que, para cualesquiera $x, y \in \mathbb{R}$,

$$\big|f(t,x) - f(t,y)\big| \leq L|x-y|, \qquad t \in [0,b],$$

entonces la solución de (2.8) existe y es única. Este hecho puede demostrarse de muchas formas. Nosotros vamos a elegir una aproximación que servirá para ilustrar nuestra discusión del principio de contracción de Banach.

Consideremos el espacio $C\big([0,b]\big)$ de las funciones reales continuas con la norma estándar del supremo. Integrando ambos lados de (2.8) obtenemos

$$x(t) = \xi + \int_0^t f\big(s,x(s)\big)\,ds.$$

Vamos a representar por Fx la función del lado derecho de la igualdad anterior, esto es,

$$(Fx)(t) = \xi + \int_0^t f\big(s,x(s)\big)\,ds.$$

Así, $F : C\big([0,b]\big) \longrightarrow C\big([0,b]\big)$, y una solución de (2.8) se corresponde a un punto fijo x de F.

Esta aproximación es la más comúnmente presentada en los libros de texto sobre ecuaciones diferenciales. Obsérvese que, para cualesquiera $x, y \in C\big([0,b]\big)$,

$$\begin{aligned} \big|(Fx)(t) - (Fy)(t)\big| &= \left|\int_0^t f\big(s,x(s)\big)\,ds - \int_0^t f\big(s,y(s)\big)\,ds\right| \\ &\leq \int_0^t \Big|f\big(s,x(s)\big) - f\big(s,y(s)\big)\Big|\,ds \\ &\leq \int_0^t L\big|x(s) - y(s)\big|\,ds \\ &\leq Lt\|x-y\|, \end{aligned}$$

de donde se deduce que

$$\|Fx - Fy\| \leq Lb\,\|x-y\|,$$

i.e., $k(F) \leq Lb$. Si $Lb < 1$, entonces el resultado es inmediato vía el principio de contracción de Banach. Sin embargo, si $Lb \geq 1$, se necesitan algunos pasos adicionales. Tomemos $h > 0$ tal que $Lh < 1$, y consideremos el espacio $C\big([0,h]\big)$. Sustituyendo b por h en el argumento anterior, obtenemos una *solución local* de (2.8), digamos x_0, en $C\big([0,h]\big)$. Ahora, consideremos el problema de Cauchy siguiente en $[h,2h]$:

$$\begin{cases} x_1'(t) = f\big(t,x_1(t)\big), \\ x_1(h) = x_0(h). \end{cases} \tag{2.9}$$

Aplicando la técnica utilizada al principio de este problema, obtenemos una única solución x_1 de (2.9) y, dado que $x_1(h) = x_0(h)$, x_1 extiende a x_0 de $[0,h]$ a $[0,2h]$. Esta extensión es diferenciable en h, ya que el problema de Cauchy tiene una *única* solución en un entorno de h. Ahora, es claro que el procedimiento descrito puede repetirse para el intervalo $[2h,3h]$, y por tanto, después de un número finito de pasos, obtendremos una solución de (2.8) válida en $[0,b]$. $\square$

Un sencillo argumento de división entera y el teorema 2.2.1 permiten obtener la siguiente versión reforzada de este último: *Sean (M,d) un espacio métrico completo y $T : M \longrightarrow M$ una aplicación continua tal que T^m es una contracción, para algún $m \in \mathbb{N}$. Entonces, T tiene un único punto fijo en M y, para cada $x_0 \in M$, la sucesión $(T^n x_0)_n$ converge a dicho punto fijo.* Este argumento permite dar una demostración ligeramente distinta del teorema de Picard 2.2.5, donde se puede evitar la discusión final sobre los casos $Lb \geq 1$ y $Lb < 1$, véase [55, p. 15-16].

El teorema del punto fijo de Banach. S. Banach, 1892-1945, contribuyó de forma decisiva en el desarrollo sistemático del análisis funcional, área en la que antes de él sólo habían aparecido resultados aislados. Banach demostró un buen número de resultados en espacios normados, y muchos teoremas importantes, que son todavía frecuentemente utilizados, llevan su nombre: los teoremas de Hahn-Banach, Banach-Steinhaus, Banach-Alaoglu, o Banach-Tarski. Muchos de estos resultados cuentan con aplicaciones numerosas. El teorema del punto fijo ha encontrado aplicaciones al estudio de la existencia y unicidad de soluciones de ecuaciones diferenciales ordinarias (problema de Cauchy), al estudio de funciones holomorfas del disco en sí mismo, (toda función holomorfa $f : D \longrightarrow D$ del disco unidad en sí mismo, tal que $\sup\{|f(z)| : z \in D\} < 1$, tiene un único punto fijo en D, [55, Theorem 2.5]), para dar una demostración del teorema de la función inversa para funciones de clase C^1 en $\mathbb{R}^n$, [107, Theorem 4.B], al análisis numérico (método de Newton), a la existencia de conjuntos fractales, etc.

2.3 El teorema del punto fijo de Brouwer

A lo largo de esta sección consideraremos en $\mathbb{R}^n$ la norma euclídea $\|\cdot\| (= \|\cdot\|_2)$ asociada al producto escalar natural $\langle x,y \rangle := \sum_{i=1}^n x_i y_i$, para $x = (x_i)_{i=1}^n$, $y = (y_i)_{i=1}^n$ en $\mathbb{R}^n$. Denotaremos, en general, por B^n y S^{n-1} la bola y la esfera unidad, respectivamente, dadas por

$$B^n := \{x \in \mathbb{R}^n : \|x\| \leq 1\} = \left\{x = (x_i)_{i=1}^n \in \mathbb{R}^n : \sum_{i=1}^n x_i^2 \leq 1\right\}$$

y

$$S^{n-1} := \{x \in \mathbb{R}^n : \|x\| = 1\} = \left\{x = (x_i)_{i=1}^n \in \mathbb{R}^n : \sum_{i=1}^n x_i^2 = 1\right\}.$$

En particular, para $n = 2$, identificaremos $\mathbb{C}$ con $\mathbb{R}^2$, escribiendo la bola unidad cerrada B^2 como $\overline{D}$, donde D es el disco unidad abierto $D := \{z \in \mathbb{C} : |z| < 1\}$; la esfera S^1 se escribirá como el *toro* $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$.

El resultado fundamental que demostraremos aquí es el siguiente:

Teorema 2.3.6 (Brouwer). *Sea $\varphi : B^n \longrightarrow B^n$ una aplicación continua de la bola unidad en sí misma. Entonces, φ tiene un punto fijo en B^n,* i.e.*, existe $x \in B^n$ tal que $\varphi(x) = x$.*

Un momento de reflexión sobre el teorema del punto fijo es suficiente para convencerse de que, cuando $n = 1$, éste se sigue del teorema del valor medio de Bolzano, tal y como ilustra la figura 2.4.

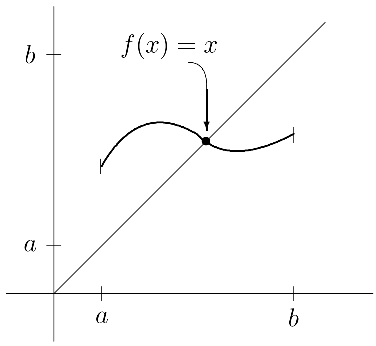


Figura 2.4: El teorema del punto fijo y el teorema de los valores intermedios

La prueba original de Brouwer fue primero establecida para $n = 2$, y posteriormente la generalizó a dimensiones mayores. En $\mathbb{C} = \mathbb{R}^2$, la demostración puede hacerse con el concurso del concepto de índice de un camino cerrado respecto a un punto. A continuación, exponemos cuáles son los argumentos que proporcionan esta prueba para $n = 2$, y después, abordaremos la demostración del caso general.

Argumentos y logaritmos de una función continua Consideremos un espacio topológico X y una función continua $f : X \longrightarrow \mathbb{C} \setminus \{0\}$. Si $g : X \longrightarrow \mathbb{C}$ es otra función continua de forma que $e^{g(x)} = f(x)$, para todo $x \in X$, entonces se dice que g es un *logaritmo continuo* de f en X. Si

$\alpha : X \longrightarrow \mathbb{R}$ es una función continua tal que $\frac{f(x)}{|f(x)|} = e^{i\alpha(x)}$, para todo $x \in X$, se dice que α es un *argumento continuo* de f en X.

Si X es un espacio topológico conexo, dos logaritmos (respectivamente, argumentos) continuos de f difieren en un múltiplo entero de $2\pi i$ (respectivamente, 2π).

Caminos Un *camino* en un espacio topológico X es una función continua $\gamma : [a,b] \longrightarrow X$. El punto $\gamma(a)$ se denomina origen del camino y el punto $\gamma(b)$ se llama extremo. El camino γ se dice cerrado si $\gamma(a) = \gamma(b)$.

El hecho notable que permite definir el concepto de índice de un camino cerrado en $\mathbb{C}$ es la siguiente proposición, cuya prueba está tomada de [106]. Una demostración más laboriosa, que utiliza el concepto de primitiva a lo largo de un camino de una forma diferencial cerrada, puede encontrarse en [19].

Proposición 2.3.1. *Todo camino continuo $\gamma : [a,b] \longrightarrow \mathbb{C} \setminus \{0\}$ posee un logaritmo continuo.*

Demostración. La siguiente observación se aplicará varias veces: si $a \leq x \leq y \leq z \leq b$ y los caminos $\gamma|_{[x,y]}$, $\gamma|_{[y,z]}$ poseen logaritmo continuo, entonces $\gamma|_{[x,z]}$ también posee logaritmo continuo. Efectivamente, si $g_1 : [x,y] \longrightarrow \mathbb{C}$ es un logaritmo continuo de $\gamma|_{[x,y]}$ y $g_2 : [y,z] \longrightarrow \mathbb{C}$ es un logaritmo continuo de $\gamma|_{[y,z]}$, se verifica que $g_1(y), g_2(y) \in \log \gamma(y)$; luego $g_1(y) - g_2(y) = 2m\pi i$, para algún $m \in \mathbb{Z}$. La función $g : [x,z] \longrightarrow \mathbb{C}$ definida por $g(t) = g_1(t)$ si $t \in [x,y]$, y $g(t) = g_2(t) + 2m\pi i$ si $t \in [y,z]$, es un logaritmo continuo de $\gamma|_{[x,z]}$.

Sea $X \subset [a,b]$ el conjunto de los puntos $x \in [a,b]$ tales que $\gamma|_{[a,x]}$ posee un logaritmo continuo. Basta probar que $b \in X$. Para ello, se empezará viendo que X no es vacío, y que si $\alpha = \sup X$, entonces $\alpha \in X$ y $\alpha = b$.

Por la continuidad de γ en a, existe $\delta > 0$ tal que $\gamma\big([a,\delta]\big) \subset D\big(\gamma(a), |\gamma(a)|\big)$. En el disco $D\big(\gamma(a), |\gamma(a)|\big)$ hay definido un logaritmo continuo de la identidad, y por tanto, existe un logaritmo continuo de $\gamma|_{[a,\delta]}$, es decir, $[a,\delta] \subset X$; luego $X \neq \emptyset$ y $\alpha = \sup X > 0$.

Por la continuidad de γ en α, encontramos $a < c < \alpha$ tal que $\gamma\big([c,\alpha]\big) \subset D\big(\gamma(\alpha), |\gamma(\alpha)|\big)$, y razonando como antes, se obtiene que $\gamma|_{[c,\alpha]}$ posee un logaritmo continuo. Además, por definición de supremo, al ser $\alpha = \sup X$, existe $x \in X$ tal que $c < x \leq \alpha$. Pero, dado que $\gamma|_{[a,x]}$ y $\gamma|_{[x,\alpha]}$ tienen logaritmo continuo, utilizando la observación preliminar se obtiene que $\alpha \in X$. La prueba se concluye demostrando, por reducción al absurdo, que $\alpha = b$. Si fuese $\alpha < b$, razonando de forma análoga, podríamos encontrar un $d \in (\alpha, b]$ de tal forma que $\gamma|_{[\alpha,d]}$ tendría logaritmo continuo. Como ya hemos visto que $\gamma|_{[a,\alpha]}$ posee logaritmo continuo, utilizando de nuevo la observación preliminar se deduciría que $\gamma|_{[a,d]}$ también lo tendría, es decir, $d \in X$, lo que contradice la definición de $\alpha = \sup X$. $\square$

Cuando γ es un camino regular a trozos, se puede dar una fórmula explícita para un logaritmo continuo de γ. La prueba del lema que sigue se puede encontrar en cualquier libro clásico de variable compleja, como [96, Teorema 10.10] o [35, Proposition 4.1].

Lema 2.3.2. *Si $\gamma : [a,b] \longrightarrow \mathbb{C} \setminus \{0\}$ es un camino regular a trozos y $\alpha \in \log \gamma(a)$, la función*

$$g(t) = \alpha + \int_a^t \frac{\gamma'(s)}{\gamma(s)}\, ds,$$

para $a \leq t \leq b$, es un logaritmo continuo de γ.

Demostración. En virtud del teorema fundamental del cálculo integral, la función g es continua en $[a,b]$, y existe un conjunto finito $H \subset [a,b]$ tal que g es derivable en cada $t \in [a,b] \setminus H$, con derivada $g'(t) = \gamma'(t)/\gamma(t)$. Puesto que la derivada de $\gamma(t)e^{-g(t)}$ existe y es nula para cada $t \in [a,b] \setminus H$, se sigue que la función $\gamma(t)e^{-g(t)}$ permanece constante entre cada par de puntos consecutivos de H. Entonces, en virtud de la continuidad, esta función es constante en todo $[a,b]$, con valor constante $\gamma(a)e^{-g(a)} = \gamma(a)e^{-\alpha} = 1$. Se obtiene así que $\gamma(t) = e^{g(t)}$, para todo $t \in [a,b]$. □

Índice de un camino cerrado respecto de un punto Si $\gamma : [a,b] \longrightarrow \mathbb{C} \setminus \{0\}$ es un camino continuo y cerrado, y $g : [a,b] \longrightarrow \mathbb{C}$ es un logaritmo continuo de γ, se verifica que $g(a)$ y $g(b)$ son logaritmos de $\gamma(a) = \gamma(b)$, por lo que existe $m \in \mathbb{N}$ tal que $g(b) - g(a) = 2m\pi i$. Obsérvese que si $h : [a,b] \longrightarrow \mathbb{C}$ es otro logaritmo continuo de γ, existe $k \in \mathbb{Z}$ tal que $h(t) = g(t) + 2k\pi i$; luego $h(b) - h(a) = g(b) - g(a)$. Por consiguiente, el número entero

$$m = \frac{1}{2\pi i}\big(g(b) - g(a)\big)$$

no depende del logaritmo continuo de γ que se haya considerado. Este valor m indica el número de vueltas, en sentido positivo, que da el camino $\gamma(t)$ alrededor de 0 cuando t recorre, de modo creciente, el intervalo $[a,b]$.

En general, el número de vueltas que da un camino cerrado y continuo γ, alrededor de un punto $z_0 \notin \gamma\big([a,b]\big)$, se obtiene contando las vueltas que da, alrededor del origen, el camino trasladado $\gamma(t) - z_0$: se define entonces el *índice* (o número de vueltas) del camino γ respecto al punto z_0, como el número entero

$$\operatorname{Ind}(\gamma, z_0) = \frac{1}{2\pi i}\big(g(b) - g(a)\big),$$

donde $g : [a,b] \longrightarrow \mathbb{C}$ es un logaritmo continuo del camino $\gamma(t) - z_0$. Si γ es un camino cerrado, regular a trozos, utilizando el logaritmo continuo dado en 2.3.2 se obtiene que

$$\operatorname{Ind}(\gamma, z_0) = \frac{1}{2\pi i} \int_a^b \frac{\gamma'(s)}{\gamma(s) - z_0}\, ds = \frac{1}{2\pi i} \int_\gamma \frac{dz}{z - z_0}.$$

La siguiente proposición recoge las propiedades que necesitamos para demostrar el teorema del punto fijo en $\mathbb{C}$; una prueba de estas propiedades puede encontrarse, por ejemplo, en [19, Proposition 8.2 y 8.3] y [35, Proposition 4.3 y Theorem 4.4].

Proposición 2.3.3. *Si $\gamma_0, \gamma_1 : [a,b] \longrightarrow \mathbb{C} \setminus \{0\}$ son caminos cerrados y continuos, se verifica:*

(i) $\operatorname{Ind}(\gamma_0\gamma_1, 0) = \operatorname{Ind}(\gamma_0, 0) + \operatorname{Ind}(\gamma_1, 0)$.

(ii) $\operatorname{Ind}(\gamma_1, 0) = \operatorname{Ind}(\gamma_0, 0)$ *si* $\big|\gamma_0(t) - \gamma_1(t)\big| \leq \big|\gamma_1(t)\big|$, *para todo* $t \in [a,b]$.

(iii) Sea $\gamma : [a,b] \longrightarrow \mathbb{C}$ *un camino cerrado y continuo. La función* $z \to \operatorname{Ind}(\gamma, z)$, *definida en el abierto* $V := \mathbb{C} \setminus \gamma\big([a,b]\big)$, *permanece constante en cada componente conexa de* V, *y vale* 0 *en la componente conexa no acotada.*

(iv) Si γ_0 *y* γ_1 *son homotópicos por caminos cerrados mediante una homotopía que no pasa por 0, entonces* $\operatorname{Ind}(\gamma_0, 0) = \operatorname{Ind}(\gamma_1, 0)$.

Proposición 2.3.4 (Teorema de Brouwer en $\mathbb{R}^2$)**.** *Si* $\overline{D}$ *es la bola unidad cerrada de* $\mathbb{C}$, *sea* $f : \overline{D} \longrightarrow \mathbb{C}$ *una función continua. Se verifica:*

(i) Si $\gamma(t) = f(e^{it})$, $t \in [0, 2\pi]$, *y* $a \notin \gamma\big([0, 2\pi]\big)$ *satisface que* $\operatorname{Ind}(\gamma, a) \neq 0$, *entonces existe* $z \in D$ *tal que* $f(z) = a$.

(ii) Si $f\big(\overline{D}\big) \subset \overline{D}$, *entonces* f *tiene un punto fijo,* i.e.*, existe* $z \in \overline{D}$ *tal que* $f(z) = z$.

Demostración. La afirmación *(i)* se demuestra razonando por reducción al absurdo: si suponemos que $a \notin f(D)$, entonces la aplicación $\Gamma : [0,1] \times [0, 2\pi] \longrightarrow \mathbb{C} \setminus \{a\}$, dada por $\Gamma(r,t) := f(re^{it})$, establece una homotopía por caminos cerrados entre γ y el camino constante $f(0)$. Consecuentemente, $\operatorname{Ind}(\gamma, a)$ debe ser cero, después de la propiedad *(iv)* en la proposición 2.3.3.

La prueba de *(ii)* es ahora una sencilla consecuencia del apartado anterior: supongamos que se satisface $f\big(\overline{D}\big) \subset \overline{D}$. Encontrar un punto fijo de f en $\overline{D}$ es encontrar un cero de la función $g(z) = f(z) - z$, para $z \in \overline{D}$. Si g se anula en $\mathbb{T} = \{w \in \mathbb{C} : |w| = 1\}$ ya hemos terminado. Si g no se anula en $\mathbb{T}$, el camino $\gamma_0(t) = g(e^{it})$, para $t \in [0, 2\pi]$, no pasa por el origen. Llamando $\gamma_1(t) = -e^{it}$, $t \in [0, 2\pi]$, se deduce que

$$\big|\gamma_0(t) - \gamma_1(t)\big| = \big|f(e^{it})\big| \leq 1 = |\gamma_1(t)|, \text{ para todo } t \in [0, 2\pi].$$

Si ahora utilizamos el apartado *(ii)* de la proposición 2.3.3, se obtiene que $\operatorname{Ind}(\gamma_0, 0) = -1$, y el apartado anterior se aplica entonces para deducir la existencia de un punto $z \in D$ tal que $g(z) = 0$, es decir, $f(z) = z$. La demostración ha terminado. □

Funciones continuamente diferenciables Sean $\Omega \subset \mathbb{R}^n$ y $f : \Omega \longrightarrow \mathbb{R}^m$. Se dice que la función f es *de clase C^1 en* Ω, y se escribe $f \in C^1(\Omega)$ (en la notación damos por supuesto el espacio donde f toma valores), si f posee derivadas parciales continuas en todo $x \in \Omega$. Equivalentemente, f es de clase C^1 en Ω si f diferenciable y, además, la aplicación diferencial $df : \Omega \longrightarrow L(\mathbb{R}^n, \mathbb{R}^m)$ es continua. Las funciones de clase C^1 se denominan también funciones continuamente diferenciables, véase [2]. La función f se dice *de clase* ∞, y se escribe $f \in C^\infty(\Omega)$, si admite derivadas parciales de todos los órdenes.

Teorema de la aplicación abierta Sea $f : \Omega \longrightarrow \mathbb{R}^n$ continua e inyectiva en un abierto $\Omega \subset \mathbb{R}^n$. Si suponemos que, para cada $x \in \Omega$, existen las derivadas parciales $\frac{\partial f_k}{\partial x_i}(x)$, $1 \leq i, k \leq n$, y que el determinante jacobiano

$$\det \left[\frac{\partial f_k}{\partial x_i}(x) \right]$$

es no nulo para cada $x \in \Omega$, entonces f es abierta, es decir, $f(V)$ es abierto para cada abierto $V \subset \Omega$, véase [2, Teorema 13.3].

La demostración del teorema del punto fijo 2.3.6 en dimensión finita arbitraria la hacemos en dos pasos: primero, para funciones diferenciables con continuidad, y después, con la ayuda del teorema de Stone-Weierstrass 2.1.4, para funciones continuas.

Teorema 2.3.5. *Sean $\Omega \subset \mathbb{R}^n$ un subconjunto abierto tal que $B^n \subset \Omega$, y $\psi : \Omega \longrightarrow \mathbb{R}^n$ una función de clase $C^1(\Omega)$ tal que $\psi(B^n) \subset B^n$. Entonces, existe $x \in B^n$ tal que $\psi(x) = x$.*

Demostración. Procedemos por reducción al absurdo. Vamos a suponer que $\psi(x) \neq x$, para cada $x \in B^n$. Definimos entonces la función

$$y(x) = \frac{\langle x, \psi(x) - x\rangle + \sqrt{d(x)}}{\|x - \psi(x)\|^2},$$

donde $d(x) = \big\langle x, x - \psi(x)\big\rangle^2 + \big(1 - \|x\|^2\big)\,\|x - \psi(x)\|^2$. Veamos, en primer lugar, que $d(x) > 0$ para cada $x \in B^n$. Si $\|x\| < 1$, entonces $d(x) \geq \big(1 - \|x\|^2\big)\,\|x - \psi(x)\|^2 > 0$. Si $\|x\| = 1$, entonces

$$\begin{aligned} d(x) &= \big\langle x, x - \psi(x)\big\rangle^2 = \Big(\|x\|^2 - \big\langle x, \psi(x)\big\rangle\Big)^2 = \frac{1}{4}\Big(2\|x\|^2 - 2\big\langle x, \psi(x)\big\rangle\Big)^2 \\ &= \frac{1}{4}\Big(\|x\|^2 + \|\psi(x)\|^2 - 2\big\langle x, \psi(x)\big\rangle + \|x\|^2 - \|\psi(x)\|^2\Big)^2 \\ &= \frac{1}{4}\Big(\|x - \psi(x)\|^2 + 1 - \|\psi(x)\|^2\Big)^2. \end{aligned}$$

Como $\|\psi(x)\| \leq 1$, obtenemos que $d(x) \geq \frac{1}{4}\|x - \psi(x)\|^4 > 0$, si $\|x\| = 1$. Ahora, del hecho de que la función $f(t) = \sqrt{t}$ sea derivable con continuidad en $(0, \infty)$, obtenemos que $\sqrt{d} \in C^1(\Omega)$; luego $y \in C^1(\Omega)$.

Es claro que $y(x)$ es raíz de la ecuación en y

$$\|x - \psi(x)\|^2 y^2 + 2\big\langle x, x - \psi(x)\big\rangle y + \|x\|^2 - 1 = 0.$$

Esto fuerza a que se verifique la siguiente igualdad para $y = y(x)$:

$$\big\|x + y\big(x - \psi(x)\big)\big\|^2 = 1. \tag{2.10}$$

En efecto,

$$\begin{aligned}\big\|x+y\big(x-\psi(x)\big)\big\|^2 &= \Big\langle x+y\big(x-\psi(x)\big), x+y\big(x-\psi(x)\big)\Big\rangle \\ &= \langle x,x\rangle + 2\big\langle x-\psi(x),x\big\rangle y + \big\langle x-\psi(x), x-\psi(x)\big\rangle y^2 \\ &= \|x\|^2 + 2\big\langle x, x-\psi(x)\big\rangle y + \big\|x-\psi(x)\big\|^2 y^2 = 1.\end{aligned}$$

Observemos que, si $\|x\| = 1$, entonces $y(x) = 0$. Veámoslo: si $\|x\| = 1$, se tiene que

$$y(x) = \frac{\big\langle x, \psi(x)-x\big\rangle + \big|\big\langle x, x-\psi(x)\big\rangle\big|}{\|x-\psi(x)\|^2}.$$

Por la desigualdad de Cauchy-Schwarz, se tiene que $\big\langle x, \psi(x)\big\rangle \leq \|x\|\,\|\psi(x)\| \leq 1$. Ahora,

$$\big\langle x, x-\psi(x)\big\rangle = \langle x,x\rangle - \big\langle x,\psi(x)\big\rangle = 1 - \big\langle x,\psi(x)\big\rangle \geq 0,$$

luego

$$\big|\big\langle x, x-\psi(x)\big\rangle\big| = \big\langle x, x-\psi(x)\big\rangle,$$

y así, $y(x) = 0$ para $\|x\| = 1$.

Definimos las aplicaciones

$$g(x) := y(x)\big(x-\psi(x)\big) \quad \text{y} \quad f_t(x) := x + tg(x),$$

para $x \in B^n$ y $0 \leq t \leq 1$. Como $y, \psi \in C^1(\Omega)$, se tiene que $g, f_t \in C^1(\Omega)$. Entonces, existe un $c > 0$ tal que

$$\big\|g(u) - g(v)\big\| \leq c\|u-v\|, \tag{2.11}$$

para todo $u, v \in B^n$.

Sean $f_t(x) = \big(f_{t,1}(x), f_{t,2}(x), \ldots, f_{t,n}(x)\big)$ y $g(x) = \big(g_1(x), g_2(x), \ldots, g_n(x)\big)$. Vamos a representar por F_t y G las matrices jacobianas de f_t y g respectivamente, *i.e.*,

$$F_t(x) := Df_t(x) = \left[\frac{\partial f_{t,i}}{\partial x_j}(x)\right] \quad \text{y} \quad G(x) := Dg(x) = \left[\frac{\partial g_i}{\partial x_j}(x)\right].$$

Se tiene entonces que

$$F_t(x) = I + tG(x), \tag{2.12}$$

donde I es la matriz unidad.

Como $g \in C^1(\Omega)$, se deduce que $\det F_t(x)$ es una aplicación continua sobre $[0,1] \times B^n$ y, dado que $[0,1] \times B^n$ es un compacto, concluimos que $\det F_t(x)$ es uniformemente continua en $[0,1] \times B^n$. Además, al ser $\det F_0(x) = 1$, existe un $t_0 \in (0,1]$ tal que

$$\det F_t(x) > 0, \tag{2.13}$$

para todo $(t,x) \in [0,t_0] \times B^n$. Sin pérdida de generalidad, podemos suponer que $t_0 < 1/c$. Vamos a demostrar que f_t es una aplicación inyectiva que satisface

$$f_t(B^n) = B^n, \tag{2.14}$$

si $0 \leq t \leq t_0$. Para ello, fijamos $t \in [0,t_0]$. Probaremos, en primer lugar, que f_t es inyectiva. Si $f_t(u) = f_t(v)$, se tiene, por (2.11), que

$$\|u - v\| = \big\|tg(u) - tg(v)\big\| \leq ct\|u - v\|.$$

Pero como $ct < 1$, obtenemos que $\|u - v\| = 0$, esto es, que $u = v$. Por otro lado, dado que $y(x) = 0$ si $x \in S^{n-1}$, deducimos que $g(x) = 0$. Luego,

$$f_t(x) = x, \quad \text{si } x \in S^{n-1}. \tag{2.15}$$

De (2.10) se tiene que, para todo $x \in B^n$,

$$f_1(x) \in S^{n-1}. \tag{2.16}$$

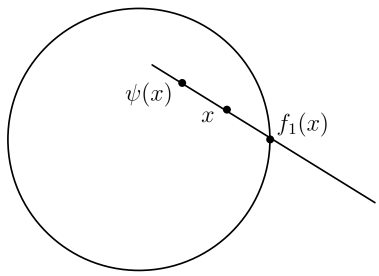


Figura 2.5: S^{n-1} no es un retracto de B^n

Por tanto, como B^n es convexo y $f_t(x) = (1-t)x + tf_1(x)$, obtenemos que $f_t(B^n) \subset B^n$. Utilizando ahora (2.15) para demostrar (2.14), es suficiente ver que

$$f_t(\operatorname{int} B^n) = \operatorname{int} B^n. \tag{2.17}$$

La condición (2.13) nos asegura que f_t es una aplicación abierta; luego $f_t(\operatorname{int} B^n)$ es abierto y $f_t(\operatorname{int} B^n) \subset \operatorname{int} B^n$. Para probar (2.17) basta demostrar que $B^n \setminus f_t(\operatorname{int} B^n) \subset S^{n-1}$. En efecto; sea $z \in B^n \setminus f_t(\operatorname{int} B^n)$. Vamos a ver que $\|z\| = 1$. Si cogemos un punto u del conjunto abierto $f_t(\operatorname{int} B^n)$, podemos encontrar otro punto v que pertenezca a la intersección del segmento $[u,z]$ con la frontera de $f_t(\operatorname{int} B^n)$, *i.e.*,

$$v \in [u,z] \cap \left(\overline{f_t(\operatorname{int} B^n)} \setminus f_t(\operatorname{int} B^n)\right).$$

Entonces, existen $v_n \in f_t(\operatorname{int} B^n)$ tales que $\lim_n \|v_n - v\| = 0$. Pongamos $v_n = f_t(w_n)$. Dado que B^n es compacto, sin pérdida alguna de generalidad podemos suponer que $\lim_n \|w_n - w\| = 0$, para algún $w \in B^n$. Además, ya que f_t es continua, obtenemos que $f_t(w) = v$ y, como $v \notin f_t(\operatorname{int} B^n)$, $w \notin \operatorname{int} B^n$. Luego $\|w\| = 1$, y por (2.15), se sigue que $w = v$. Ahora bien, v pertenece al segmento $[u, z]$ y $u \in \operatorname{int} B^n$; en consecuencia, $z \in S^{n-1}$. Esto concluye la demostración de que $B^n \setminus f_t(\operatorname{int} B^n) \subset S^{n-1}$ y, por tanto, podemos asegurar (2.17) y (2.14).

Sea ahora

$$P(t) := \int_{f_t^{-1}(B^n)} \det F_t(x)\, dx_1 dx_2 \dots dx_n,$$

para $0 \leq t \leq 1$. Utilizando el teorema del cambio de variable, (2.13) y (2.14), obtenemos que, si $0 \leq t \leq t_0$, entonces

$$P(t) = \int_{f_t^{-1}(B^n)} \det F_t(x)\, dx_1 \dots dx_n = \int_{B^n} dx_1 \dots dx_n = \operatorname{Vol}(B^n) > 0.$$

Como se verifica (2.12) y (2.14) entonces $P(t)$ es un polinomio, y así, $P(t) = \operatorname{Vol}(B^n)$, para todo $0 \leq t \leq 1$. Por otro lado, de (2.16) se tiene que, para cualquier $x \in B^n$,

$$\sum_{k=1}^{n} f_{1,k}^2(x) = 1; \tag{2.18}$$

luego

$$\sum_{k=1}^{n} \frac{\partial f_{1,k}}{\partial x_i}(x) f_{1,k}(x) = 0, \qquad i = 1, 2, \dots, n.$$

Además, por (2.18), obtenemos que el sistema lineal

$$\sum_{k=1}^{n} \frac{\partial f_{1,k}}{\partial x_i}(x) u_k = 0, \qquad i = 1, 2, \dots, n,$$

con incógnitas $u_1, u_2, \dots, u_n$, tiene solución no trivial. Entonces, $\det F_1(x) = 0$ para todo $x \in B^n$, lo que implica que $P(1) = 0$. Sin embargo, $P(1) = \operatorname{Vol}(B^n) > 0$, lo que nos da una contradicción que concluye la prueba. □

Teorema 2.3.6 (Brouwer). *Sea $\varphi : B^n \longrightarrow B^n$ una aplicación continua de la bola unidad en sí misma. Entonces, φ tiene un punto fijo en B^n,* i.e.*, existe $x \in B^n$ tal que $\varphi(x) = x$.*

Demostración. Probaremos primero que, dado $\varepsilon > 0$, existe $\psi_\varepsilon = (\psi_{1,\varepsilon}, \dots, \psi_{n,\varepsilon})$ de forma que $\psi_{i,\varepsilon} \in C^\infty(\mathbb{R}^n)$, $\psi_\varepsilon : B^n \longrightarrow B^n$ y $\|\varphi(x) - \psi_\varepsilon(x)\| < \varepsilon$ para cada $x \in B^n$.

Sea $\varphi = (\varphi_1, \dots, \varphi_n)$. Definimos $\eta := \min\left\{1, \frac{\varepsilon}{2\sqrt{n}}\right\}$. Como $\varphi_i \in C(B^n)$, en virtud del teorema de Weierstrass, corolario 2.1.5, existe un polinomio $\psi_{i,\varepsilon}$ en las variables $[x_1, \dots, x_n]$ tal que, para cada $x=(x_1,\dots,x_n)\in B^n$, $\left|(1-\eta)\varphi_i(x)-\psi_{i,\varepsilon}(x)\right|<\eta/(2n)$, para todo $i=1,2,\dots,n$. Ahora, dado que $\left|\varphi_i(x)\right|\le\left\|\varphi(x)\right\|\le 1$, para cada $x\in B^n$ y cada $i=1,2,\dots,n$, obtenemos que

$$\left|\varphi_i(x)-\psi_{i,\varepsilon}(x)\right|\le\left|(1-\eta)\varphi_i(x)-\psi_{i,\varepsilon}(x)\right|+\eta\left|\varphi_i(x)\right|<\frac{\eta}{2n}+\eta\le 2\eta,$$

y por tanto, para cada $x\in B^n$, se deduce que

$$\left\|\varphi(x)-\psi_\varepsilon(x)\right\|<2\eta\sqrt{n}\le\varepsilon,$$

donde $\psi_\varepsilon=(\psi_{1,\varepsilon},\dots,\psi_{n,\varepsilon})$. Como cada $\psi_{i,\varepsilon}$ es un polinomio, se tiene que $\psi_{i,\varepsilon}\in C^\infty(\mathbb{R}^n)$. Veamos que $\psi_\varepsilon(x)\in B^n$, para cada $x\in B^n$. Dado $x\in B^n$

$$\begin{aligned}
\left\|\psi_\varepsilon(x)\right\|^2&=\sum_{i=1}^n\psi_{i,\varepsilon}(x)^2=\sum_{i=1}^n\Big((1-\eta)^2\varphi_i(x)^2+\psi_{i,\varepsilon}(x)^2-(1-\eta)^2\varphi_i(x)^2\Big)\\
&=(1-\eta)^2\left\|\varphi(x)\right\|^2+\sum_{i=1}^n\Big(\psi_{i,\varepsilon}(x)^2-(1-\eta)^2\varphi_i(x)^2\Big)\\
&\le(1-\eta)^2+\sum_{i=1}^n\left|\psi_{i,\varepsilon}(x)+(1-\eta)\varphi_i(x)\right|\cdot\left|\psi_{i,\varepsilon}(x)-(1-\eta)\varphi_i(x)\right|\\
&\le(1-\eta)^2+\sum_{i=1}^n\Big(2(1-\eta)\left|\varphi_i(x)\right|+\left|\psi_{i,\varepsilon}(x)-(1-\eta)\varphi_i(x)\right|\Big)\left|\psi_{i,\varepsilon}(x)-(1-\eta)\varphi_i(x)\right|\\
&<(1-\eta)^2+n\Big(2(1-\eta)+\frac{\eta}{2n}\Big)\frac{\eta}{2n}<1-\eta+2n\frac{\eta}{2n}=1.
\end{aligned}$$

Para concluir la prueba tomamos, para cada $k\in\mathbb{N}$, una función $\psi_k=(\psi_{1,k},\dots,\psi_{n,k})$ tal que $\psi_{i,k}\in C^\infty(\mathbb{R}^n)$, $\psi_k:B^n\longrightarrow B^n$ y $\left\|\varphi(x)-\psi_k(x)\right\|\le 1/k$, si $x\in B^n$. El teorema 2.3.5 anterior nos asegura que, para cada $k\in\mathbb{N}$, existe $x_k\in B^n$ con $\psi_k(x_k)=x_k$. Usando ahora que B^n es compacto, existe una subsucesión convergente $(x_{k_m})_m$ hacia un cierto $x\in B^n$ y, por ser φ continua, $\big(\varphi(x_{k_m})\big)_m$ converge a $\varphi(x)$. Veamos, por último, que $\varphi(x)=x$. En efecto, como $\psi_{k_m}(x_{k_m})=x_{k_m}$, tendremos

$$\begin{aligned}
\left\|\varphi(x)-x\right\|&\le\left\|\varphi(x)-\varphi(x_{k_m})\right\|+\left\|\varphi(x_{k_m})-\psi_{k_m}(x_{k_m})\right\|+\|x_{k_m}-x\|\\
&\le\left\|\varphi(x)-\varphi(x_{k_m})\right\|+\frac{1}{k_m}+\|x_{k_m}-x\|\underset{m}{\longrightarrow}0;
\end{aligned}$$

luego $\left\|\varphi(x)-x\right\|=0$ y $\varphi(x)=x$. $\square$

H **Teorema del punto fijo de Brouwer, 1817-1966.** L. E. J. Brouwer, 1881-1966, llevó a cabo casi todo su trabajo en topología entre los años 1909 y 1913. Como ocurre a veces en matemáticas, Brouwer no fue el primero en demostrar *su teorema* del punto fijo para aplicaciones $\varphi:B^n\longrightarrow B^n$. Para $n=1$, el teorema es una sencilla consecuencia del teorema del valor medio de Bolzano, 1817. Para $n = 2$ y $n = 3$, Brouwer proporcionó una prueba en 1909. Un año más tarde J. Hadamard obtuvo una demostración para n arbitrario; finalmente, en 1912, el propio Brouwer propuso una nueva demostración, en el caso general, distinta de la Hadamard. El teorema del punto fijo de Brouwer cuenta con demostraciones y generalizaciones muy diversas en topología algebraica (véase [110]), geometría diferencial (véase [10]), y análisis (véase [93]). Se conoce que el teorema del punto fijo de Brouwer es equivalente a otros resultados clásicos: S^{n-1} no es un retracto de B^n; no existe un campo de vectores continuo, que no se anule en ningún punto, y tangente a la esfera S^{2n} (teorema de la bola peluda), véase [55, Chapter 18].

2.4 Los teoremas del punto fijo de Schauder y de Tychonoff

EN esta sección demostraremos los teoremas de Schauder 2.4.9 y de Tychonoff 2.4.15, que son la versiones infinito-dimensionales del teorema de Brouwer 2.3.6 demostrado en la sección anterior. Los primeros resultados que establecemos son resultados preliminares que utilizaremos con posterioridad.

Normas equivalentes Dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$ en un espacio vectorial X se dicen *equivalentes* si las topologías asociadas coinciden. O lo que es lo mismo, $\|\cdot\|_1$ y $\|\cdot\|_2$ en un espacio vectorial X son equivalentes si existen constantes $c, k > 0$ tales que $c\|x\|_1 \leq \|x\|_2 \leq k\|x\|_1$, para todo $x \in X$. En los espacios de dimensión finita (en $\mathbb{R}^n$) todas las normas son equivalentes.

Lema 2.4.1. *Sea X un espacio vectorial con dos normas equivalentes $\|\cdot\|_1$ y $\|\cdot\|_2$. Representamos por $B_1 = \{x \in X : \|x\|_1 \leq 1\}$ y $B_2 = \{x \in X : \|x\|_2 \leq 1\}$. Entonces, B_1 es homeomorfo a B_2.*

Demostración. Definimos $T : X \longrightarrow X$ mediante la expresión

$$T(x) := \begin{cases} 0 & \text{si} \quad x = 0, \\ \dfrac{\|x\|_1}{\|x\|_2} x & \text{si} \quad x \neq 0. \end{cases}$$

Tomemos $x \neq 0$ y pongamos $y = T(x)$. Entonces, $\|y\|_2 = \|x\|_1$ y $\|y\|_1 = \|x\|_1^2 / \|x\|_2$. Como

$$x = \frac{\|x\|_2}{\|x\|_1} y = \frac{\|x\|_1^2}{\|y\|_1 \|x\|_1} y = \frac{\|y\|_2}{\|y\|_1} y,$$

obtenemos que

$$T^{-1}(y) = \begin{cases} 0 & \text{si} \quad y = 0, \\ \dfrac{\|y\|_2}{\|y\|_1} y & \text{si} \quad y \neq 0. \end{cases}$$

Además, al ser

$$\big\|T(x)\big\|_2 = \|x\|_1 \quad \text{y} \quad \big\|T^{-1}(y)\big\|_1 = \|y\|_2, \tag{2.19}$$

se deduce que $T(B_1) \subset B_2$ y $T^{-1}(B_2) \subset B_1$. Luego $T(B_1) = B_2$. Por otra parte, si $x, u \in X$,

$$\begin{aligned} T(x) - T(u) &= \frac{\|x\|_1}{\|x\|_2}(x-u) + \left(\frac{\|x\|_1}{\|x\|_2} - \frac{\|u\|_1}{\|u\|_2}\right) u \\ &= \frac{\|x\|_1}{\|x\|_2}(x-u) + \frac{\|x\|_1\big(\|u\|_2 - \|x\|_2\big) + \|x\|_2\big(\|x\|_1 - \|u\|_1\big)}{\|x\|_2\|u\|_2} u. \end{aligned}$$

Luego

$$\big\|T(x) - T(u)\big\|_2 \leq 2\frac{\|x\|_1}{\|x\|_2}\|x-u\|_2 + \|x-u\|_1,$$

lo cual establece claramente que T es continua en $x \neq 0$. La igualdad (2.19) nos proporciona la continuidad de T en 0. De forma similar, se obtiene que T^{-1} es continua, y la prueba termina. □

Corolario 2.4.2. *Sea X un espacio vectorial real finito-dimensional. Entonces:*

(i) Si $\|\cdot\|$ es una norma cualquiera en X, B_X es la bola unidad cerrada de X y $\varphi : B_X \longrightarrow B_X$ es continua, entonces existe $x \in B_X$ tal que $\varphi(x) = x$.

(ii) Si $X = \mathbb{R}^n$, $a = (a_i)_{i=1}^n$ es tal que $a_i > 0$, $B_a = \{x = (x_i)_{i=1}^n \in \mathbb{R}^n : |x_i| \leq a_i\}$ y $\varphi : B_a \longrightarrow B_a$ es continua, entonces existe $x \in B_a$ tal que $\varphi(x) = x$.

Demostración. La afirmación en *(i)* se demuestra como sigue. Si $\dim X = n$, entonces existe un isomorfismo lineal $\phi : \mathbb{R}^n \longrightarrow X$. Para $x \in \mathbb{R}^n$, definimos $\|x\| := \big\|\phi(x)\big\|$. La norma $\|\cdot\|$ así definida en $\mathbb{R}^n$ es equivalente a la norma euclídea $\|\cdot\|_2$. Entonces, por el lema anterior, existe un homeomorfismo T de la bola unidad euclídea B^n sobre la bola unidad cerrada B asociada a la norma $\|\cdot\|$. Observando que $\phi(B) = B_X$, el teorema de Brouwer 2.3.6 utilizado para la aplicación continua

$$T^{-1} \circ \phi^{-1} \circ \varphi \circ \phi \circ T : B^n \longrightarrow B^n,$$

nos garantiza la existencia de un punto $y \in B^n$ tal que $(T^{-1} \circ \phi^{-1} \circ \varphi \circ \phi \circ T)(y) = y$. De aquí se obtiene que $x = \phi\big(T(y)\big)$ es un punto fijo para φ.

La afirmación en *(ii)* se sigue de *(i)*, teniendo en cuenta que B_a es la bola unidad cerrada asociada a la norma en $\mathbb{R}^n$ dada por

$$\|x\| = \max_{1 \leq i \leq n} \frac{|x_i|}{a_i}, \qquad \text{para} \quad x = (x_i)_{i=1}^n \in \mathbb{R}^n. \qquad \square$$

Conjuntos con la propiedad del punto fijo Un espacio topológico K se dice que tiene la *propiedad del punto fijo* si cada función continua $\varphi : K \longrightarrow K$ tiene un punto fijo. La propiedad del punto fijo se conserva por homeomorfismos. Este hecho ya ha sido utilizado en el corolario 2.4.2, y será empleado con profusión en el resto del capítulo. Si X es un espacio topológico, un subconjunto $K \subset X$ se dice que tiene la propiedad del punto fijo si K, con la topología inducida, tiene la propiedad del punto fijo.

Cubo de Hilbert Sea $(e_n)_n$ la base canónica en ℓ^2 (todas las coordenadas de e_n son 0 salvo la coordenada n-ésima, que es 1). Si $x=(x_i)_{i=1}^{\infty}\in\ell^2$, tenemos que

$$x=\sum_{i=1}^{\infty}x_ie_i, \quad i.e., \quad \lim_{n\to\infty}\|x-P_nx\|=0,$$

donde $P_nx=\sum_{i=1}^{n}x_ie_i$, para $n\in\mathbb{N}$, y la norma considerada es la canónica de ℓ^2 dada por $\|x\|^2=\big\|(x_i)_i\big\|^2=\sum_{i=1}^{\infty}|x_i|^2$. En lo que sigue, utilizaremos la notación

$$R_nx=x-P_nx=\sum_{i=n+1}^{\infty}x_ie_i.$$

Entonces, para cada $x\in\ell^2$ se tiene que

$$\|x\|^2=\|P_nx\|^2+\|R_nx\|^2.$$

El *cubo de Hilbert* es el subconjunto de ℓ^2 definido por

$$Q:=\Big\{x=(x_i)_{i=1}^{\infty}\in\ell^2:|x_i|\le\frac{1}{2^{i-1}},\ i=1,2,\dots\Big\}.$$

Lema 2.4.3. *Si $x,y\in Q$, entonces, para cada $n=1,2,\dots$, se tienen las desigualdades*

$$\|R_nx\|^2\le\frac{1}{3\cdot4^{n-1}}, \tag{2.20}$$

$$\|x-y\|^2\le\big\|P_n(x-y)\big\|^2+\frac{1}{3\cdot4^{n-3}}. \tag{2.21}$$

Demostración. La desigualdad (2.20) se deduce directamente de la definición de Q, sumando la serie geométrica $\frac{1}{4^n}+\frac{1}{4^{n+1}}+\cdots$. Utilizando ahora la igualdad $\|x\|^2=\|P_nx\|^2+\|R_nx\|^2$ y (2.20), obtenemos

$$\begin{aligned}\|x-y\|^2&=\big\|P_n(x-y)\big\|^2+\big\|R_n(x-y)\big\|^2\le\big\|P_n(x-y)\big\|^2+\big(\|R_nx\|+\|R_ny\|\big)^2\\&\le\big\|P_n(x-y)\big\|^2+\left(\frac{2}{\sqrt{3}\cdot2^{n-2}}\right)^2=\big\|P_n(x-y)\big\|^2+\frac{1}{3\cdot4^{n-3}}.\end{aligned}$$ □

Lema 2.4.4. *El cubo de Hilbert Q es compacto con la topología inducida por ℓ^2.*

Demostración. Como ℓ^2 es un espacio métrico y completo y Q es un subconjunto cerrado, para demostrar que Q es compacto es suficiente probar que Q es totalmente acotado (véase la página 115): veremos que, para cada $\varepsilon>0$, existe una cantidad finita de puntos $y_1,y_2,\dots,y_m\in\ell^2$ tal que $Q\subset\bigcup_{k=1}^{m}(y_k+\varepsilon B_{\ell^2})$. Así, dado $\varepsilon>0$, encontramos un $N\in\mathbb{N}$ tal que $\frac{1}{4^{N-3}}<\varepsilon^2$. Como $Q_N=\{P_Nx:x\in Q\}$ es un conjunto acotado en un espacio de dimensión finita, existe un conjunto finito $\{y_k:k=1,2,\dots,m\}\subset Q_N$ de forma que, para cada $y\in Q_N$, existe un y_k con $\|y-y_k\|<\varepsilon/2$.

Fijado $x \in Q$, encontramos y_k tal que $\|P_N x - y_k\| < \varepsilon/2$. Como $P_N y_k = y_k$, de la desigualdad (2.21) obtenemos que

$$\|x - y_k\|^2 \le \|P_N x - y_k\|^2 + \frac{1}{3 \cdot 4^{N-3}} < \frac{\varepsilon^2}{4} + \frac{\varepsilon^2}{3} < \varepsilon^2. \qquad \square$$

Lema 2.4.5. *Sean $Q \subset \ell^2$ el cubo de Hilbert y $\varphi : Q \longrightarrow Q$ una aplicación continua. Entonces, existe $y \in Q$ tal que $\varphi(y) = y$.*

Demostración. Consideremos de nuevo las proyecciones definidas en la página 87 por la expresión $P_n x = \sum_{i=1}^n x_i e_i$, $n \in \mathbb{N}$. En el espacio $P_n(\ell^2)$ de dimensión n, el conjunto $P_n(Q)$ es la bola unidad para una norma que se identifica, mediante el isomorfismo natural, con la bola de $\mathbb{R}^n$ dada por $B_{a_n} = \big\{(x_i)_{i=1}^n : |x_i| \le a_i,\ i = 1, 2, \ldots\big\}$, para $a_n = \big(1, \frac{1}{2}, \frac{1}{2^2}, \ldots, \frac{1}{2^{n-1}}\big)$.

Para cada $n \in \mathbb{N}$, $\varphi_n = P_n \circ \varphi \circ P_n : P_n(Q) \longrightarrow P_n(Q)$ es continua. Por el corolario 2.4.2, existe $y_n \in P_n(Q)$ tal que $\varphi_n(y_n) = y_n$. Como $y_n \in P_n(Q)$, entonces $P_n y_n = y_n$, y así, $P_n\big(\varphi(y_n)\big) = y_n$. Además, cada y_n lo podemos ver como un elemento de Q, que es compacto, por lo que existirán $y \in Q$ y una subsucesión $(y_{n_m})_{m \in \mathbb{N}}$, tales que $\|y_{n_m} - y\| \to 0$. Dado que φ es continua, obtenemos $\big\|\varphi(y_{n_m}) - \varphi(y)\big\| \to 0$. Por otro lado, como $P_n^2 = P_n$, la desigualdad (2.21) nos asegura que

$$\big\|\varphi(y_n) - y_n\big\|^2 = \big\|\varphi(y_n) - P_n(\varphi(y_n))\big\|^2 \le \frac{1}{3 \cdot 4^{n-3}} \xrightarrow[n \to \infty]{} 0.$$

En consecuencia, $\varphi(y) = y$, y la demostración termina. $\square$

$\boxed{\text{Mejor aproximación}}$ Sean (M, d) un espacio métrico y $D \subset M$. Un elemento $x_0 \in M$ tiene una *mejor aproximación* $y_0 \in D$, si $d(x_0, D) = \operatorname{ínf}_{y \in D} d(x_0, y) = d(x_0, y_0)$.

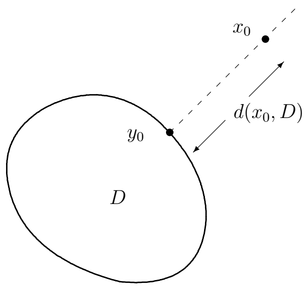


Figura 2.6: Mejor aproximación

Lema 2.4.6. *Sea K un subconjunto compacto de un espacio de Banach $\big(X, \|\cdot\|\big)$. Se tienen las siguientes propiedades:*

(i) Para cada $x \in X$, existe una mejor aproximación $P_K(x)$ de x en K.

(ii) Si $\big(X, \|\cdot\|\big)$ es estrictamente convexo y K es, además, convexo, entonces $P_K(x)$ es único para cada $x \in X$, y la aplicación $P_K : X \longrightarrow X$ es continua y satisface $P_K(x) = x$ para todo $x \in K$ (K es un retracto de X).

Demostración. La demostración de *(i)* es muy sencilla: fijemos $x \in X$ y sea $(y_n)_n$ una sucesión en K tal que $\lim_n \|x - y_n\| = d(x,K) =: d$. Como K es compacto, podemos suponer sin pérdida de generalidad, pasando a una subsucesión si es necesario, que existe $y \in K$ tal que $\lim_n \|y_n - y\| = 0$. Claramente, se tiene que $\|x - y\| = d$.

Veamos ahora si se satisfacen las hipótesis en *(ii)*. Supongamos que, para $x \in X$, existen dos puntos u, y en K tales que $\|x - u\| = \|x - y\| = d$. Como K es convexo, tenemos que $\frac{y+u}{2} \in K$. Así,

$$d \leq \left\| x - \frac{y+u}{2} \right\| = \frac{1}{2} \big\| (x-y) + (x-u) \big\| \leq \frac{1}{2} \big(\|x-y\| + \|x-u\| \big) = d.$$

Por tanto, tenemos la igualdad

$$d = \|x-u\| = \|x-y\| = \left\| \frac{(x-y)+(x-u)}{2} \right\|.$$

Como $\big(X, \|\cdot\|\big)$ es estrictamente convexo, se concluye que $x - y = x - u$ y, en consecuencia, $y = u$.

Veamos por último que P_K es continua. Sean $(x_n)_n$ y x en X tales que $\lim_n \|x_n - x\| = 0$. Dado que K es compacto, para ver que $\big(P_K(x_n)\big)_n$ converge hacia $P_K(x)$ es suficiente probar que $P_K(x)$ es el único punto de aglomeración de $\big(P_K(x_n)\big)_n$. Supongamos que $y \in K$ es un punto de aglomeración de $\big(P_K(x_n)\big)_n$. Existe, por tanto, una subsucesión $\big(P_K(x_{n_k})\big)_k$ tal que $\lim_k \big\| P_K(x_{n_k}) - y \big\| = 0$. Tenemos entonces que

$$d = \big\| x - P_K(x) \big\| = \lim_{k\to\infty} \big\| x_{n_k} - P_K(x) \big\| \geq \lim_{k\to\infty} \big\| x_{n_k} - P_K(x_{n_k}) \big\| = \|x - y\| \geq d.$$

De aquí se sigue que $d = \big\| x - P_K(x) \big\| = \|x - y\|$ y, finalmente, como $y \in K$, concluimos que $y = P_K(x)$ y la prueba termina. $\square$

Lema 2.4.7. *Sea K un subconjunto convexo y cerrado del cubo de Hilbert Q y sea $\varphi : K \longrightarrow K$ una aplicación continua. Entonces, existe $x \in K$ tal que $\varphi(x) = x$.*

Demostración. Sea $P_K : \ell^2 \longrightarrow \ell^2$ la aplicación que envía cada $x \in \ell^2$ sobre su mejor aproximación $P_K(x) \in K$. Definimos $\psi := \varphi \circ P_K|_Q$. Entonces, $\psi : Q \longrightarrow K \subset Q$. En virtud del lema 2.4.5 anterior, existe un punto fijo $x \in Q$ de ψ, es decir $\varphi\big(P_K(x)\big) = x$. Como $\varphi(K) \subset K$, obtenemos que $x \in K$, y así se tiene que $P_K(x) = x$. De aquí se concluye que $\varphi(x) = x$, y la prueba termina. $\square$

Es claro que, si el teorema del punto fijo se verifica para un espacio topológico dado, entonces se satisface para cualquier espacio homeomorfo. Esta observación y el lema anterior nos permiten demostrar fácilmente el teorema de Schauder.

Lema 2.4.8. *Si $(X, \|\cdot\|)$ es un espacio normado separable, entonces existe un conjunto numerable en X^* que separa los puntos de X. Más concretamente, si $\{x_n : n \in \mathbb{N}\}$ es un conjunto numerable denso en S_X, y para cada $n \in \mathbb{N}$ elegimos $f_n \in X^*$ tal que $f_n(x_n) \geq 1/2$, entonces*

$$D^* := \{f_n : n \in \mathbb{N}\}$$

separa los puntos de X.

Demostración. Es suficiente demostrar que, si $x \in S_X$, entonces, para algún $n \in \mathbb{N}$, se tiene que $f_n(x) \neq 0$. Para $x \in S_X$ dado, tomamos x_n tal que $\|x - x_n\| < 1/3$. Entonces,

$$f_n(x) = f_n(x_n) + f_n(x - x_n) \geq \frac{1}{2} - \|x - x_n\| \geq \frac{1}{2} - \frac{1}{3} > 0,$$

y la prueba termina. □

Teorema 2.4.9 (Schauder)**.** *Sea X un espacio normado y sea $K \subset X$ un conjunto convexo y compacto para la norma. Si $\varphi : K \longrightarrow K$ es continua, entonces existe $x \in K$ tal que $\varphi(x) = x$.*

Demostración. Vamos a demostrar que K es homeomorfo a un subconjunto compacto del cubo de Hilbert Q, para lo que suponemos, sin pérdida de generalidad, que $K \subset B_X$. Observemos que, si $K \subset X$ es compacto, entonces K es totalmente acotado y separable, y por tanto, $Y := \operatorname{span} K$ es un espacio normado separable. Efectivamente. Si $\{k_n : n \in \mathbb{N}\}$ es un subconjunto denso de K, entonces $D := \left\{\sum_{n \in N} r_n k_n : r_n \in \mathbb{Q},\ N \text{ subconjunto finito de } \mathbb{N}\right\}$ es numerable y denso en Y. Por el lema 2.4.8, existe un conjunto numerable D^* de Y^* que separa los puntos de Y. Consideremos el operador lineal $T : Y \longrightarrow \ell^2$ dado por

$$Ty := \left(\frac{1}{2^{n-1}} f_n(y)\right)_n, \quad y \in Y.$$

Claramente, T es lineal. Como D^* separa los puntos de Y, se tiene que T es inyectivo. Por otra parte, la desigualdad

$$\|Ty\|^2 \leq \frac{4}{3}\|y\|^2$$

nos dice que T es continuo para las normas respectivas de Y y ℓ^2. Con toda esta información, dado que K es compacto, concluimos que $T(K)$ también es compacto, y así, T establece un homeomorfismo entre K y el compacto convexo $T(K)$, véase el corolario 1.1.12, que es, claramente, un subconjunto de Q ya que $K \subset B_Y$. La demostración acaba ahora utilizando el lema 2.4.7. □

STOP Obsérvese que en la demostración del teorema de Schauder se ha obtenido que todo subconjunto compacto convexo de un espacio normado es homeomorfo a un subconjunto compacto convexo del cubo de Hilbert.

Como aplicación del teorema del punto fijo de Schauder se puede demostrar el teorema de Peano sobre existencia de soluciones de una ecuación diferencial ordinaria.

Teorema 2.4.10 (Peano)**.** *Sean $I = [t_0 - h, t_0 + h] \times [x_0 - k, x_0 + k] \subset \mathbb{R}^2$ y $f \in C(I)$ tales que $0 < \|f\|_\infty h \leq k$. Entonces, para cada $x_0 \in \mathbb{R}$, la ecuación*

$$x'(t) = f(t, x)$$

tiene, al menos, una solución en $[t_0 - h, t_0 + h]$ que satisface la condición inicial $x(t_0) = x_0$.

Demostración. Sea B la bola en $C\big([t_0-h,t_0+h]\big)$ con centro $x_0(t)\equiv x_0$ y radio k, es decir,

$$B=\Big\{x\in C\big([t_0-h,t_0+h]\big):\|x-x_0\|_\infty\le k\Big\}.$$

Consideremos la aplicación $\varphi : B\longrightarrow C\big([t_0-h,t_0+h]\big)$ definida mediante la fórmula $z=\varphi(x)$, donde

$$z(t)=x_0+\int_{t_0}^{t} f\big(u,x(u)\big)\,du.$$

Para $z\in\varphi(B)$ se tiene que

$$\|z-x_0\|_\infty\le\|f\|_\infty h\le k, \tag{2.22}$$

y por tanto,

$$\varphi(B)\subset B. \tag{2.23}$$

Si $v,w\in[t_0-h,t_0+h]$, obtenemos que

$$\big|z(v)-z(w)\big|\le\left|\int_{w}^{v} f\big(u,x(u)\big)\,du\right|\le\|f\|_\infty\,|v-w|\,. \tag{2.24}$$

Ahora bien, por (2.22) se deduce que

$$\|z\|_\infty\le|x_0|+k. \tag{2.25}$$

Es sencillo comprobar que (2.24) y (2.25) también se satisfacen para $z\in\operatorname{co}\varphi(B)$. Además, en virtud del teorema de Ascoli-Arzelá, [69, Teorema 17. p. 265], obtenemos que $\operatorname{co}\varphi(B)$ es totalmente acotado y, en consecuencia, $K:=\overline{\operatorname{co}\varphi(B)}$ es compacto. Como B es convexo y cerrado, y $K\subset B$ (por (2.23)), podemos asegurar que

$$\varphi(K)\subset\varphi(B)\subset\overline{\operatorname{co}\varphi(B)}=K.$$

Vamos a demostrar ahora que φ es continua respecto a $\|\cdot\|_\infty$. Tomamos $\varepsilon>0$. Al ser f uniformemente continua en I, existe $\delta>0$ tal que, si $|x_1-x_2|<\delta$, entonces $\big|f(t,x_1)-f(t,x_2)\big|<\varepsilon$, para cada $t\in[t_0-h,t_0+h]$. Por lo tanto, si $x_1,x_2\in B$ y $\|x_1-x_2\|_\infty<\delta$, tendremos que, para todo $t\in[t_0-h,t_0+h]$,

$$\left|\int_{t_0}^{t} f\big(u,x_1(u)\big)\,du-\int_{t_0}^{t} f\big(u,x_2(u)\big)\,du\right|<\varepsilon h.$$

Luego $\big\|\varphi(x_1)-\varphi(x_2)\big\|_\infty<\varepsilon h$. Finalmente, gracias al teorema de Schauder 2.4.9, podemos afirmar que φ tiene un punto fijo $x\in K$. Luego

$$x(t)=x_0+\int_{t_0}^{t} f\big(u,x(u)\big)\,du.$$

Esto implica que $x(t_0)=x_0$, y que

$$x'(t)=f\big(t,x(t)\big),$$

lo que concluye la demostración. $\square$

Las ideas utilizadas en la prueba del teorema de Peano sirven para demostrar la siguiente versión global: *Sea $f : [a,b] \times \mathbb{R} \longrightarrow \mathbb{R}$ una función continua y acotada. Entonces, para cada $x_0 \in \mathbb{R}$, la ecuación $x'(t) = f(t,x)$ tiene, al menos, una solución en $[a,b]$ que satisface la condición inicial $x(t_0) = x_0$.* La demostración es la misma que la del teorema 2.4.10: para establecer la continuidad de φ para $\|\cdot\|_\infty$ en $x_1 \in B$, basta utilizar la continuidad uniforme de f en el compacto $[a,b] \times K$, donde $K = x_1\big([a,b]\big) + [-1,1]$.

El teorema del punto fijo también es cierto para subconjuntos convexos compactos de un espacio localmente convexo, véase el teorema 2.4.15. La demostración es, sin embargo, más elaborada que la que hemos realizado para el teorema de Schauder, y requiere de algún trabajo preliminar que hacemos a continuación.

Determinación de conjuntos funcionales respecto de funciones Consideremos E un espacio localmente convexo, K un subconjunto de E, $\varphi : K \longrightarrow K$ una función, y F y G dos subconjuntos de E'. Se dice que G *determina a* F *respecto de* φ si, para cada $f \in F$ y cada $\varepsilon > 0$, existen $\delta > 0$ y $A \subset G$ finito tales que, si $x, y \in K$ con

$$x - y \in V(0, A, \delta) = \big\{z \in E : |g(z)| < \delta,\ g \in A\big\},$$

entonces

$$\big|f\big(\varphi(x)\big) - f\big(\varphi(y)\big)\big| < \varepsilon.$$

El lema que sigue es consecuencia inmediata de la definición anterior.

Lema 2.4.11. *Sean E un espacio localmente convexo, K un subconjunto de E, $\varphi : K \longrightarrow K$ una función, y F y G dos subconjuntos de E'. Si G determina a F respecto de φ, y para $x_0, y_0 \in K$ se tiene que*

$$g(x_0) = g(y_0) \ \text{ para cada } \ g \in G,$$

entonces, para cada $f \in F$,

$$f\big(\varphi(x_0)\big) = f\big(\varphi(y_0)\big).$$

Lema 2.4.12. *Sean E un espacio localmente convexo y K un subconjunto $\sigma(E,E')$-compacto de E. Si $\eta : K \longrightarrow \mathbb{R}$ es una aplicación $\sigma(E,E')$-continua, entonces, para cada $\varepsilon > 0$, existen un conjunto finito $A \subset E'$ y $\delta > 0$ tales que*

$$\big|\eta(x) - \eta(y)\big| < \varepsilon,$$

para todo $x, y \in K$ que satisfacen $x - y \in V(0, A, \delta)$.

Demostración. Como η es $\sigma(E,E')$-continua, para cada $x \in K$ existen un conjunto finito $A_x \subset E'$ y $\delta_x > 0$ tales que $\big|\eta(y) - \eta(x)\big| < \varepsilon/2$, si $y \in \big(x + V(0, A_x, \delta_x)\big) \cap K$. La inclusión

$$K \subset \bigcup_{x \in K} \Big(x + V\big(0, A_x, \tfrac{\delta_x}{2}\big)\Big),$$

y el hecho de que K es compacto y cada $x+V(0,A_x,\delta_x/2)$ es abierto en $\big(E,\sigma(E,E')\big)$, nos permite determinar $x_1,\dots,x_n\in K$ tales que

$$K\subset\bigcup_{i=1}^{n}\Big(x_i+V\big(0,A_{x_i},\tfrac{\delta_{x_i}}{2}\big)\Big). \tag{2.26}$$

Definimos ahora $A:=\bigcup_{i=1}^n A_{x_i}$ y $\delta:=\text{mín}_{1\le i\le n}\delta_{x_i}/2$. Veamos, para terminar, que A y δ satisfacen la condición exigida en el lema. Fijamos $x,y\in K$ con $x-y\in V(0,A,\delta)$. Gracias a la inclusión (2.26), existe un x_i tal que $x\in\big(x_i+V(0,A_{x_i},\delta_{x_i}/2)\big)$, es decir, $\big|g(x-x_i)\big|<\delta_{x_i}/2$, para cada $g\in A_{x_i}$. Como $\big|g(x-y)\big|<\delta$ si $g\in A$, $A_{x_i}\subset A$ y $\delta\le\delta_{x_i}/2$, se tiene que

$$\big|g(y-x_i)\big|\le\big|g(y-x)\big|+\big|g(x-x_i)\big|<\delta+\frac{\delta_{x_i}}{2}\le\frac{\delta_{x_i}}{2}+\frac{\delta_{x_i}}{2}=\delta_{x_i},$$

para cada $g\in A_{x_i}$, y así, concluimos que $y\in x_i+V(0,A_{x_i},\delta_{x_i})$. Por la definición de $V(0,A_{x_i},\delta_{x_i})$ obtenemos que

$$\big|\eta(x)-\eta(x_i)\big|<\frac{\varepsilon}{2}\quad\text{y}\quad\big|\eta(y)-\eta(x_i)\big|<\frac{\varepsilon}{2}.$$

Finalmente,

$$\big|\eta(x)-\eta(y)\big|<\big|\eta(x)-\eta(x_i)\big|+\big|\eta(x_i)-\eta(y)\big|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,$$

y la demostración termina. □

STOP Todo espacio localmente convexo (E,τ) es un espacio uniforme en el sentido de [69, p. 204]. Si $\mathscr{U}$ es una base de τ-entornos del origen en E, y para cada $U\in\mathscr{U}$ definimos

$$N_U:=\big\{(x,y)\in E\times E: x-y\in U\big\},$$

entonces $\{N_U:U\in\mathscr{U}\}$ es base para una uniformidad en E cuya topología asociada es τ. El lema anterior 2.4.12 es un caso particular de que, en un espacio uniforme, las funciones continuas en los compactos son uniformemente continuas, véase [69, Teorema 31, p. 226].

Lema 2.4.13. *Sean E un espacio localmente convexo y $K\subset E$ un conjunto $\sigma(E,E')$-compacto. Si $\varphi:K\longrightarrow K$ es una aplicación $\sigma(E,E')$-continua y $f\in E'$, entonces existe $G\subset E'$ numerable, con $f\in G$, de forma que G se determina a sí mismo respecto de φ.*

Demostración. Etapa 1. Definimos $F_0:=\{f\}$. Como $f\circ\varphi$ es $\sigma(E,E')$-continua, en virtud del lema 2.4.12, existen $\delta_1>0$ y $G_1\subset E'$ finito tales que

$$\big|(f\circ\varphi)(x)-(f\circ\varphi)(y)\big|<1,\ \text{ si }\ x-y\in V(0,G_1,\delta_1).$$

Etapa 2. Definimos $F_1:=F_0\cup G_1$. Aplicamos de nuevo el lema 2.4.12 y encontramos $G_2\subset E'$ finito y $\delta_2>0$ tales que

$$\big|(h\circ\varphi)(x)-(h\circ\varphi)(y)\big|<\frac{1}{2},\ \text{ si }\ x-y\in V(0,G_2,\delta_2)\ \text{ y }\ h\in F_1.$$

Reiterando este procedimiento, el conjunto $G:=\bigcup_{k=0}^{\infty}F_k$ tiene las propiedades requeridas. □

Proposición 2.4.14. *Sean $(E,\mathfrak{T})$ un espacio localmente convexo, $K \subset E$ un conjunto $\mathfrak{T}$-compacto y convexo, y $\varphi : K \longrightarrow K$ una aplicación continua. Si K tiene, al menos, dos puntos distintos, entonces existe un subconjunto $K_0 \subset K$, propio y cerrado, tal que $\varphi(K_0) \subset K_0$.*

Demostración. Como K es $\mathfrak{T}$-compacto, las topologías débil y $\mathfrak{T}$ restringidas a K coinciden (véase el corolario 1.1.13). Así pues, φ es $\sigma(E,E')$-continua. Tomemos dos puntos $x_1, x_2 \in K$, $x_1 \neq x_2$, y fijemos $f \in E'$ con $f(x_1) \neq f(x_2)$. Por el lema 2.4.13, existe $G = \{g_i\}_{i=1}^{\infty} \subset E'$ tal que $f \in G$ y G se determina a sí mismo respecto de φ. Como $\sup_{x\in K}\big|g_i(x)\big| < \infty$ para cada $i \in \mathbb{N}$, podemos suponer, sin pérdida de generalidad, que

$$\sup_{x\in K}\big|g_i(x)\big| < \frac{1}{2^{i-1}}, \tag{2.27}$$

para cada $i \in \mathbb{N}$. Definimos $Y := \operatorname{span} K$ y $T : Y \longrightarrow \ell^2$ el operador lineal dado por $Ty = \big\{g_i(y)\big\}_{i=1}^{\infty}$. Escribamos $L = T(K)$. Se tienen las siguientes propiedades:

– *L es convexo y está contenido en Q:* Esto es consecuencia de que T es lineal. El hecho de que $T(K) \subset Q$ se sigue de la definición de T y de las desigualdades (2.27).

– *$T|_K$ es $\sigma(Y,E')$-$\|\cdot\|$-continua:* Fijamos $\varepsilon > 0$. Sea $N \in \mathbb{N}$ tal que $\frac{1}{4^{N-3}} < \varepsilon^2$. Sean $x, y \in K$ tales que $\big|g_i(x-y)\big| < \varepsilon/\sqrt{2N}$, para cada $i = 1,\dots,N$. Entonces, considerando de nuevo las proyecciones definidas en la página 87 por la expresión $P_n x = \sum_{i=1}^{n} x_i e_i$, $n \in \mathbb{N}$, se tiene que $\big\|P_N T(x-y)\big\| \leq \varepsilon^2/2$. Así, usando la desigualdad (2.21) del lema 2.4.3 obtenemos que

$$\|Tx - Ty\|^2 = \big\|T(x-y)\big\|^2 \leq \big\|P_N T(x-y)\big\|^2 + \frac{\varepsilon^2}{3} < \frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{2} = \varepsilon^2,$$

lo que nos asegura que $T|_K$ es $\sigma(Y,E')$-$\|\cdot\|$-continua.

– *L es compacto:* $L = T(K)$ es compacto en ℓ^2 como consecuencia del apartado anterior.

– *$\psi = T\circ\varphi\circ T^{-1}$ está bien definida:* Tomamos $u \in L$ y $x, y \in T^{-1}(u)$. Entonces $Tx = Ty = u$, es decir, $g_i(x) = g_i(y)$ para cada $i \in \mathbb{N}$. Como G se determina a sí mismo respecto de φ, por el lema 2.4.11 podemos asegurar que $g_i\big(\varphi(x)\big) = g_i\big(\varphi(y)\big)$, para cada $i \in \mathbb{N}$, es decir, se tiene que $(T\circ\varphi)(x) = (T\circ\varphi)(y)$. De aquí se deduce que la aplicación $\psi : L \longrightarrow L$ dada por $\psi(u) = (T\circ\varphi\circ T^{-1})(u)$ está bien definida.

– *ψ es continua:* Fijamos $\varepsilon > 0$. Sea $N \in \mathbb{N}$ tal que $\frac{1}{4^{N-3}} < \varepsilon^2$. Como G se determina a sí mismo, podemos encontrar $m \in \mathbb{N}$ y $\delta > 0$ tales que, si $w, z \in K$ satisfacen

$$\big|g_i(w) - g_i(z)\big| < \delta, \text{ para cada } i = 1,\dots,m,$$

entonces

$$\Big|g_i\big(\varphi(w)\big) - g_i\big(\varphi(z)\big)\Big| < \frac{\varepsilon}{\sqrt{2N}}, \text{ para cada } i = 1,\dots,N. \tag{2.28}$$

Vamos a probar que, si $u, v \in L$ son tales que $\|u - v\| < \delta$, entonces $\big\|\psi(u) - \psi(v)\big\| < \varepsilon$. Efectivamente, como $u, v \in L = T(K)$, existen $x, y \in K$ tales que $u = Tx$ y $v = Ty$, es decir,

$u = \big(g_i(x)\big)_i$ y $v = \big(g_i(y)\big)_i$. Entonces, se tiene que

$$\big|g_i(x) - g_i(y)\big| \leq \left(\sum_{j=1}^{\infty} \big|g_j(x) - g_j(y)\big|^2\right)^{1/2} = \|u - v\| < \delta, \text{ para } i = 1, 2, \ldots$$

Así, por (2.28) se sigue que

$$\Big|g_i\big(\varphi(x)\big) - g_i\big(\varphi(y)\big)\Big| < \frac{\varepsilon}{\sqrt{2N}}, \text{ para } i = 1, 2, \ldots, N.$$

Ahora, como $\psi(u) = (T \circ \varphi)(x) = \big(g_i(\varphi(x))\big)_i$ y $\psi(v) = (T \circ \varphi)(y) = \big(g_i(\varphi(y))\big)_i$, los cálculos anteriores y la desigualdad (2.21) nos permiten concluir que

$$\begin{aligned} \big\|\psi(u) - \psi(v)\big\|^2 &= \Big\|T\big(\varphi(x) - \varphi(y)\big)\Big\|^2 \leq \Big\|P_N T\big(\varphi(x) - \varphi(y)\big)\Big\|^2 + \frac{\varepsilon^2}{3} \\ &= \sum_{i=1}^{N} g_i\big(\varphi(x) - \varphi(y)\big)^2 + \frac{\varepsilon^2}{3} < N\frac{\varepsilon^2}{2N} + \frac{\varepsilon^2}{2} = \varepsilon^2. \end{aligned}$$

Como L es un subconjunto cerrado y convexo del cubo de Hilbert, el lema 2.4.7 nos asegura que ψ tiene un punto fijo. Sea $u_0 = \psi(u_0)$. Tomamos $K_0 = T^{-1}(u_0)$. Vamos a ver que este K_0 es el conjunto que andábamos buscando. Dado $x \in K_0 = T^{-1}(u_0)$, como $\psi = T \circ \varphi \circ T^{-1}$, tenemos que $\varphi(x) \in K_0$. Hemos demostrado así que $\varphi(K_0) \subset K_0$. Veamos, por último, que K_0 es un subconjunto propio de K. Si $K_0 = K$, entonces los puntos x_1, x_2 fijados al principio de la demostración pertenecerían a $K_0 = T^{-1}(u_0)$, de donde se deduciría que $Tx_1 = Tx_2$. Pero esto implicaría que $f(x_1) = f(x_2)$, pues $f \in G$, contradicción que concluye la prueba. □

Teorema 2.4.15 (Schauder-Tychonoff)**.** *Sean E un espacio localmente convexo, $K \subset E$ convexo y compacto, y $\varphi : K \longrightarrow K$ una aplicación continua. Entonces, existe $x \in K$ tal que $\varphi(x) = x$.*

Demostración. Sea $\mathscr{K}$ la familia de todos los subconjuntos C de K, convexos y compactos, tales que $\varphi(C) \subset C$. Introducimos en $\mathscr{K}$ un orden: $C_1 \preceq C_2$ si, y sólo si, $C_1 \supset C_2$. Como, para cada subfamilia de $\mathscr{K}$ totalmente ordenada existe un elemento maximal, el Lema de Zorn nos asegura la existencia de un elemento maximal de $\mathscr{K}$. Aplicando ahora la proposición 2.4.14, obtenemos que el elemento maximal contiene un único punto. □

2.5 El teorema de Lomonosov

EL resultado central de esta sección, lema 2.5.7 y teorema 2.5.8, que se debe a Lomonosov, [75], implica, en particular, que *cualquier operador compacto en un espacio de Banach tiene un subespacio invariante no trivial*. Cuando este resultado apareció, causó un gran impacto, tanto por la fuerza de su conclusión como por la simplicidad de su prueba. La demostración presentada aquí utiliza el teorema del punto fijo de Schauder 2.4.9, en su versión para espacios normados.

$\boxed{\textit{Espacio de operadores acotados}}$ En lo que sigue, $\mathscr{L}(X)$ denotará el espacio de operadores acotados del espacio de Banach X en sí mismo. Un operador $T \in \mathscr{L}(X)$ se dice compacto si $T(B_X)$ es relativamente compacto para la topología de la norma: $\mathscr{K}(X)$ denota el espacio de operadores compactos de X en si mismo.

$\boxed{\textit{Subespacios invariantes}}$ Si X es un espacio de Banach y $T \in \mathscr{L}(X)$, un *subespacio invariante* para T es un subespacio lineal cerrado Y de X tal que $Tx \in Y$ para $x \in Y$. Y es *no trivial* si $Y \neq \{0\}, X$. Representaremos por $\operatorname{Lat} T$ la familia de todos los subespacios invariantes para T. Si $\mathscr{A} \subset \mathscr{L}(X)$, entonces $\operatorname{Lat}\mathscr{A} = \bigcap\{\operatorname{Lat} T : T \in \mathscr{A}\}$.

La siguiente proposición, cuya demostración se deja como ejercicio, justifica la utilización del símbolo Lat para representar la familia de los subespacios invariantes: con las operaciones $\vee$ y $\wedge$, $\operatorname{Lat} T$ es un retículo *(i)* que es completo *(ii)*. Más aún, $\operatorname{Lat} T$ tiene un mayor elemento, X, y un menor elemento, $\{0\}$.

Proposición 2.5.1. *Se verifican las siguientes propiedades:*

(i) *Si* $Y_1, Y_2 \in \operatorname{Lat} T$, *entonces* $Y_1 \vee Y_2 := \overline{Y_1 + Y_2} \in \operatorname{Lat} T$, *e* $Y_1 \wedge Y_2 := Y_1 \cap Y_2 \in \operatorname{Lat} T$.

(ii) *Si* $\{Y_i : i \in I\} \subset \operatorname{Lat} T$, *entonces* $\bigvee\{Y_i : i \in I\}$, *el espacio lineal cerrado generado por* $\bigcup_i Y_i$ *y* $\bigwedge\{Y_i : i \in I\} \equiv \bigcap_i Y_i$ *pertenecen a* $\operatorname{Lat} T$.

La pregunta principal que se plantea es la siguiente: ¿tiene $\operatorname{Lat} T$ algún elemento además de los evidentes $\{0\}$ y X? O, en otras palabras, ¿tiene T un subespacio invariante no trivial? Una primera respuesta obvia es la siguiente: si X es un espacio de Banach no separable, todo operador acotado $T \in \mathscr{L}(X)$ tiene un subespacio invariante no trivial; a saber: para cualquier $x \neq 0$ en X, el subespacio $Y := \overline{\operatorname{span}\{T^n x : n \in \mathbb{N}\}}$ es invariante, no nulo y distinto de X dado que es separable. Otra respuesta conocida se tiene cuando $X = H$ es un espacio de Hilbert complejo y $T \in \mathscr{L}(H)$, $T \neq 0$, es un operador compacto y normal; en tal caso, T tiene un valor propio no nulo α y, consecuentemente, el espacio propio (finito-dimensional) $\ker(T - \alpha I) \neq \{0\}$ es invariante por T. Como cada subespacio de $\ker(T - \alpha I) \neq \{0\}$ es también invariante por T, se deduce que T tiene subespacios invariantes no triviales.

Ejemplo 2.5.2. Si X es un espacio finito-dimensional sobre $\mathbb{C}$ y $T \in \mathscr{L}(X)$, entonces $\operatorname{Lat} T$ es no trivial. En efecto, sea $X = \mathbb{C}^n$ y sea T una matriz. Entonces, $p(z) = \det(T - zI)$ es un polinomio de grado n, y por tanto, tiene un cero, digamos α. Si $\det(T - \alpha I) = 0$, entonces $T - \alpha I$ es no invertible. Pero en espacios finito-dimensionales esto significa que $T - \alpha I$ no es inyectiva. En consecuencia, $\ker(T - \alpha I) \neq \{0\}$. Sea $Y \subset \ker(T - \alpha I)$, tal que $Y \neq \{0\}$. Si $x \in Y$, entonces $Tx = \alpha x \in Y$ y, por tanto, $Y \in \operatorname{Lat} T$. □

Ejemplo 2.5.3. Si $T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ en $\mathbb{R}^2$, entonces $\operatorname{Lat} T$ es trivial. En efecto, si $\operatorname{Lat} T$ no fuese trivial, existiría un espacio uno-dimensional Y en $\operatorname{Lat} T$. Supongamos que $Y = \{\alpha e : \alpha \in \mathbb{R}\}$. Dado que $Y \in \operatorname{Lat} T$, $Te = \lambda e$ para algún λ en $\mathbb{R}$. Por tanto, $T^2 e = T(Te) = \lambda Te = \lambda^2 e$. Pero $T^2 = -I$ y, en consecuencia, $-e = \lambda^2 e$, de donde se deduciría que $\lambda^2 = -1$ si $e \neq 0$, lo cual no puede ocurrir si λ es real. Sin embargo, en el caso de que $n \geq 3$ y $T \in \mathscr{L}(\mathbb{R}^n)$, entonces $\operatorname{Lat} T$ es no trivial (véase [36, §4, ejercicio 6]). $\square$

Ejemplo 2.5.4. Si $V : L^2([0,1]) \longrightarrow L^2([0,1])$ es el operador de Volterra, $Vf(x) = \displaystyle\int_0^x f(t)\,dt$ y $0 \leq \alpha \leq 1$, sea

$$Y_\alpha = \Big\{ f \in L^2([0,1]) : f(t) = 0 \ \text{ para } \ 0 \leq t \leq \alpha \Big\}.$$

Entonces, $Y_\alpha \in \operatorname{Lat} V$. Más aún, puede demostrarse que $\operatorname{Lat} V = \{Y_\alpha : 0 \leq \alpha \leq 1\}$ (véanse [67] y [88, p. 68]). $\square$

Ejemplo 2.5.5. Si $S : \ell^p \longrightarrow \ell^p$ es el operador desplazamiento $S(\alpha_1, \alpha_2, \ldots) = (0, \alpha_1, \alpha_2, \ldots)$, e

$$Y_n = \Big\{ x \in \ell^p : x(k) = 0 \ \text{ para } \ 1 \leq k \leq n \Big\},$$

entonces $Y_n \in \operatorname{Lat} S$. $\square$

Ejemplo 2.5.6. Sea (Ω, Σ, μ) un espacio de medida σ-finito y, para $\phi \in L^\infty(\mu)$, representaremos por M_ϕ el operador multiplicación sobre $L^p(\mu)$, $1 \leq p \leq \infty$. Si $A \in \Sigma$, sea

$$Y_A = \Big\{ f \in L^p(\mu) : f(w) = 0 \ \ \mu\text{–p.c.t. } \ w \in \Omega \setminus A \Big\}.$$

Entonces, $Y_A \in \operatorname{Lat} M_\phi$. $\square$

Es una tarea difícil determinar todos los subespacios invariantes de un operador específico. El operador de Volterra y el operador desplazamiento son ejemplos donde esto es posible. Sin embargo, hay operadores para los que no existe ninguna caracterización de sus subespacios invariantes: por ejemplo, sea μ la medida de Lebesgue sobre el disco unidad abierto D, y sea $Af(z) = zf(z)$, para $f \in L^2(\mu)$; no existe ninguna caracterización conocida de $\operatorname{Lat} A$.

Lema 2.5.7 (Lomonosov, 1973)**.** *Si $\mathscr{A}$ es una subálgebra de $\mathscr{L}(X)$ de tal forma que $I \in \mathscr{A}$ y $\operatorname{Lat} \mathscr{A} = \big\{ \{0\}, X \big\}$, y si K es un operador compacto no nulo sobre X, entonces existe un $A \in \mathscr{A}$ tal que* $\ker(AK - I) \neq \{0\}$.

Demostración. Podemos suponer que $\|K\| = 1$. Fijamos un x_0 en X de modo que $\|Kx_0\| > 1$, y tomamos $S = \big\{ x \in X : \|x - x_0\| \leq 1 \big\}$. Entonces,

$$0 \notin \overline{K(S)}. \tag{2.29}$$

Efectivamente, para cada $x \in S$, se tiene la cadena de desigualdades

$$\|Kx\| \geq \|Kx_0\| - \big\|K(x - x_0)\big\| \geq \|Kx_0\| - \|x - x_0\| \geq \|Kx_0\| - 1 > 0.$$

Ahora, si $y \in X$ y $y \neq 0$, $\overline{\{Ty : T \in \mathscr{A}\}}$ es un subespacio invariante para $\mathscr{A}$ (dado que $\mathscr{A}$ es un álgebra) que contiene al vector no nulo y (pues $I \in \mathscr{A}$). Por hipótesis, $\overline{\{Ty : T \in \mathscr{A}\}} = X$.

Pero, por (2.29), esto nos dice que, para todo $y \in \overline{K(S)}$, existe un $T \in \mathscr{A}$ con $\|Ty - x_0\| < 1$. Equivalentemente,

$$\overline{K(S)} \subset \bigcup_{T \in \mathscr{A}} \big\{y \in X : \|Ty - x_0\| < 1\big\}.$$

Debido a que $\overline{K(S)}$ es compacto, existen $T_1, \dots, T_n$ en $\mathscr{A}$ tales que

$$\overline{K(S)} \subset \bigcup_{j=1}^{n} \big\{y \in X : \|T_j y - x_0\| < 1\big\}. \tag{2.30}$$

Dados $y \in \overline{K(S)}$ y $1 \leq j \leq n$, sea $a_j(y) = \text{máx}\big\{0, 1 - \|T_j y - x_0\|\big\}$. Por (2.30), se tiene que $\sum_{j=1}^n a_j(y) > 0$, para todo $y \in \overline{K(S)}$. Definimos entonces $b_j : \overline{K(S)} \longrightarrow \mathbb{R}$ por

$$b_j(y) := \frac{a_j(y)}{\sum_{i=1}^n a_i(y)},$$

y definimos $\psi : S \longrightarrow X$ como

$$\psi(x) := \sum_{j=1}^{n} b_j(Kx) T_j Kx.$$

Es fácil ver que $a_j : \overline{K(S)} \longrightarrow [0,1]$ es una función continua. Por lo tanto, b_j y ψ son también continuas.

Si $x \in S$, entonces $Kx \in K(S)$. Si $b_j(Kx) > 0$, $a_j(Kx) > 0$, y por tanto, $\|T_j Kx - x_0\| < 1$. Esto es, $T_j Kx \in S$ siempre que $b_j(Kx) > 0$. Ya que S es un conjunto convexo y $\sum_{j=1}^n b_j(Kx) = 1$ para $x \in S$, entonces

$$\psi(S) \subset S.$$

Obsérvese que $T_j K$ es un operador compacto para cada j, por lo que $\overline{T_j K(S)}$ es convexo y compacto. La envoltura convexa $\text{co}\big(\bigcup_{j=1}^n \overline{T_j K(S)}\big)$ es compacta gracias al lema 3.1.7. Como este conjunto convexo contiene a $\psi(S)$, se deduce que $\overline{\text{co}(\psi(S))} \subset S$ (S es cerrado y convexo) es también compacto. Obsérvese que, si escribimos $S_0 = \overline{\text{co}(\psi(S))}$, entonces $\psi(S_0) \subset \psi(S) \subset S_0$. Utilizando ahora el teorema del punto fijo de Schauder 2.4.9, existe un vector x_1 en S tal que $\psi(x_1) = x_1$.

Sea $\beta_j := b_j(Kx_1)$, y consideremos $A := \sum_{j=1}^n \beta_j T_j$. Así, $A \in \mathscr{A}$ y $AKx_1 = \psi(x_1) = x_1$. Dado que $x_1 \neq 0$ por (2.29), $\ker(AK - I) \neq \{0\}$. □

$\boxed{\text{Subespacio hiperinvariante}}$ Si $T \in \mathscr{L}(X)$, un *subespacio hiperinvariante* para T es un subespacio Y de X tal que $AY \subset Y$ para todo operador A en el centralizador $C_{\mathscr{L}(X)}(T)$ de T, es decir, $AY \subset Y$ cuando $AT = TA$. Obsérvese que cualquier subespacio hiperinvariante para T es invariante.

Teorema 2.5.8 (Lomonosov)**.** *Si X es un espacio de Banach sobre $\mathbb{C}$, $T \in \mathscr{L}(X)$ no es un múltiplo de la identidad y $TK = KT$ para algún operador compacto no nulo K, entonces T tiene un subespacio hiperinvariante no trivial.*

Demostración. Sea $\mathscr{A} = C_{\mathscr{L}(X)}(T)$. Queremos demostrar que $\mathrm{Lat}\mathscr{A} \neq \{\{0\}, X\}$. Si esto no es así, el lema de Lomonosov 2.5.7 implica la existencia de un operador $A \in \mathscr{A}$ para el que se tiene $Y = \ker(AK - I) \neq \{0\}$. Pero $Y \in \mathrm{Lat}(AK)$, y la restricción $AK|_Y$ es el operador identidad. Como $AK \in \mathscr{K}(X)$, entonces $AK\big|_Y \in \mathscr{K}(Y)$ y, por tanto, $\dim Y < \infty$. Dado que $AK \in \mathscr{A} = C_{\mathscr{L}(X)}(T)$, para todo $y \in Y$ se tiene que $AK(Ty) = T(AKy) = Ty$; en consecuencia, $Ty \in \ker(AK - I)$, y así, $TY \subset Y$. Pero al ser $\dim Y < \infty$, $T|_Y$ debe tener un valor propio λ. Así, $\ker(T - \lambda I) = Z \neq \{0\}$. Ahora bien, $Z \neq X$, ya que T no es un múltiplo de la identidad. Es fácil ver entonces que Z es hiperinvariante para T. Efectivamente, si tomamos $B \in C_{\mathscr{L}(X)}(T)$ y $z \in Z$, entonces se tiene que $T(Bz) = B(Tz) = B(\lambda z) = \lambda Bz$, y por tanto $Bz \in Z$. $\square$

Observemos que, en el teorema anterior, los espacios de Banach tienen que ser necesariamente complejos: el ejemplo 2.5.3 muestra que, en efecto, para espacios de Banach reales el teorema anterior no es cierto.

Corolario 2.5.9 (Aronszajn-Smith, 1954, [3])**.** *Sea X un espacio de Banach complejo y sea $K : X \longrightarrow X$ un operador compacto. Entonces,* $\mathrm{Lat}K$ *es no trivial.*

El siguiente resultado apareció en [7], donde se demuestra utilizando técnicas de análisis no-estándar. Halmos, [59], dio una prueba en la que empleaba análisis estándar. Ahora, es una sencilla consecuencia del teorema de Lomonosov.

Corolario 2.5.10. *Si X es infinito-dimensional, $A \in \mathscr{L}(X)$ y existe un polinomio en una variable, p, tal que $p(A)$ es compacto, entonces* $\mathrm{Lat}A$ *es no trivial.*

Demostración. Si $p(A) \neq 0$, entonces se aplica el teorema de Lomonosov. Si $p(A) = 0$, sea

$$p(z) = \alpha_0 + \alpha_1 z + \ldots + \alpha_n z^n, \qquad \alpha_n \neq 0.$$

Para $x \neq 0$, sea $Y = \mathrm{span}\{x, Ax, \ldots, A^{n-1}x\}$. Dado que

$$A^n = -\alpha_n^{-1}\left[\alpha_0 + \alpha_1 A + \ldots + \alpha_{n-1}A^{n-1}\right],$$

$Y \in \mathrm{Lat}A$. Además, como $x \in Y$, $Y \neq \{0\}$; finalmente, al ser $\dim Y < \infty$, $Y \neq X$. $\square$

Corolario 2.5.11. *Si $K_1, K_2 \in \mathscr{K}(X)$ y $K_1K_2 = K_2K_1$, entonces K_1 y K_2 tienen un subespacio invariante no trivial común a ambos.*

Demostración. Si $K_2 = 0$ el resultado es claro. Si $K_2 \neq 0$ el teorema 2.5.8 nos asegura la existencia de un subespacio hiperinvariante no trivial para K_1; por definición de hiperinvariante, este subespacio es invariante para K_2. $\square$

Existencia de subespacios invariantes, V. Lomonosov, 1951-1987.

«La mejor aplicación del teorema de Schauder del punto fijo que puedo pensar es la prueba del teorema de Lomonosov de existencia de subespacios invariantes. Hay varias generalizaciones de este resultado, pero lo que Lomonosov demostró fue que: si X es un espacio de Banach y T un operador acotado en X tal que $TK = KT$ para algún operador compacto K no nulo de X en X, entonces algún subespacio Y propio de X es invariante bajo T, i.e., $T(Y) \subset Y$.

*Un poco de historia sobre este resultado. Parece ser que J. von Neumann demostró que todo operador compacto en un espacio de Hilbert tiene un subespacio invariante no trivial, pero no publicó la prueba. En 1951 ó 1952, N. Aronszajn, en Kansas, probó este resultado independientemente de von Neumann y mostró su prueba a K. T. Smith. Esa noche, Smith fue a casa y trató de reproducir la demostración de Aronszajn, pero encontró otra diferente que también servía para espacios de Banach (conozco esto porque en aquellos tiempos yo era un estudiante en la Universidad de Kansas). Aronszajn y Smith escribieron un artículo juntos. Al final del artículo, dejaron una cuestión abierta: si un operador T tiene cuadrado compacto, ¿tiene T un subespacio invariante no trivial? Mucho después, creo que en 1963, A. R. Bernstein y A. Robinson demostraron, con técnicas de análisis no-estándar, que si T es un operador con la propiedad de que $p(T)$ es compacto, donde p es un polinomio de grado positivo, entonces T tiene un subespacio invariante propio. Esto fue demostrado originalmente para espacios de Hilbert, pero pronto fue generalizado a espacios de Banach usando otra vez técnicas de análisis no-estándar. Date cuenta de que el teorema de Lomonosov generaliza todo lo de arriba, y la demostración es ¡*tan bella y tan simple*! Recuerdo que cuando Lomonosov demostró su resultado, J. Lindenstrauss nos visitó en la Universidad de Washington desde Berkeley. Sacó del bolsillo una hoja de papel arrugada (¡una sólo!), que era la fotocopia de la prueba completa manuscrita que el propio Lomonosov había hecho de su resultado. Joram dijo: "esta demostración está propiciando que mucha gente en Berkeley sea infeliz". Leí la prueba una vez, y* no la podré olvidar nunca.

Demostración. Probaremos que si T es un operador acotado en un espacio de Banach infinito-dimensional X que conmuta con algún operador compacto no nulo K, entonces T tiene un subespacio invariante propio.

Supongamos que $\|K\| = 1$ y fijemos un punto a de forma que $\|K(a)\| > 1$. Sean $S = a + B_X$ y $C = \overline{K(S)}$. Entonces, C es compacto en norma, convexo y $0 \notin C$. Observemos que si $\{P(T)x : P \text{ polinomio}\}$ no es denso para algún elemento no nulo x en X, entonces su clausura es un subespacio invariante propio de T. Así, podemos suponer que $\{P(T)x : P \text{ polinomio}\}$ es denso en X para cada elemento no nulo $x \in X$. Entonces, para cada $x \in C$, existe un polinomio P_x tal que $\|a - P_x(T)(x)\| < 1$. Definimos $U_x = \{y \in X : \|a - P_x(T)(y)\| < 1\}$. Cada U_x es un entorno abierto de x, y así, por compacidad, existen $x_1, x_2, \ldots, x_n$ de forma que C está recubierto por $U_1, \ldots, U_n$, donde $U_i = U_{x_i}$. Sea $b_1, \ldots, b_n$ una partición de la unidad en C subordinada al cubrimiento $U_1, \ldots, U_n$, y consideremos $F : C \longrightarrow C$ la aplicación continua dada por

$$F(x) = \sum_{i=1}^{n} b_i(x) K P_i(T)(x)$$

(aquí, $P_i = P_{x_i}$). Por el teorema de Schauder, existe un punto $c \in C$ tal que $F(c) = c$. Sea H el operador lineal en X dado por

$$H = \sum_{i=1}^{n} b_i(c) K P_i(T),$$

y sea $H_1 = \{z \in X : H(z) = z\}$ el auto-espacio asociado al autovalor 1. Dado que H conmuta con T, H_1 es un subespacio invariante para T. Obsérvese que H_1 no es nulo, puesto que $c \in H_1$. Por otro lado, si tuviéramos que $H_1 = X$, H sería la identidad en X y, como H es compacto, entonces X sería de dimensión finita, que no es el caso. Consecuentemente, H_1 es propio y la prueba termina. $\square$

I. Namioka (2 de enero de 2004)»

C. J. Read, en 1984, [90], demostró la existencia de un espacio de Banach y de un operador sobre dicho espacio de Banach que no tenía ningún subespacio invariante no trivial, lo que estuvo precedido por algún trabajo de P. Enflo (no publicado, aunque conocido) en el que se resolvía esta misma cuestión. Más tarde, B. Beauzamy, en 1985, [6], recogió y arregló las ideas de Enflo y dio una exposición y simplificación de su construcción. El trabajo de Enflo apareció finalmente en 1987, [47]. Read, en 1986, [91], presentó una exposición autocontenida demostrando la existencia de un operador acotado en ℓ^1 que no tenía ningún subespacio invariante no trivial.

Pero todo este profundo trabajo no cierra completamente el tema. ¿Qué espacios de Banach X tienen la propiedad de que existe un operador acotado sobre X con ningún subespacio invariante no trivial? Si X es reflexivo, ¿es $\operatorname{Lat} T$ no trivial, para todo T en $\mathscr{L}(X)$? Esta pregunta no tiene respuesta incluso si X es un espacio de Hilbert.

Una demostración de una versión ligeramente más débil del teorema de Lomonosov debida a A. J. Michaels, que elude el teorema del punto fijo de Schauder, puede encontrarse en [77].

PARA SABER MÁS

- El teorema de Stone-Weierstrass puede obtenerse como consecuencia del teorema de Krein-Milman 3.1.5, véase [36, p. 148-19]. Una versión del teorema de Stone-Weierstrass más general se encuentra en [95, p. 118-120].
- Los libros [55] y [107] son dos buenas referencias para el estudio de los teoremas del punto fijo. Ambos textos ofrecen demostraciones de los teoremas de Banach, Brouwer y Schauder. El primero de ellos estudia diversos tópicos en la teoría métrica del punto fijo (aplicaciones no expansivas), y el segundo recoge innumerables aplicaciones al estudio de ecuaciones diferenciales (soluciones periódicas, estabilidad, etc.), desarrollando además la teoría del grado de Leray-Schauder.
- Otras extensiones clásicas de los teoremas del punto fijo son las que tienen que ver con puntos fijos simultáneos para familias de aplicaciones, y con existencia de puntos fijos para multifunciones. Una referencia para el estudio de puntos fijos de multifunciones es [55]. Entre los resultados sobre puntos fijos simultáneos para familias de funciones, destacamos los teoremas de Markov-Kakutani y Ryll-Nardzewski. Este último se puede utilizar para demostrar la existencia de la medida de Haar en todo grupo topológico compacto, véase [36, p. 155-163].
- El libro [56] es un tratado extenso y moderno sobre teoremas del punto fijo, donde se trata una gran variedad de aspectos.

Capítulo 3

Optimización: funcionales que alcanzan la norma

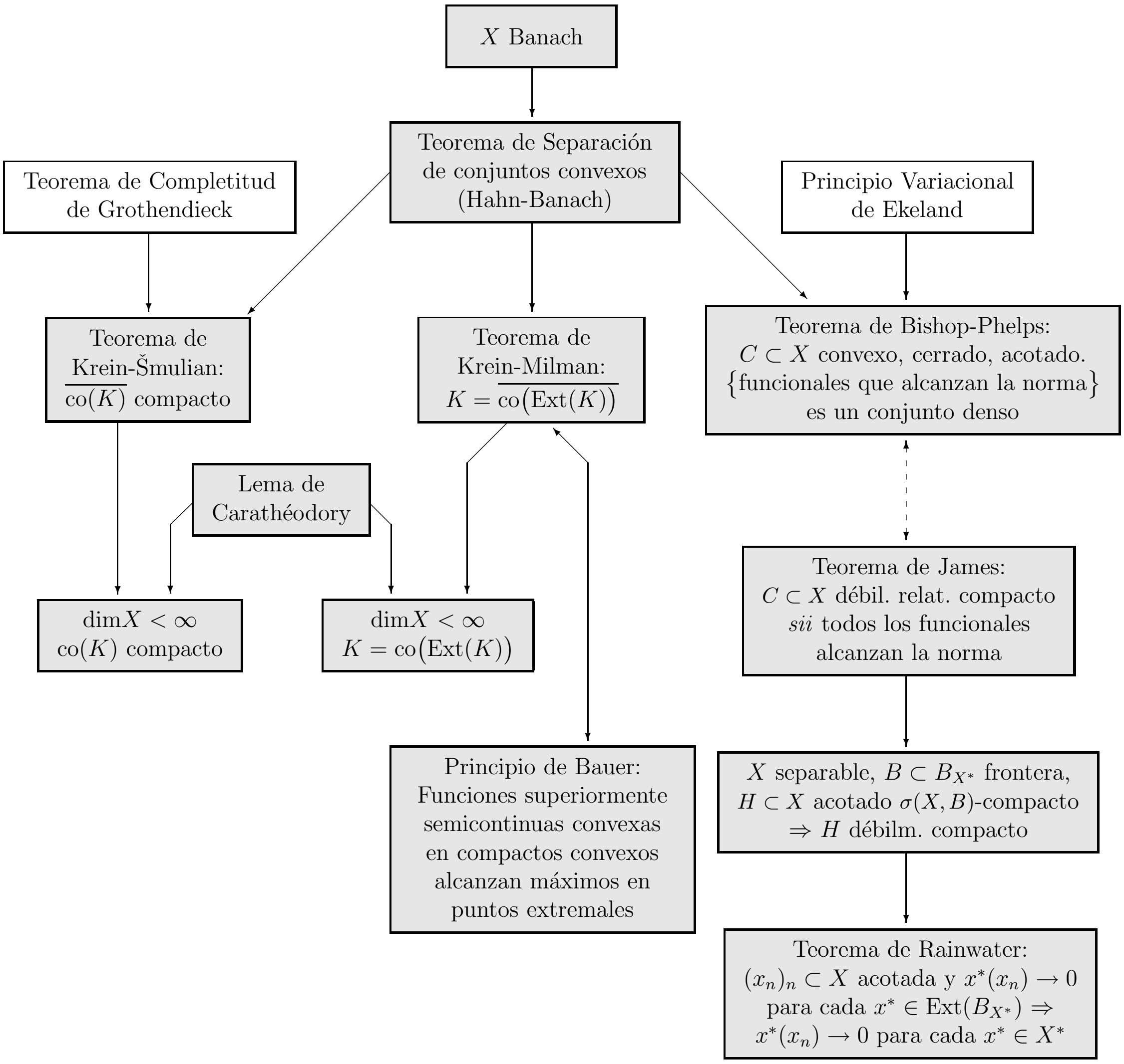
«OBJETIVOS»

- Estudiar los teoremas de Krein-Milman y Krein-Šmulian, tanto en sus versiones infinito como finito-dimensionales (teorema de Minkowski). Establecer la equivalencia entre el teorema de Krein-Milman y el Principio del Máximo de Bauer.
- Establecer el Principio Variacional de Ekeland y obtener, como consecuencia, el teorema de Bishop-Phelps sobre existencia de *suficientes* formas lineales continuas que alcanzan su supremo en conjuntos acotados de espacios de Banach.
- Presentar el teorema de James, que caracteriza la compacidad débil de subconjuntos débil cerrados de un espacio de Banach, vía la propiedad de que las formas lineales continuas alcancen su supremo en los mismos.
- Ligar el teorema de James, con ayuda del teorema de Krein-Šmulian, con problemas de aproximación y *tests* de convergencia y compacidad débil.

LA determinación de *máximos* y *mínimos* de funciones ha tenido, y tiene, un lugar preeminente en el Análisis Matemático y sus aplicaciones a cuestiones de *optimización*. El hilo conductor de este capítulo es el estudio de los conjuntos convexos (convexos compactos muchas de las veces) en conexión con las formas lineales que alcanzan su supremo en ellos. La herramienta básica para el estudio aquí presentado es la teoría de dualidad, desarrollada en la sección 1.3, y en especial, los teoremas de separación de conjuntos convexos, que son consecuencia del teorema de extensión de Hahn-Banach. Presentamos, en primer lugar, el teorema de Krein-Milman 3.1.5, el cual establece que si $K \subset E$ es un conjunto convexo y compacto en un espacio localmente convexo, entonces

$$K = \overline{\mathrm{co}\big(\mathrm{Ext}(K)\big)}. \tag{3.1}$$

Mostramos que este resultado es equivalente al Principio del Máximo de Bauer 3.1.11. A la vista de la igualdad (3.1), se plantean dos cuestiones de forma natural: (a) ¿Cuándo no es necesario



Cuadro 3.1: Esquema del capítulo *Optimización: funcionales que alcanzan la norma*

tomar clausuras en la fórmula establecida? (b) Si partimos de $K \subset E$ compacto, ¿cuándo es cierto que $\overline{\mathrm{co}\big(\mathrm{Ext}(K)\big)}$ es compacto? Damos respuesta a (b) cuando E es un espacio de Banach, teorema de Krein-Šmulian 3.2.1 y 3.2.3. La respuesta a (a) es positiva en los espacios de dimensión finita gracias al Lema de Carathéodory 3.2.4, véanse los teoremas de Minkowski 3.2.6 y 3.2.5. El Principio Variacional de Ekeland 3.3.1 abre las puertas al teorema de Bishop-Phelps 3.3.4, el cual asegura que, para un espacio de Banach X, el conjunto de las formas $x^* \in X^*$ que alcanzan su supremo en B_X es denso para la norma de X^*. La condición de que todas las formas $x^* \in X^*$ alcancen su supremo en B_X obliga a que B_X sea débilmente compacto, *i.e.*, X reflexivo, teorema de James 3.4.5. Presentamos algunas aplicaciones del teorema de James a la teoría general de la aproximación en espacios normados, véase el corolario 3.4.6. El capítulo se completa con una sección en la que ligamos el teorema de James, gracias a la ayuda del teorema de Krein-Šmulian, con ciertos *tests* de convergencia y compacidad débil.

3.1 El teorema de Krein-Milman

EN esta sección estudiamos los conceptos de conjunto y punto extremal junto con sus propiedades. El resultado fundamental que demostramos es lo que se conoce como teorema de Krein-Milman, el cual establece que todo conjunto compacto convexo de un espacio localmente convexo es la envoltura convexa cerrada del conjunto de sus puntos extremales. El teorema de Krein-Milman (llamado teorema de Minkowski en dimensión finita) representa uno de los primeros y más elegantes resultados para conjuntos compactos convexos. En apariencia, el teorema de Krein-Milman es un resultado inocente, pero tiene numerosas e importantes aplicaciones, véanse las páginas 110-113 para alguna de ellas y el artículo [94] para una colección más exhaustiva.

Conjunto y punto extremal Sean E un espacio vectorial y $K \subset E$ un conjunto. Un subconjunto no vacío $S \subset K$ se llama *conjunto extremal* de K cuando se satisface la siguiente condición:

$$\text{si } x,y \in K,\ 0<t<1 \text{ y } tx+(1-t)y \in S, \text{ entonces } x,y \in S.$$

Un punto $e \in K$ se dice que es un *punto extremal* si el conjunto $\{e\}$ es extremal. Así, si K es un conjunto convexo y $e \in K$ es extremal, entonces e no es punto interior de ningún segmento contenido en K.

Lema 3.1.1. *Si $S_1 \subset K$ es un conjunto extremal de K y $S_2 \subset S_1$ es un conjunto extremal de S_1, entonces S_2 es un conjunto extremal de K.*

Demostración. Sean $x,y \in K$, $0<t<1$ y $tx+(1-t)y \in S_2$. Como $tx+(1-t)y \in S_1$ y S_1 es un conjunto extremal de K, entonces $x,y \in S_1$. Así, $tx+(1-t)y \in S_2$ y $x,y \in S_1$. Si utilizamos ahora que S_2 un conjunto extremal de S_1, concluimos que $x,y \in S_2$, y la prueba termina. □

La siguiente proposición caracteriza los puntos extremales de conjuntos convexos.

Proposición 3.1.2. *Sean E un espacio vectorial y $K \subset E$ convexo. Son equivalentes:*

(i) *e es un punto extremal.*
(ii) *Si $x_1,x_2 \in K$, $x_1 \neq x_2$, $0 \le t \le 1$ y $e = tx_1+(1-t)x_2$, entonces, o bien $e=x_1$, o bien $e=x_2$.*
(iii) *Si $x_1,x_2 \in K$ y $e=(x_1+x_2)/2$, entonces $e=x_1=x_2$.*
(iv) *$K \setminus \{e\}$ es convexo.*

Demostración. Que *(i)*, *(ii)* y *(iii)* son equivalentes se sigue del hecho de que, para $t \ge 1/2$, se tiene la igualdad

$$e=(x_1+y)/2 \text{ si, y sólo si, } e=tx_1+(1-t)x_2,$$

donde $y=(2t-1)x_1+2(1-t)x_2\in K$. Claramente, *(i)* implica *(iv)*. Recíprocamente, si *(i)* no se cumple, podemos escribir

$$e=tx_1+(1-t)x_2,$$

para $0<t<1$ y $x_1,x_2\neq e$, contradiciendo así la convexidad de $K\setminus\{e\}$. $\square$

El siguiente lema pone de manifiesto cómo construir conjuntos extremales de forma natural.

Lema 3.1.3. *Sean E un espacio vectorial y $A\subset E$ un conjunto. Sea además $f\in E^{\#}$ de tal forma que $s=\sup\{f(x):x\in A\}<\infty$. Si $A_f=\{x\in A:f(x)=s\}$ es no vacío, entonces A_f es un conjunto extremal de A.*

Demostración. Sean $x,y\in A$ y $0<t<1$ tales que $tx+(1-t)y=z\in A_f$. Supongamos que $x\notin A_f$. Entonces $f(x)<s$, y por tanto,

$$f(z)=tf(x)+(1-t)f(y)<ts+(1-t)s=s,$$

es decir, $z\notin A_f$. De forma similar se razona que si $y\notin A_f$, entonces $z\notin A_f$, y la prueba termina. $\square$

Lema 3.1.4. *Sea $E[\mathfrak{T}]$ un* e.l.c. *y sea K un conjunto compacto de E. Entonces K contiene, al menos, un punto extremal.*

Demostración. Sea $\mathscr{P}$ la colección de todos los conjuntos extremales compactos de K. $\mathscr{P}$ es no vacía, puesto que $K\in\mathscr{P}$. Se ordena parcialmente $\mathscr{P}$ mediante la inclusión *hacia abajo*. Sea entonces $\mathscr{B}$ un subcolección de $\mathscr{P}$ totalmente ordenada, y escribamos $S:=\bigcap_{B\in\mathscr{B}}B$. Como $\mathscr{B}$ está totalmente ordenada, $\mathscr{B}$ es una familia de subconjuntos cerrados con la propiedad de la intersección finita en el compacto K. Por tanto, S es no vacío, compacto y se comprueba fácilmente que es un conjunto extremal de K. Ahora, en virtud del lema de Zorn 1.2.1, existe un elemento maximal A de $\mathscr{P}$. Utilizando el lema 3.1.3, para cada $f\in E'$, el conjunto A_f es compacto y extremal de A. Además, por el lema 3.1.1, el conjunto compacto A_f es extremal de K y, dado que A es maximal, $A_f=A$. Esto significa que cada $f\in E'$ es constante en A. Por último, como E' separa los puntos de E, obtenemos que A contiene un solo punto, que es, claro está, un punto extremal de K. $\square$

Teorema 3.1.5 (Krein-Milman)**.** *Sean $E[\mathfrak{T}]$ un* e.l.c. *y $K\subset E$ un conjunto convexo y compacto. Entonces, K es la envoltura convexa y cerrada del conjunto de sus puntos extremales* $\operatorname{Ext}(K)$*,* i.e.*,* $\overline{\operatorname{co}(\operatorname{Ext}(K))}=K$.

Demostración. Escribamos $H=\overline{\operatorname{co}(\operatorname{Ext}(K))}$. Como K es convexo y compacto, obtenemos que $H\subset K$, de donde se deduce que H es también compacto. La inclusión contraria, $K\subset H$, la probamos por reducción al absurdo: supongamos que existe $x_0\in K\setminus H$ y busquemos una contradicción. Si esto es así, podemos utilizar el corolario 1.3.26 para deducir la existencia de $f\in E'$ tal que

$$\sup\{f(x):x\in H\}<f(x_0)\leq\sup\{f(x):x\in K\}=:s. \tag{3.2}$$

Consideremos ahora el conjunto no vacío $K_f = \{x \in K : f(x) = s\}$. El lema 3.1.4 nos asegura que K_f contiene un punto extremal, que llamaremos e. Por el lema 3.1.3, el conjunto K_f es un subconjunto extremal de K, y así obtenemos que e es un punto extremal de K después de utilizar el lema 3.1.1. En particular, $K_f \cap \mathrm{Ext}(K) \neq \emptyset$. Por otro lado, de la desigualdad (3.2) deducimos que $K_f \cap H = \emptyset$ y, como $\mathrm{Ext}(K) \subset H$, obtenemos una contradicción que termina la prueba. $\square$

STOP Obsérvese que, como consecuencia del teorema de Krein-Milman, la bola unidad B_{X^*} del dual de un espacio de Banach X tiene *muchos* puntos extremales. A partir de esta observación se obtiene que el espacio de las sucesiones nulas c_0, con la norma del supremo, no es isométricamente isomorfo al dual de ningún espacio de Banach, dado que su bola unidad cerrada no tiene puntos extremales.

La demostración del teorema 3.1.5 encierra la siguiente versión del teorema de Krein-Milman.

Teorema 3.1.6. *Sean* $E[\mathfrak{T}]$ *un* e.l.c. *y* $K \subset E$ *un conjunto compacto. Entonces* $\overline{\mathrm{co}(\mathrm{Ext}(K))} = \overline{\mathrm{co}(K)}$.

Hacemos notar que, en general, el conjunto de los puntos extremales de un convexo compacto no es un conjunto cerrado, como pone de manifiesto el ejemplo de la figura 3.1.

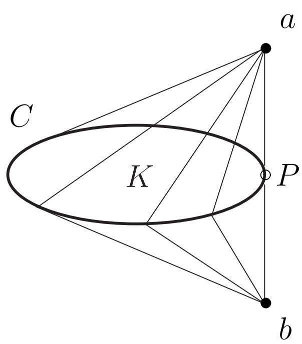


Figura 3.1: $K = \mathrm{co}(C \cup \{a,b\})$, $\mathrm{Ext}(K) = (C \setminus \{P\}) \cup \{a,b\}$

La igualdad en el teorema 3.1.6 da lugar a la siguiente pregunta: *¿puede decirse algo sobre la envoltura convexa de un conjunto compacto,* $\mathrm{co}(K)$*?* Incluso en el caso de un espacio de Hilbert, puede ocurrir que $\mathrm{co}(K)$ no sea cerrada, y existen casos en los que $\overline{\mathrm{co}(K)}$ no es compacta, véase [96, Ejericio 22. p. 84]. Esta última situación no es posible en espacios de Banach X para topologías localmente convexas $\mathfrak{T}$ compatibles con la dualidad $\langle X, X^* \rangle$, como pone de manifiesto el teorema de Krein-Šmulian 3.2.3. En espacios de dimensión finita, si $K \subset \mathbb{R}^n$ es compacto, entonces $\mathrm{co}(K)$ es compacto, corolario 3.2.5.

STOP En espacios de dimensión finita si $K \subset \mathbb{R}^n$ es compacto convexo, entonces $K = \mathrm{co}(\mathrm{Ext}(K))$ sin necesidad de tomar clausuras, teorema de Minkowski 3.2.6.

Combinando los lemas 3.1.3 y 3.1.4, es fácil concluir que, dado un conjunto compacto y convexo K de un *e.l.c.* $E[\mathfrak{T}]$, toda forma lineal y continua $f \in E'$ alcanza su supremo en K en un punto extremal de K. Prácticamente con el mismo trabajo, se puede demostrar que el resultado anterior es cierto para funciones convexas superiormente semicontinuas, véanse el principio del máximo de Bauer 3.1.11 y el gráfico de la figura 3.5, que ilustra esta afirmación.

El siguiente resultado se utiliza para demostrar el lema de Choquet 3.1.8.

Lema 3.1.7. *Sean* $E[\mathfrak{T}]$ *un* e.v.t. *y* $K_1, K_2, \dots, K_m$ *subconjuntos convexos compactos de* E. *Entonces, la envoltura convexa* $\operatorname{co}\left(\bigcup_{i=1}^m K_i\right)$ *es un conjunto compacto (y por tanto, cerrado).*

Demostración. Sea

$$K = \left\{ \sum_{i=1}^m t_i x_i : t_i \geq 0,\ \sum_{i=1}^m t_i = 1,\ x_i \in K_i,\ i = 1, 2, \dots, m \right\}.$$

Como los K_i son convexos, obtenemos que K también lo es y, por tanto, $K = \operatorname{co}\left(\bigcup_{i=1}^m K_i\right)$. Veamos ahora que K es compacto. Sea $\{e_i\}_{i=1}^m$ la base canónica de $\mathbb{R}^m$ y escribamos

$$C = \left\{ \sum_{i=1}^m t_i e_i \in \mathbb{R}^m : t_i \geq 0,\ \sum_{i=1}^m t_i = 1 \right\}.$$

Sea $M = C \times \left(\prod_{i=1}^m K_i \right)$. Dado que C y cada K_i son compactos, se tiene que M es compacto, véase la proposición 1.1.14. Consideremos ahora la aplicación $F : M \longrightarrow K$ definida por la fórmula

$$F\left(\sum_{i=1}^m t_i e_i, x_1, x_2, \dots, x_m \right) = \sum_{i=1}^m t_i x_i.$$

Probemos que F es continua. Fijamos $v = \sum_{i=1}^m t_i x_i \in K$ y V un entorno de 0 en $E[\mathfrak{T}]$. Entonces, existe un $\mathfrak{T}$-entorno equilibrado del origen, W, tal que $W + \overbrace{\dots}^{2m \text{ veces}} + W \subset V$. Sea $\delta > 0$ de forma que $\delta x_i \in W$, para cada $i = 1, 2, \dots, m$. Si tomamos $d = \left(\sum_{i=1}^m s_i e_i, k_1, \dots, k_m \right) \in M$ verificando $\left\| \sum s_i e_i - \sum t_i e_i \right\|_2 < \delta$ y $k_i \in x_i + W$, $i = 1, 2, \dots, m$, como $|s_i - t_i| < \delta$, obtenemos que

$$\begin{aligned}
F(d) &= s_1 k_1 + \dots + s_m k_m \\
&\in s_1(x_1 + W) + \dots + s_m(x_m + W) \\
&\subset s_1 x_1 + \dots + s_m x_m + W + \overbrace{\dots}^{m \text{ veces}} + W \\
&= t_1 x_1 + \dots + t_m x_m + (s_1 - t_1) x_1 + \dots + (s_m - t_m) x_m + W + \overbrace{\dots}^{m \text{ veces}} + W \\
&\subset v + W + \overbrace{\dots}^{2m \text{ veces}} + W \\
&\subset v + V,
\end{aligned}$$

lo que prueba la continuidad de F. Dado que $F(M) = K$, concluimos que K es compacto, como queríamos demostrar. $\square$

Semiespacios Decimos que H es un *semiespacio abierto* de un espacio vectorial topológico $E[\mathfrak{T}]$ si existen $f \in E'$ y $a \in \mathbb{R}$ tales que $H = \{y \in E : f(y) > a\}$. Denotemos por $\mathscr{H}_x$ la familia de todos los semiespacios abiertos que contienen el punto $x \in E$.

Puntos fuertemente extremales Sea K un subconjunto convexo de un *e.l.c.* $E[\mathfrak{T}]$. Un punto $e \in K$ se dice que es *fuertemente extremal* si la familia $\mathscr{U}_e = \{K \cap H : H \in \mathscr{H}_e\}$ forma una base de $\mathfrak{T}$-entornos en K de e.

Teorema 3.1.8 (Lema de Choquet)**.** *Sea K un conjunto convexo y compacto de un* e.l.c. $E[\mathfrak{T}]$ *y sea $e \in K$. El punto e es extremal de K si, y sólo si, es fuertemente extremal.*

Demostración. Observemos primero que, al ser K $\mathfrak{T}$-compacto y $\sigma(E,E')$ Hausdorff, se tiene que $\mathfrak{T}|_K = \sigma(E,E')|_K$, véase el corolario 1.1.13. Supongamos ahora que e es un punto extremal de K y veamos que $\mathscr{U}_e$ es base de entornos de e. Fijemos un $\mathfrak{T}$-entorno U de e. Como $\sigma(E,E')$ coincide con $\mathfrak{T}$ sobre K, existen $H_1, H_2, \dots, H_m \in \mathscr{H}_e$ tales que

$$\left(\bigcap_{i=1}^{m} H_i\right) \cap K \subset U.$$

Definimos $L_i = (E \setminus H_i) \cap K$, $i = 1,2,\dots,m$. Dado que cada $E \setminus H_i$ es convexo y cerrado, obtenemos que L_i es compacto y convexo. Por tanto, $L := \operatorname{co}\left(\bigcup_{i=1}^m L_i\right)$ es $\mathfrak{T}$-compacto, después del lema 3.1.7. Ahora probamos que $e \notin L$. Procedamos por reducción al absurdo y supongamos que sí. Existirán entonces $\mu_i \geq 0$, $\sum_{i=1}^m \mu_i = 1$, y $l_i \in L_i$, tales que $e = \sum_{i=1}^m \mu_i l_i$. Al ser e un punto extremal de K, tenemos que $e = l_j \in L_j$ para algún j (utilícese la caracterización *(iv)* de la proposición 3.1.2). Como, por otro lado, $e \in H_j$ y $H_j \cap L_j = \emptyset$, llegamos a una contradicción que nos asegura que $e \notin L$. En virtud del corolario 1.3.26, existe $f \in E'$ tal que $f(x) > s := \sup\{f(y) : y \in L\}$. Definimos $H = \{y \in E : f(y) > s\}$. Entonces, $e \in H \cap K$ y $H \cap L = \emptyset$.

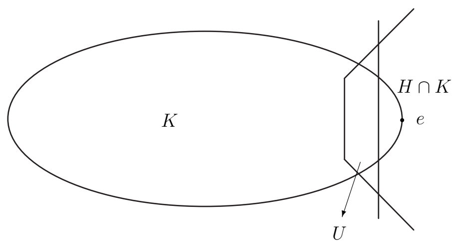


Figura 3.2: El lema de Choquet

Tenemos así que $H \cap L_i = \emptyset$, para cada $i = 1,2,\dots,m$, y por tanto,

$$H \cap K \subset \bigcap_{i=1}^{m} (E \setminus L_i) \cap K = \bigcap_{i=1}^{m} (H_i \cap K) \subset U,$$

lo que significa que la colección $\mathscr{U}_e$ forma una base de entornos del punto e y, por definición, e es fuertemente extremal, véase la figura 3.2.

Recíprocamente, supongamos que e es fuertemente extremal y demostremos que e es un punto extremal. Procedemos por reducción al absurdo. Si e no es extremal en K, existen $y,z \in K$, $y \neq z$, tales que $e = (y+z)/2$. Pongamos $u = y - e$. Se tiene entonces que $y = e+u$, $z = 2e - y = e - u$. Como E' separa los puntos de E y $u \neq 0$, existe $f \in E'$ tal que $a = f(u) > 0$. Sea

$$V = \left\{ v \in X : \left| f(e-v) \right| < \frac{a}{2} \right\}.$$

Es claro que V es un entorno de e. Al ser $\{H \cap K : H \in \mathscr{H}_e\}$ una base de entornos del punto e, existe $H \in \mathscr{H}_e$ tal que $e \in H \cap K \subset V \cap K$. Como $e \pm u \notin V$, obtenemos que $e \pm u \notin H \cap K$. Pero $e \pm u \in K$, luego $e \pm u \notin H$. Entonces $e \pm u \in E \setminus H$. Dado que $E \setminus H$ es convexo, se tiene que $e = \frac{e+u}{2} + \frac{e-u}{2} \in E \setminus H$, es decir, $e \notin H$, y llegamos así a la contradicción que acaba la prueba. $\square$

Corolario 3.1.9. *Sean $E[\mathfrak{T}]$ un* e.l.c. *y $K \subset E$ un conjunto compacto convexo. Sea $F \subset K$ un conjunto cerrado tal que $K = \operatorname{co}(F)$. Entonces,* $\operatorname{Ext}(K) \subset F$.

Demostración. Sea e un punto extremal de K que no está en F. Como, por el lema de Choquet, e es fuertemente extremal, existe $H \in \mathscr{H}_e$ tal que $H \cap F = \emptyset$. Por tanto, $F \subset E \setminus H$. Al ser $E \setminus H$ convexo y cerrado, se tiene que $K = \operatorname{co}(F) \subset E \setminus H$, y así, $e \in E \setminus H$, lo cual contradice que $e \in H$ y la prueba termina. $\square$

Aplicaciones del teorema de Krein-Milman

Tal y como ha sido comentado previamente, el teorema de Krein-Milman puede ser utilizado para demostrar otros prominentes resultados; por ejemplo, el teorema de Stone-Weierstrass 2.1.4, el principio del máximo de Bauer 3.1.11, o el teorema de Bernstein 3.1.10 que caracteriza las funciones completamente monótonas.

$\boxed{\textit{Funciones completamente monótonas}}$ Una función f definida en $(0,\infty)$ se dice que es *completamente monótona* si f tiene derivadas $f^{(0)} := f$, $f^{(1)}$, $f^{(2)}, \ldots$ de todos los órdenes y, además, $(-1)^n f^{(n)} \geq 0$, para $n = 0, 1, \ldots$ Así, f y cada una de las funciones $(-1)^n f^{(n)}$ es no negativa y no creciente. Ejemplos de tales funciones son $x^{-\alpha}$ y $e^{-\alpha x}$, para $\alpha \geq 0$.

S. Bernstein estableció un teorema fundamental sobre funciones completamente monótonas, que puede ser demostrado utilizando el teorema de Krein-Milman, véase [86, p. 9-12].

Teorema 3.1.10 (Bernstein)**.** *Sea f una función acotada definida en $(0,\infty)$. Son equivalentes:*

(i) f es completamente monótona.

(ii) Existe una (única) medida de Borel no negativa μ en $[0,\infty)$ tal que $\mu\big([0,\infty)\big) = f(0^+)$, satisfaciendo

$$f(x) = \int_0^\infty e^{-\alpha x} d\mu(\alpha).$$

Funciones convexas y cóncavas Una función $f : C \longrightarrow \mathbb{R}$ definida en un subconjunto convexo C de un espacio vectorial E se dice *convexa* (respectivamente, *cóncava*) si

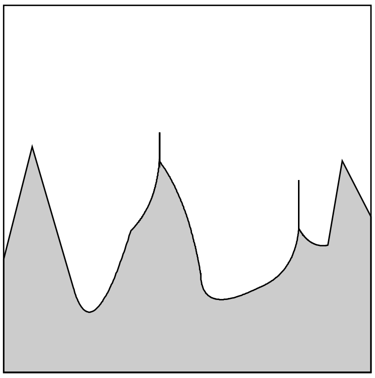
$$f\big(tx+(1-t)y\big) \leq tf(x)+(1-t)f(y)$$

(respectivamente, $f\big(tx+(1-t)y\big) \geq tf(x)+(1-t)f(y)$), para puntos cualesquiera $x,y \in C$ y cada $0 \leq t \leq 1$.

Función superiormente semicontinua Una función $f : X \longrightarrow \mathbb{R}$ definida en un espacio topológico X se dice *superiormente semicontinua* si, para cada $t \in \mathbb{R}$, el conjunto $\big\{x \in X : f(x) \geq t\big\}$ es cerrado. Equivalentemente, f es superiormente semicontinua si el subgrafo

$$\mathrm{Sub}(f) := \big\{(x,t) : f(x) \geq t\big\}$$

es cerrado en $X \times \mathbb{R}$, véase la figura 3.3. Las funciones superiormente semicontinuas definidas en un espacio compacto son acotadas superiormente, y alcanzan el supremo en un punto del mismo, [32, p. 137-142].



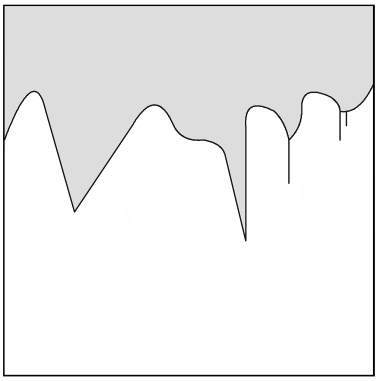
$\mathrm{Sub}(f) = \big\{(x,t) : f(x) \geq t\big\}$

Figura 3.3: Función superiormente semicontinua: subgrafo

Funciones inferiormente semicontinuas Una función $f : X \longrightarrow \mathbb{R}$ definida en un espacio topológico X se dice que es *inferiormente semicontinua* si $-f$ es superiormente semicontinua, *i.e.*, si para cada $t \in \mathbb{R}$, el conjunto $\big\{x \in X : f(x) \leq t\big\}$ es cerrado. Equivalentemente, f es inferiormente semicontinua si el epígrafo

$$\mathrm{Epi}(f) := \big\{(x,t) : f(x) \leq t\big\}$$

es cerrado en $X \times \mathbb{R}$, véase la figura 3.4.



$\mathrm{Epi}(f) = \{(x, t) : f(x) \le t\}$

Figura 3.4: Función inferiormente semicontinua: epígrafo

Operaciones con funciones inferiormente semicontinuas Si φ, ψ son funciones inferiormente semicontinuas en un espacio topológico X, entonces la suma $\varphi + \psi$ es inferiormente semicontinua. Más en general, si $(\varphi_n)_{n=1}^{\infty}$ es una sucesión de funciones inferiormente semicontinuas de forma que la serie $\varphi = \sum_{n=1}^{\infty} \varphi_n$ converge uniformemente en X, entonces φ es inferiormente semicontinua en X. Si φ es una función inferiormente semicontinua en X y ϕ es continua y creciente en $(a,b) \supset \varphi(X)$, entonces la composición $\phi \circ \varphi$ es inferiormente semicontinua en el dominio de ϕ. Si $(\varphi_i)_{i\in I}$ es una familia de funciones inferiormente semicontinuas, entonces la función $\varphi(x) := \sup\{\varphi_i(x) : i \in I\}$ es inferiormente semicontinua.

Teorema 3.1.11 (Bauer)**.** *Sean $E[\mathfrak{T}]$ un* e.l.c. *y $K \subset E$ un conjunto compacto convexo no vacío. Supongamos que $f : K \longrightarrow \mathbb{R}$ es una función convexa y superiormente semicontinua. Entonces, existe un punto extremal de K (no necesariamente único) en el que f alcanza el supremo.*

Demostración. Por la compacidad de K y por ser f superiormente semicontinua, sabemos que $s := \sup\{f(y) : y \in K\} < \infty$ y que f alcanza su supremo en K. Así, el conjunto

$$K_f := \{x \in K : f(x) = s\}$$

es no vacío. Obsérvese que la igualdad

$$K_f = \bigcap_{n=1}^{\infty} \{x \in K : f(x) \ge s - 1/n\},$$

nos permite concluir que K_f es un conjunto cerrado, gracias, de nuevo, a la semicontinuidad superior de f. Por otro lado, K_f es un conjunto extremal de K. Para demostrar esto, razonamos de forma similar a como hicimos en el lema 3.1.3, utilizando ahora que f es convexa: sean $x, y \in K$ y

$0 < t < 1$ tales que $tx + (1-t)y = z \in K_f$. Supongamos que $x \notin K_f$. Entonces $f(x) < s$, luego

$$f(z) \leq tf(x) + (1-t)f(y) < ts + (1-t)s = s,$$

es decir, $z \notin K_f$; si $y \notin K_f$ se razona de modo análogo, concluyéndose entonces que $z \notin K_f$.

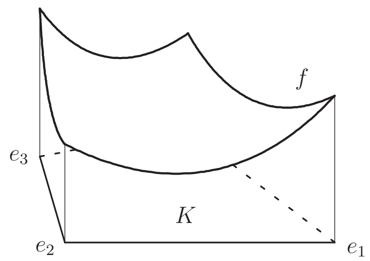


Figura 3.5: Principio del Máximo de Bauer

Ahora, utilizamos el lema 3.1.4 para deducir la existencia de un punto extremal e de K_f, que es, por el lema 3.1.1, un punto extremal de K, y la prueba termina, véase la figura 3.5. $\square$

En [33, Theorem 25.9 y Theorem 25.12] se establece primero el principio del máximo de Bauer, y de él se obtiene el teorema de Krein-Milman. Consecuentemente, el teorema de Krein-Milman y el principio del máximo de Bauer son equivalentes.

El teorema de Krein-Milman, 1911-1940. La noción de punto extremal fue introducida por H. Minkowski en 1911, quien probó que si K es un subconjunto compacto convexo de $\mathbb{R}^3$, entonces cada punto de K puede expresarse como combinación convexa de puntos extremales de K. Este resultado, conocido (en $\mathbb{R}^n$) como el *teorema de Minkowski*, fue *precisado* por C. Carathéodory, quien demostró que si K es un subconjunto compacto convexo de $\mathbb{R}^n$, entonces cada punto de K puede expresarse como una combinación convexa de, a lo más, $n+1$ puntos extremales de K. En 1940, M. Krein y D. Milman extendieron el teorema de Minkowski a dimensión infinita. El teorema de Minkowski ha servido como el punto de partida de toda la investigación desarrollada sobre la estructura externa de los conjuntos convexos en espacios de dimensión finita.

3.2 El teorema de Krein-Šmulian

En esta sección estudiamos cuándo la envoltura convexa cerrada de un conjunto compacto es de nuevo compacta. Aunque algunos de los resultados que presentamos aquí son ciertos en situaciones más generales, hemos decidido, sin embargo, restringirnos al caso de los espacios de Banach y topologías compatibles con la dualidad asociada, lo que proporciona un marco suficientemente amplio e interesante.

Teorema 3.2.1 (Krein-Šmulian, Topología débil)**.** *Sea X un espacio de Banach y sea K un subconjunto débilmente compacto de X. Entonces la envoltura convexa y cerrada de K, $\overline{\mathrm{co}(K)}$, es débilmente compacto.*

Demostración. Vamos a razonar, en primer lugar, que el resultado general se puede reducir al caso de espacios de Banach separables. Efectivamente, de acuerdo con el teorema de Eberlein-Šmulian, [51, p. 38-39], para demostrar que $\overline{\mathrm{co}(K)}$ es débilmente compacto es suficiente probar que cada sucesión $(x_n)_n$ en $\mathrm{co}(K)$ tiene una subsucesión débilmente convergente hacia algún punto de X. Para cada $n \in \mathbb{N}$, tomamos un conjunto finito $H_n \subset \mathbb{N}$, y puntos $y_j \in K$ y escalares $t_j \geq 0$, para $j \in H_n$, tales que $\sum_{j\in H_n} t_j = 1$ y $x_n = \sum_{j\in H_n} t_j y_j$. Definimos $Y := \overline{\mathrm{span}\{y_j : j \in H_n, n \in \mathbb{N}\}}$. Claramente, Y es un espacio de Banach separable con la norma inducida por la de X, y $K_0 := \overline{\{y_j : j \in H_n, n \in \mathbb{N}\}}$ es $\sigma(Y,Y^*)$-compacto, teniéndose que $x_n \in \mathrm{co}(K_0)$. Si el resultado que queremos demostrar es cierto para espacios de Banach separables, se tendrá, de nuevo gracias al teorema de Eberlein-Šmulian, que $(x_n)_n$ tiene una subsucesión $\sigma(Y,Y^*)$-convergente a un punto $x \in Y$ o, equivalentemente, que $(x_n)_n$ tiene una subsucesión débilmente convergente en X al punto x. Así, el resultado general se sigue del caso separable.

Supongamos a partir de ahora que X es separable, y demostremos que $\overline{\mathrm{co}(K)}$ es débilmente compacto en X. Consideremos el espacio de las funciones continuas en K, $C(K)$, dotado de su norma del supremo. El dual de $C(K)$ puede identificarse, de acuerdo con el teorema de Riesz, [34, Theorem 7.3.5], con el espacio $M(K)$ de las medidas de Radon definidas en la σ-álgebra de Borel de K: esta identificación asocia, a cada medida $\mu \in M(K)$, el elemento de $C(K)^*$ definido vía la integral $f \mapsto \int_K f\,d\mu$, para $f \in C(K)$. Si $P(K)$ es el conjunto de las probabilidades de Radon en K, entonces $P(K)$ es un conjunto convexo y $\sigma\big(M(K),C(K)\big)$-cerrado, contenido en la bola unidad cerrada $B_{M(K)}$. Es fácil probar que $P(K)$ es un subconjunto $\sigma\big(M(K),C(K)\big)$-cerrado de $B_{M(K)}$, y así, si tenemos ahora en cuenta el teorema de Alaoglu-Bourbaki, corolario 1.3.42, concluimos que $P(K)$ es $\sigma\big(M(K),C(K)\big)$-compacto. Para cada $\mu \in P(K)$, consideramos $T_\mu(x^*) = \int_K x^*|_K\,d\mu$, $x^* \in X^*$. Por un lado, $T_\mu \in X^{**}$, y por otro, T_μ restringido a B_{X^*} es $\sigma(X^*,X)$-continua. Veamos esto último: como X es separable, $\big(B_{X^*},\sigma(X^*,X)\big)$ es metrizable, y así, para probar que T_μ es continuo es suficiente demostrar que es sucesionalmente continuo; si $(x_n^*)_n$ y x^* están en B_{X^*}, y (x_n^*) converge hacia x^* en la topología $\sigma(X^*,X)$, el teorema de convergencia dominada de Lebesgue nos garantiza que

$$\lim_n T_\mu(x_n^*) = \lim_n \int_K x_n^*|_K\,d\mu = \int_K x^*|_K\,d\mu = T_\mu(x^*),$$

y por tanto, T_μ es $\sigma(X^*,X)$-sucesionalmente continuo restringido a B_{X^*}. Apelando ahora al teorema de completitud de Grothendieck 1.3.55, concluimos que existe un único $x_\mu \in X$ de forma que $T_\mu(x^*) = x^*(x_\mu)$, para cada $x^* \in X^*$. Es fácil ver que la aplicación $\phi : P(K) \longrightarrow X$, dada por $\phi(\mu) = x_\mu$, es afín y $\sigma\big(M(K),C(K)\big)$-$\sigma(X,X^*)$-continua, y que $K \subset \phi\big(P(K)\big)$. En consecuencia, $\phi\big(P(K)\big)$ es un conjunto convexo y $\sigma(X,X^*)$-compacto que contiene a K. Por lo tanto, $\overline{\mathrm{co}(K)} \subset \phi\big(P(K)\big)$ es $\sigma(X,X^*)$-compacto, y la prueba termina. $\square$

Conjuntos totalmente acotados en espacios métricos Recordemos que un subconjunto A de un espacio métrico M es *totalmente acotado* o *precompacto* cuando, para cada $\varepsilon > 0$, A está contenido en la unión de una familia finita de bolas abiertas de radio ε. Conviene tener presente que un conjunto A de un espacio métrico M es compacto si, y sólo si, A es completo para la distancia inducida y totalmente acotado, [69, Teorema 32, p. 227].

Conjuntos totalmente acotados en e.v.t. Un conjunto A de un espacio vectorial topológico E es *totalmente acotado* si, para cada entorno V de 0 en E, existe un conjunto finito $F \subset E$ tal que $A \subset F + V$. Al igual que en el caso de espacios métricos, A es un subconjunto compacto del *e.v.t.* E si, y sólo si, A es completo y totalmente acotado, [69, Teorema 32, p. 227].

En el caso de que E sea un *e.v.t.* metrizable, las dos nociones de acotación total coinciden, siempre que nos limitemos a considerar métricas invariantes por traslaciones compatibles con la topología de E.

Lema 3.2.2. *Sean E un* e.l.c. *y $H \subset E$ totalmente acotado. Entonces, la envoltura convexa* $\mathrm{co}(H)$ *también es totalmente acotada.*

Demostración. Fijemos U entorno de 0 en E y tomemos un entorno convexo V de 0 en E tal que $V + V \subset U$. Existe un conjunto finito $H_0 \subset E$ tal que $H \subset H_0 + V$. Es claro que se tiene la inclusión

$$\mathrm{co}(H) = \mathrm{co}(H_0) + V. \tag{3.3}$$

Como la envoltura convexa $\mathrm{co}(H_0)$ es un conjunto compacto después del lema 3.1.7, existe un conjunto finito $F \subset E$ tal que $\mathrm{co}(H_0) \subset F + V$. De aquí, teniendo en cuenta la inclusión (3.3), concluimos que

$$\mathrm{co}(H) \subset F + V + V \subset F + U.$$

Como U era arbitrario, hemos probado que H es totalmente acotado. $\square$

Corolario 3.2.3 (Teorema de Mazur y Krein-Šmulian)**.** *Sea X un espacio de Banach y sea $\mathfrak{T}$ una topología localmente convexa en X, compatible con el par dual $\langle X, X^* \rangle$. Si $K \subset X$ es $\mathfrak{T}$-compacto, entonces* $\overline{\mathrm{co}(K)}$ *es $\mathfrak{T}$-compacto.*

Demostración. La envoltura convexa $\mathrm{co}(K)$ es totalmente acotado, después del lema 3.2.2. Es fácil comprobar que $\overline{\mathrm{co}(K)}$ es de nuevo totalmente acotado, y así, para terminar la prueba, basta demostrar que $\overline{\mathrm{co}(K)}$ es $\mathfrak{T}$-completo.

Obsérvese que, dado que $\mathfrak{T}$ es una topología localmente convexa compatible con el par dual $\langle X, X^* \rangle$, se tiene que $\overline{\mathrm{co}(K)}^{\mathfrak{T}} = \overline{\mathrm{co}(K)}^{\sigma(X,X^*)}$ es un conjunto $\sigma(X,X^*)$-compacto después del teorema 3.2.1 y, en consecuencia, $\sigma(X,X^*)$-completo. Obsérvese también que $(E, \mathfrak{T})$ tiene una base de entornos del origen formada por conjuntos absolutamente convexos y $\mathfrak{T}$-cerrados, por ende, $\sigma(X,X^*)$-cerrados, corolario 1.3.29. Para demostrar que $\overline{\mathrm{co}(K)}$ es $\mathfrak{T}$-completo razonamos como sigue. Sea $(x_j)_{j\in D}$ una red de Cauchy para $\mathfrak{T}$ en $\overline{\mathrm{co}(K)}$. En particular, $(x_j)_{j\in D}$ es $\sigma(X,X^*)$-Cauchy, y por tanto, existe $x \in \overline{\mathrm{co}(K)}$ tal que $(x_j)_{j\in D}$ converge a x en la topología $\sigma(X,X^*)$. Fijemos U un $\mathfrak{T}$-entorno del origen en X, absolutamente convexo y $\sigma(X,X^*)$-cerrado. Sea $j_U \in D$ tal que

$$x_i - x_j \in U, \quad \text{para cada } \; i, j \geq j_U. \tag{3.4}$$

Tomando límites en la topología $\sigma(X,X^*)$ para la red $(x_i)_{i\geq j_U}$, la ecuación (3.4) conduce a que

$$x - x_j \in U, \quad \text{para cada } \; j \geq j_U,$$

lo cual significa que $(x_j)_{j\in D}$ converge hacia x en la topología $\mathfrak{T}$, y la prueba termina. □

En espacios de dimensión finita no es necesario tomar la clausura en el teorema de Krein-Šmulian. Esta mejora se demuestra utilizando el lema de Carathéodory que sigue:

Lema 3.2.4 (Carathéodory)**.** *Si x está en la envoltura convexa de un conjunto $H \subset \mathbb{R}^n$, entonces x pertenece a la envoltura convexa de algún subconjunto de H que tiene, a lo sumo, $n+1$ puntos.*

Demostración. Basta probar que, si $r > n$ y $x = \sum_{i=1}^{r+1} t_i x_i$ es una combinación convexa de $r+1$ vectores $x_i \in H$, entonces x es, efectivamente, combinación convexa de r vectores extraídos de entre ellos. Sin pérdida de generalidad, podemos suponer que todos los $t_i > 0$, para $1 \leq i \leq r+1$. Los r vectores $x_i - x_{r+1}$, $1 \leq i \leq r$, son linealmente dependientes, pues $r > n$. Se sigue entonces que existen números reales b_i, no todos nulos, tales que

$$\sum_{i=1}^{r} b_i(x_i - x_{r+1}) = 0.$$

Sean $a_i = b_i$, para $i \in \{1,2,\ldots,r\}$, y $a_{r+1} = -\sum_{i=1}^{r} b_i$. Entonces,

$$\sum_{i=1}^{r+1} a_i x_i = 0 \quad \text{y} \quad \sum_{i=1}^{r+1} a_i = 0.$$

Tomemos m de forma que $\left|\dfrac{a_i}{t_i}\right| \leq \left|\dfrac{a_m}{t_m}\right|$, para $1 \leq i \leq r+1$, y definamos

$$c_i = t_i - \frac{a_i t_m}{a_m}, \quad 1 \leq i \leq r+1.$$

Entonces, $c_i \geq 0$, $\sum c_i = \sum t_i = 1$, $x = \sum c_i x_i$ y $c_m = 0$. □

Corolario 3.2.5. *Sean $X = \mathbb{R}^n$ y $K \subset X$ un subconjunto compacto. Entonces,* $\mathrm{co}(K)$ *es compacto.*

Demostración. Sea S el símplice de $\mathbb{R}^{n+1}$ formado por todos los $t = (t_1,\ldots,t_{n+1})$ tales que $t_i \geq 0$ y $\sum t_i = 1$. Por el lema 3.2.4, $x \in \mathrm{co}(K)$ si, y sólo si,

$$x = \sum_{i=1}^{n+1} t_i x_i,$$

para ciertos $t \in S$ y $x_i \in K$, $1 \leq i \leq n+1$. Dicho de otra forma, $\mathrm{co}(K)$ es la imagen del compacto

$$S \times K \times \overset{(n+1)}{\cdots} \times K$$

por la aplicación continua

$$(t, x_1, \ldots, x_{n+1}) \to \sum_{i=1}^{n+1} t_i x_i.$$

Por consiguiente, $\mathrm{co}(K)$ es compacto. □

Vamos ahora a utilizar el lema de Carathéodory anterior para demostrar que, en el teorema de Krein-Milman 3.1.5, se puede suprimir la clausura siempre que trabajemos con espacios de dimensión finita.

Teorema 3.2.6 (Minkowski, 1911). *Sea K un subconjunto compacto convexo de un* e.l.c. *finito-dimensional. Entonces $K = \mathrm{co}\big(\mathrm{Ext}(K)\big)$.*

Demostración. Suponemos que los espacios localmente convexos considerados son espacios reales, y la demostración la hacemos por inducción sobre su dimensión, n. Para $n = 1$, esto es, para convexos compactos de $\mathbb{R}$, el resultado es cierto: si $K \subset \mathbb{R}$ es compacto convexo, entonces $K = [a,b]$, siendo, por tanto, $\mathrm{Ext}\big([a,b]\big) = \{a,b\}$; claramente se satisface que $K = \mathrm{co}\big(\mathrm{Ext}(K)\big)$.

Supongamos que hemos probado el resultado para espacios de dimensión $n-1$, y vamos a demostrarlo para un espacio arbitrario E de dimensión n. Sea $K \subset E$ compacto y convexo. Por el teorema de Krein-Milman 3.1.5, sabemos que $K = \overline{\mathrm{co}(\mathrm{Ext}(K))}$. Para probar que $K = \mathrm{co}\big(\mathrm{Ext}(K)\big)$ es suficiente demostrar que

$$\overline{\mathrm{Ext}(K)} \subset \mathrm{co}\big(\mathrm{Ext}(K)\big). \tag{3.5}$$

Efectivamente, si la última inclusión se da, tenemos que $\mathrm{co}\big(\overline{\mathrm{Ext}(K)}\big)$ es compacto, después del corolario 3.2.5. En particular, $\mathrm{co}\big(\overline{\mathrm{Ext}(K)}\big)$ es cerrado, y por tanto, se tiene que

$$K = \overline{\mathrm{co}\big(\mathrm{Ext}(K)\big)} \subset \mathrm{co}\big(\overline{\mathrm{Ext}(K)}\big) \subset \mathrm{co}\big(\mathrm{Ext}(K)\big)$$

que, claramente, proporciona $K = \mathrm{co}\big(\mathrm{Ext}(K)\big)$. Probemos la inclusión (3.5) por reducción al absurdo. Supongamos que existe $e \in \overline{\mathrm{Ext}(K)}$ tal que $e \notin \mathrm{co}\big(\mathrm{Ext}(K)\big) =: A$. Como $K = \overline{A}$, nuestra suposición significa que $e \in \overline{A} \setminus A$. Por tanto, existen un funcional $f : E \longrightarrow \mathbb{R}$ y $\alpha \in \mathbb{R}$ tales que $f(e) = \alpha$ y $A \subset \big\{x \in E : f(x) \leq \alpha\big\}$. Es claro que $f(e) = \alpha = \max\big\{f(x) : x \in K\big\}$. El compacto convexo $K_f = \big\{x \in K : f(x) = \alpha\big\}$ está en una variedad afín de dimensión $n-1$, y en consecuencia, por hipótesis de inducción, $e \in K_f = \mathrm{co}\big(\mathrm{Ext}(K_f)\big)$. De acuerdo a los lemas 3.1.1 y 3.1.3, cada punto extremal de K_f es un punto extremal de K, luego $e \in \mathrm{co}\big(\mathrm{Ext}(K)\big)$, lo cual es una contradicción que acaba la prueba. □

3.3 Principio variacional de Ekeland y teorema de Bishop-Phelps

EL principio variacional de Ekeland 3.3.1 apareció en la prueba original del teorema de Bishop-Phelps 3.3.4. En [44] se encuentran aplicaciones del principio variacional para determinar la existencia de autovalores aproximados en problemas no lineales y soluciones aproximadas en problemas de control óptimo.

Teorema 3.3.1 (Principio Variacional de Ekeland)**.** *Sean (M,d) un espacio métrico completo y $f : M \longrightarrow \mathbb{R} \cup \{\infty\}$ una función inferiormente semicontinua y acotada inferiormente tal que $\{x \in M : f(x) < \infty\} \neq \emptyset$. Entonces, para cada $\varepsilon > 0$, existe $x_\varepsilon \in M$ tal que*

$$f(x) \geq f(x_\varepsilon) - \varepsilon d(x_\varepsilon, x),$$

para cada $x \in M$.

Demostración. Tomemos $u_0 \in M$ tal que $f(u_0) < \infty$. Procediendo por recurrencia, encontramos una sucesión $u_0, u_1, \ldots, u_n, \ldots$ en M que satisface

$$f(u_{n+1}) + \varepsilon d(u_{n+1}, u_n) \leq \operatorname{ínf}\big\{f(x) + \varepsilon d(x, u_n) : x \in M\big\} + \frac{1}{2^n}, \tag{3.6}$$

para $n = 0, 1, 2, \ldots$ De la desigualdad anterior se sigue que

$$f(u_n) \geq \operatorname{ínf}\big\{f(x) + \varepsilon d(x, u_n) : x \in M\big\} \geq f(u_{n+1}) + \varepsilon d(u_{n+1}, u_n) - \frac{1}{2^n} \geq f(u_{n+1}) - \frac{1}{2^n}. \tag{3.7}$$

Entonces,

$$f(u_n) + \frac{1}{2^{n-1}} \geq f(u_{n+1}) + \frac{1}{2^n}, \qquad \text{para cada } n \in \mathbb{N}.$$

Esto implica que la sucesión $\big(f(u_n) + 1/2^{n-1}\big)_n$ es decreciente. Como f está acotada inferiormente, podemos deducir la existencia del límite $\operatorname{lím}_{n\to\infty}\big(f(u_n) + 1/2^{n-1}\big)$. En consecuencia, también existe $\operatorname{lím}_{n\to\infty} f(u_n)$. De la desigualdad (3.7) se sigue que

$$d(u_{n+1}, u_n) \leq \frac{f(u_n) - f(u_{n+1})}{\varepsilon} + \frac{1}{\varepsilon 2^n}, \quad n \in \mathbb{N},$$

y de aquí se tiene que

$$\begin{aligned} d(u_{n+p}, u_n) &\leq \sum_{k=n}^{n+p-1} d(u_{k+1}, u_k) \leq \frac{1}{\varepsilon} \sum_{k=n}^{n+p-1} \big(f(u_k) - f(u_{k+1})\big) + \frac{1}{\varepsilon 2^{n-1}} \\ &= \frac{1}{\varepsilon}\big(f(u_n) - f(u_{n+p})\big) + \frac{1}{\varepsilon 2^{n-1}}, \end{aligned}$$

para cualesquiera $n, p \in \mathbb{N}$. Esto implica que la sucesión $(u_n)_n$ es de Cauchy en (M,d). Como (M,d) es completo, existe $x_\varepsilon \in M$ tal que $\operatorname{lím}_{n\to\infty} d(u_n, x_\varepsilon) = 0$. Usando que f es inferiormente semicontinua, obtenemos que

$$\liminf_{n\to\infty} f(u_n) \geq f(x_\varepsilon).$$

De la desigualdad (3.6) concluimos que

$$f(u_{n+1}) + \varepsilon d(u_{n+1}, u_n) \leq f(x) + \varepsilon d(x, u_n) + \frac{1}{2^n},$$

para cada $x \in M$. Tomando ahora límites, llegamos finalmente a la desigualdad

$$f(x_\varepsilon) \leq \liminf_{n\to\infty} \big(f(u_{n+1}) + \varepsilon d(u_n, u_{n+1})\big) \leq \liminf_{n\to\infty} \left(f(x) + \varepsilon d(x, u_n) + \frac{1}{2^n} \right) = f(x) + \varepsilon d(x, x_\varepsilon),$$

y la prueba termina. □

La siguiente observación elemental será utilizada en los resultados que exponemos a continuación: si $\varphi : C \longrightarrow \mathbb{R}$, definida en un subconjunto convexo C de un espacio vectorial E, es convexa (respectivamente, cóncava), entonces el epígrafo

$$\operatorname{Epi}(\varphi) := \big\{(x, t) \in C \times \mathbb{R} : \varphi(x) \leq t\big\}$$

(respectivamente, subgrafo

$$\operatorname{Sub}(\varphi) := \big\{(x, t) \in C \times \mathbb{R} : \varphi(x) \geq t\big\}\big)$$

es un conjunto convexo.

Lema 3.3.2. *Sean X un espacio de Banach, $C \subset X$ y $\varphi : C \longrightarrow \mathbb{R}$ una función continua en un punto interior $x \in C$. Entonces, para todo $t \in \mathbb{R}$ tal que $\varphi(x) < t$ (respectivamente, $\varphi(x) > t$), el punto (x, t) es un punto interior de* $\operatorname{Epi}(\varphi)$ *(respectivamente, de* $\operatorname{Sub}(\varphi)$*).*

Demostración. Vamos a demostrar sólo la parte concerniente al epígrafo. Si tomamos $t > \varphi(x)$, existe $\varepsilon > 0$ tal que $t - 2\varepsilon > \varphi(x)$. Como φ es continua en x y $x \in \operatorname{int}(C)$, podemos encontrar $\delta > 0$ tal que, si $\|x - z\| < \delta$, entonces $z \in C$ y $\big\{\varphi(x) - \varphi(z)\big\} < \varepsilon$. Tomando $|s - t| < \varepsilon$, tendremos que

$$s > t - \varepsilon > \varphi(x) + \varepsilon > \varphi(z).$$

Luego $\big\{z \in X : \|x - z\| < \delta\big\} \times (t - \varepsilon, t + \varepsilon) \subset \operatorname{Epi}(\varphi)$, y así, $(x, t) \in \operatorname{int}\big(\operatorname{Epi}(\varphi)\big)$. □

Corolario 3.3.3. *Sean X un espacio de Banach, $C \subset X$ y $\varphi : C \longrightarrow \mathbb{R}$ una función continua. Entonces:*

(i) $\operatorname{Epi}(\varphi)$ *y* $\operatorname{Sub}(\varphi)$ *son cerrados en $X \times \mathbb{R}$, y se tiene que*

$$\operatorname{int}\big(\operatorname{Epi}(\varphi)\big) = \big\{(x, t) \in \operatorname{int}(C) \times \mathbb{R} : \varphi(x) < t\big\},$$

$$\operatorname{int}\big(\operatorname{Sub}(\varphi)\big) = \big\{(x, t) \in \operatorname{int}(C) \times \mathbb{R} : \varphi(x) > t\big\}.$$

(ii) Si además C es convexo y φ es convexa (respectivamente, cóncava) entonces $\operatorname{int}\big(\operatorname{Epi}(\varphi)\big)$ *(respectivamente,* $\operatorname{int}\big(\operatorname{Sub}(\varphi)\big)$*) es convexo.*

Teorema 3.3.4 (Bishop-Phelps)**.** *Sea C un conjunto convexo, cerrado y acotado, de un espacio de Banach real X. Entonces, el conjunto de todos los $x^* \in X^*$ que alcanzan su supremo en C es denso en X^*.*

Demostración. Fijamos $f \in X^*$, $\varepsilon > 0$. Vamos a demostrar la existencia de $g \in X^*$ tal que $f+g$ alcanza el ínfimo en C y satisfaciendo $\|g\| \leq \varepsilon$. En virtud del principio variacional de Ekeland 3.3.1, existe $x_\varepsilon \in C$ tal que

$$f(x) \geq f(x_\varepsilon) - \varepsilon\|x - x_\varepsilon\|, \quad \text{para cada } x \in C. \tag{3.8}$$

Escribimos

$$C_1 = \big\{(x,t) \in C \times \mathbb{R} : t \geq f(x)\big\},$$

$$C_2 = \big\{(x,t) \in X \times \mathbb{R} : t < f(x_\varepsilon) - \varepsilon\|x - x_\varepsilon\|\big\}.$$

Como f es continua y lineal (por tanto, convexa), tenemos que $C_1 = \operatorname{Epi}\big(f|_C\big)$ es convexo y cerrado en $X \times \mathbb{R}$ gracias al corolario 3.3.3. Como $\varphi(x) = f(x_\varepsilon) - \varepsilon\|x - x_\varepsilon\|$ es cóncava y continua en X, el corolario 3.3.3 nos asegura de nuevo que $C_2 := \operatorname{int}\big(\operatorname{Sub}(\varphi)\big)$ es convexo y abierto.

La desigualdad (3.8) implica que C_1 y C_2 son disjuntos, y así, el corolario 1.3.24 garantiza la existencia de $h \in (X \times \mathbb{R})^*$, $h \neq 0$, y $a \in \mathbb{R}$, tales que

$$\operatorname*{ínf}_{C_1} h \geq a \geq \sup_{C_2} h.$$

Como h es continua, obtenemos que

$$\operatorname*{ínf}_{C_1} h \geq a \geq \sup_{\overline{C_2}} h. \tag{3.9}$$

Es fácil razonar que $h(x,t) = g(x) + \alpha t$, para cada $(x,t) \in X \times \mathbb{R}$, donde $g \in X^*$ y $\alpha \in \mathbb{R}$ están unívocamente determinados por h. La desigualdad (3.9) conduce a las dos desigualdades:

$$g(x) + \alpha t \geq a, \quad \text{si } x \in C \text{ y } t \geq f(x), \text{ y} \tag{3.10}$$

$$g(x) + \alpha t \leq a, \quad \text{si } x \in X \text{ y } t \leq f(x_\varepsilon) - \varepsilon\|x - x_\varepsilon\|. \tag{3.11}$$

Vamos a demostrar que $\alpha > 0$. Procedamos por reducción al absurdo y supongamos, en primer lugar, que $\alpha = 0$. Dado $x \in X$, tenemos que $\big(x, f(x_\varepsilon) - \varepsilon\|x - x_\varepsilon\|\big) \in \overline{C_2}$ y, consecuentemente, $g(x) \leq a$. Esto implica que $g = 0$, luego h sería cero. Así, $\alpha \neq 0$. Por otro lado, como $(x_\varepsilon, t) \in C_1$ para $t \geq f(x_\varepsilon)$, de (3.10) obtenemos que $g(x_\varepsilon) + \alpha t \geq a$. Entonces, para $t > \operatorname{máx}\big\{0, f(x_\varepsilon)\big\}$, se tiene que $g(x_\varepsilon)/t + \alpha \geq a/t$. Hacemos tender t hacia infinito, lo que permite deducir que $\alpha \geq 0$ y, como ya sabemos que $\alpha \neq 0$, concluimos que $\alpha > 0$.

Sin pérdida de generalidad, podemos suponer que $\alpha = 1$ en las desigualdades (3.10) y (3.11). Como $\big(x_\varepsilon, f(x_\varepsilon)\big) \in C_1 \cap \overline{C_2}$, obtenemos, de (3.10) y (3.11), que

$$g(x_\varepsilon) + f(x_\varepsilon) \geq a \quad \text{y} \quad g(x_\varepsilon) + f(x_\varepsilon) \leq a;$$

por lo tanto,

$$g(x_\varepsilon) + f(x_\varepsilon) = a. \tag{3.12}$$

Ahora bien, dado que $\big(x, f(x)\big) \in C_1$ para todo $x \in C$, a partir de (3.10) y (3.12) se deduce que

$$g(x) + f(x) \geq g(x_\varepsilon) + f(x_\varepsilon),$$

es decir, $g + f$ alcanza su ínfimo en C.

Por otro lado, como $\big(x, f(x_\varepsilon) - \varepsilon\|x - x_\varepsilon\|\big) \in \overline{C_2}$ para todo $x \in X$, obtenemos, de las relaciones (3.11) y (3.12), que

$$g(x) + f(x_\varepsilon) - \varepsilon\|x - x_\varepsilon\| \leq g(x_\varepsilon) + f(x_\varepsilon),$$

lo cual implica que, para cada $x \in X$, $g(x - x_\varepsilon) \leq \varepsilon\|x - x_\varepsilon\|$; en consecuencia, $\|g\| \leq \varepsilon$.

Finalmente, tomando $x^* \in X^*$ arbitrario, ponemos $f = -x^*$, y utilizamos lo demostrado anteriormente para deducir la existencia de $g \in X^*$ de forma que $\|g\| \leq \varepsilon$ y $-x^* + g$ alcanza el ínfimo en C. En tal caso, $x^* - g$ alcanza el supremo en C, y además, $\big\|x^* - (x^* - g)\big\| = \|g\| \leq \varepsilon$, con lo que la prueba termina. □

STOP Mientras que muchos de los resultados que se han presentado hasta ahora sirven, tanto para espacios de Banach reales como complejos, el teorema de Bishop-Phelps no es cierto para espacios complejos, después de un contraejemplo de Lomonosov [76].

Proposición 3.3.5. *Sean X un espacio de Banach, $C \subset X$ convexo y abierto, y $\varphi : C \longrightarrow \mathbb{R}$ una función convexa y continua en algún $x_0 \in C$. Entonces φ es localmente Lipschitz en x_0,* i.e.*, existen $M > 0$ y $\delta > 0$ tales que*

$$\big|\varphi(x) - \varphi(y)\big| \leq M\|x - y\|, \tag{3.13}$$

para cada $x, y \in C$ con $\|x - x_0\| < \delta$ y $\|y - x_0\| < \delta$.

Demostración. Como φ es continua en x_0, existen $M_1 > 0$ y $\delta > 0$ tales que, si $\|x - x_0\| < 2\delta$, entonces $x \in C$ y $\big|\varphi(x)\big| \leq M_1$. Fijemos dos puntos distintos $x, y \in C$ verificando $\|x - x_0\| < \delta$ y $\|y - x_0\| < \delta$, y sea

$$z := y + \frac{\delta}{\alpha}(y - x), \qquad \text{donde } \alpha = \|x - y\|.$$

Al ser $\|z - x_0\| \leq \|y - x_0\| + \delta < 2\delta$, obtenemos que $z \in C$. Por otro lado, la igualdad

$$y = \frac{\alpha}{\alpha + \delta} z + \frac{\delta}{\alpha + \delta} x$$

y la convexidad de φ nos permiten deducir que

$$\varphi(y) \leq \frac{\alpha}{\alpha + \delta}\varphi(z) + \frac{\delta}{\alpha + \delta}\varphi(x).$$

De aquí se obtiene que

$$\varphi(y)-\varphi(x) \leq \frac{\alpha}{\alpha+\delta}\big(\varphi(z)-\varphi(x)\big) \leq \frac{\alpha}{\delta}\big(\varphi(z)-\varphi(x)\big) \leq \frac{2M_1}{\delta}\|x-y\|.$$

Intercambiando los papeles de x e y se tiene la desigualdad (3.13), con $M=2M_1/\delta$. $\square$

Subdiferencial Sean X un espacio de Banach, $C\subset X$ convexo y abierto, $\varphi: C \longrightarrow \mathbb{R}$ una función convexa y $x_0\in C$. El conjunto

$$\partial\varphi(x_0) := \big\{f\in X^* : \varphi(y)-\varphi(x_0) \geq f(y-x_0), \text{ para todo } y\in C\big\}$$

se denomina *subdiferencial* de φ en x_0, véase la figura 3.6. Es claro que, si $\partial\varphi(x_0)\neq\emptyset$, entonces $\partial\varphi(x_0)$ es cerrado y convexo.

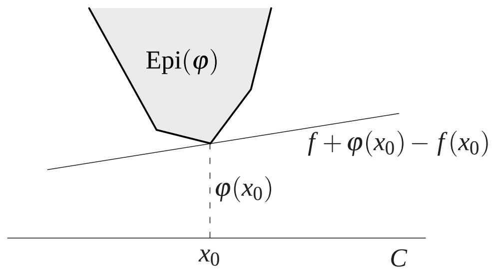


Figura 3.6: Subdiferencial

Teorema 3.3.6. *Sean X un espacio de Banach, $C\subset X$ convexo y abierto, $x_0\in C$, y $\varphi: C\longrightarrow\mathbb{R}$ una función convexa y continua en x_0. Entonces, $\partial\varphi(x_0)$ es un conjunto no vacío y $\sigma(X^*,X)$-compacto. Además, la aplicación $x\rightarrow\partial\varphi(x)$ es localmente acotada,* i.e.*, existen $U\subset C$ entorno de x_0 y $M>0$ tales que $\|f\|\leq M$, si $f\in\partial\varphi(x)$ y $x\in U$.*

Demostración. La proposición 3.3.5 nos asegura que φ es localmente Lipschitz en x_0, y por tanto, podemos garantizar la existencia de $M>0$ y un entorno U de x_0 tal que $U\subset C$, satisfaciendo

$$\big|\varphi(x)-\varphi(y)\big| \leq M\|x-y\|,$$

para cada $x,y\in U$. Si tomamos $x,y\in U$ y $f\in\partial\varphi(x)$, entonces

$$f(y-x)\leq\varphi(y)-\varphi(x)\leq M\|y-x\|.$$

Como U es abierto, esto implica que $\|f\|\leq M$. El teorema de Alaoglu-Bourbaki 1.3.41 nos asegura que $\partial\varphi(x_0)$ es $\sigma(X^*,X)$-compacto.

Vamos a probar ahora que $\partial\varphi(x_0)\neq\emptyset$. Sin pérdida de generalidad podemos suponer que $x_0=0$ y que $\varphi(0)<0$. Pongamos $C_\varphi := \mathrm{Epi}(\varphi)$. C_φ es convexo y, dado que $\varphi(0)<0$, el lema 3.3.2

nos garantiza que 0 es un punto interior de C_φ en $X \times \mathbb{R}$. Denotemos por P_φ el funcional de Minkowski de C_φ. Como $\big(x, \varphi(x)\big) \in C_\varphi$ para todo $x \in C$, obtenemos que $P_\varphi\big(x, \varphi(x)\big) \leq 1$ para todo $x \in C$. Sea $w_0 := \big(0, \varphi(0)\big)$. Vamos a probar que $P_\varphi(w_0) = 1$. Procedamos por reducción al absurdo, y supongamos que $P_\varphi(w_0) < 1$. Entonces, existe $\lambda \in (0,1)$ tal que $w_0/\lambda \in C_\varphi = \mathrm{Epi}(\varphi)$. Consecuentemente, $\varphi(0) \leq \varphi(0)/\lambda$. Pero al ser $\varphi(0) < 0$, $\lambda \geq 1$, lo cual es absurdo. Denotemos por W el subespacio de $X \times \mathbb{R}$ generado por w_0, y sea $h_0 : W \longrightarrow \mathbb{R}$ definido mediante la fórmula $h_0(\alpha w_0) := \alpha$. Como $h_0(w_0) = 1 = P_\varphi(w_0)$, h_0 es homogéneo y P_φ positivamente homogéneo, obtenemos que $h_0(w) \leq P_\varphi(w)$, para todo $w \in W$. En virtud del teorema de Hahn-Banach 1.3.9, existe una aplicación lineal $h : X \times \mathbb{R} \longrightarrow \mathbb{R}$ tal que

$$h(x,t) \leq P_\varphi(x,t), \text{ para todo } (x,t) \in X \times \mathbb{R}, \quad \text{y} \quad h|_W = h_0.$$

Ahora bien, 0 es un punto interior de C_φ, de donde se deduce que h es continua en $X \times \mathbb{R}$, después de la proposición 1.3.5 y el corolario 1.3.6. Entonces, existen $f \in X^*$ y $a \in \mathbb{R}$ tales que

$$h(x,t) = f(x) + at, \quad \text{para todo } (x,t) \in X \times \mathbb{R}.$$

En consecuencia, para todo $(x,t) \in C_\varphi = \mathrm{Epi}(\varphi)$, se verifica que

$$f(x) + at = h(x,t) \leq P_\varphi(x,t) \leq 1. \tag{3.14}$$

Como, por otro lado,

$$a\varphi(0) = h(w_0) = 1 \quad \text{y} \quad \varphi(0) < 0, \tag{3.15}$$

obtenemos que $a < 0$. Dado que, para todo $x \in C$, se tiene $\big(x, \varphi(x)\big) \in C_\varphi = \mathrm{Epi}(\varphi)$, las desigualdades (3.14) y (3.15) implican que $f(x) + a\varphi(x) \leq a\varphi(0)$, de donde, finalmente, se concluye que $\varphi(x) - \varphi(0) \geq g(x - 0)$, siendo $g := (-f)/a$, es decir, $g \in \partial\varphi(0)$. □

3.4 Mejores aproximaciones y el teorema de James

En esta sección ligamos, para un espacio de Banach, la existencia de subespacios proximinales y funcionales que alcanzan la norma, vía el teorema de James 3.4.5. Empezaremos por recordar algunas definiciones básicas y por establecer unas primeras consideraciones de tipo general sobre subespacios proximinales.

Conjuntos proximinales Un subconjunto $D \subset (M,d)$ se dice *proximinal* si todo $x \in M$ tiene una mejor aproximación en D, véase la página 88 para la definición. Obsérvese que, si D es proximinal, entonces D es cerrado. En general, un subconjunto arbitrario $D \subset (M,d)$ no tiene por qué ser proximinal, y si $x \in M$ tiene una mejor aproximación en D, ésta no tiene por qué ser única. En el lema 2.4.6 se ha establecido que todo subconjunto compacto de un espacio de Banach es proximinal. Más en general, los subconjuntos débilmente compactos son proximinales, tal y como se desprende de la prueba de *(ii)* de la proposición 3.4.1.

$\boxed{\textit{Proyección métrica}}$ Sean $(X,\|\cdot\|)$ un espacio de Banach, $Y \subset X$ un subespacio proximinal, y

$$P_Y(x) := \big\{y \in Y : \|y-x\| = d(x,Y)\big\}, \qquad \text{para } x \in X.$$

La aplicación multivaluada $P_Y : X \longrightarrow 2^Y$ se llama *proyección métrica* de X sobre Y. En general, P_Y es multivaluada. Obsérvese que, cuando $Y = K$ es un compacto convexo de un espacio de Banach estrictamente convexo, entonces P_Y es univaluada y continua, lema 2.4.6.

Si $X = H$ es un espacio de Hilbert, el teorema de la proyección, [36, Theorem 2.5], garantiza que todos los subespacios cerrados son proximinales; más aún, si $Y \subset H$ es un subespacio cerrado, entonces $P_Y(x)$ consta de un único elemento para cada $x \in H$, y la *proyección métrica* $P_Y : X \longrightarrow Y$ es lineal y continua. Esto implica que todos los subespacios cerrados de un espacio de Hilbert son complementados; debe observarse que, el hecho de que todos los subespacios cerrados de un espacio de Banach X sean complementados, implica que X es isomorfo a un espacio de Hilbert, [74]. Sin embargo, veremos, como consecuencia del teorema de James 3.4.5, que el que todos los subespacios cerrados de un espacio de Banach sean proximinales caracteriza la reflexividad.

$\boxed{\textit{Aplicaciones multivaluadas superiormente semicontinuas}}$ Sean X e Y dos espacios topológicos. Una aplicación multivaluada $\psi : X \longrightarrow 2^Y$ se dice que es *superiormente semicontinua en un punto* x_0 de X si $\psi(x_0)$ es no vacío y, para cada subconjunto abierto V de Y tal que $\psi(x_0) \subset V$, existe un entorno abierto U de x_0 en X con $\psi(U) \subset V$. La multifunción ψ se dice *superiormente semicontinua* si es superiormente semicontinua en cada punto de X.

$\boxed{\textit{Aplicaciones multivaluadas usco}}$ Si X e Y son dos espacios topológicos, una aplicación multivaluada $\psi : X \longrightarrow 2^Y$ se dice que es *usco* (abreviación en inglés para *upper semi-continuous compact valued*) si es superiormente semicontinua y, para cada $x \in X$, el conjunto $\psi(x)$ es no vacío y compacto.

Proposición 3.4.1. *Sea $\big(X,\|\cdot\|\big)$ un espacio de Banach. Se tienen las siguientes propiedades:*

(i) Si Y es un subespacio proximinal de X, entonces $P_Y(x)$ es acotado, convexo y cerrado.

(ii) Si X es reflexivo, cada subespacio cerrado $Y \subset X$ es proximinal, y la aplicación multivaluada $P_Y : X \longrightarrow 2^Y$ es usco *cuando en X se considera la norma y en Y la topología débil.*

Demostración. Veamos *(i)*. Claramente, $P_Y(x)$ es cerrado en norma. Por otra parte, si $y_0, y_1 \in P_Y(x)$ y $0 \le \lambda, \mu \le 1$, con $\lambda + \mu = 1$, se tiene que

$$\begin{aligned}\big\|(\lambda y_0 + \mu y_1) - x\big\| &= \big\|(\lambda y_0 + \mu y_1) - (\lambda + \mu)x\big\| \\ &\le \lambda\|y_0 - x\| + \mu\|y_1 - x\| = \lambda\, d(x,Y) + \mu\, d(x,Y) = d(x,Y),\end{aligned}$$

lo que prueba que $P_Y(x)$ es convexo. Para terminar, obsérvese que, si $y \in P_Y(x)$, entonces

$$\|y\| \le \|x-y\| + \|x\| = \|x\| + d(x,Y) \le 2\|x\|,$$

y así, $P_Y(x)$ es acotado.

Veamos ahora la demostración de *(ii)*. Primero establecemos que, si $(X,\|\cdot\|)$ es reflexivo, entonces cada subespacio cerrado $Y\subset X$ es proximinal. Efectivamente, dado $x\in X$, si

$$d(x,Y)=\text{ínf}\big\{\|x-y\|:y\in Y\big\},$$

podemos tomar $(y_n)_n\subset Y$ con $d(x,Y)=\text{lím}_n\|x-y_n\|$. Esto implica que $(y_n)_n$ es acotada en norma. Como X es reflexivo, Y también lo es, y así, la sucesión $(y_n)_n$ tiene, gracias al teorema de Eberlein-Šmulian, [51, p. 38-39], una subsucesión $(y_{n_k})_k$ que converge hacia un cierto punto $y_0\in Y$ en la topología débil. En consecuencia, para cada $x^*\in X^*$, $\|x^*\|\leq 1$, se tiene que

$$\big|x^*(x)-x^*(y_0)\big|\leq\big|x^*(x)-x^*(y_{n_k})\big|+\big|x^*(y_{n_k})-x^*(y_0)\big|.$$

Fijado $\varepsilon>0$, existe $k_0\in\mathbb{N}$ tal que, si $k\geq k_0$,

$$\big|x^*(x)-x^*(y_0)\big|\leq\big|x^*(x)-x^*(y_{n_k})\big|+\varepsilon.$$

De aquí,

$$\sup_{x^*\in B_{X^*}}\big|x^*(x)-x^*(y_0)\big|\leq\sup_{x^*\in B_{X^*}}\big|x^*(x)-x^*(y_{n_k})\big|+\varepsilon,$$

es decir, $\|x-y_0\|\leq\|x-y_{n_k}\|+\varepsilon$, de donde, finalmente, deducimos que

$$\|x-y_0\|=\text{lím}_k\|x-y_{n_k}\|=d(x,Y)$$

y, consecuentemente, Y es proximinal.

Utilizando *(i)* y lo ya probado de *(ii)* tenemos que, si $(X,\|\cdot\|)$ es un espacio de Banach reflexivo, entonces, para cada $x\in X$, el conjunto de las mejores aproximaciones $P_Y(x)$ es no vacío, convexo y $\sigma(X,X^*)$-compacto. Vamos a demostrar que la proyección métrica es, en este caso, $\|\cdot\|$-$\sigma(X,X^*)$-usco. El lector puede convencerse de que, como P_Y toma valores débilmente compactos, para probar la semicontinuidad superior de P_Y basta demostrar que, si $(x_n)_n$ es una sucesión en X convergente a x en $\|\cdot\|$, y para cada $n\in\mathbb{N}$ tomamos $y_n\in P_Y(x_n)$, entonces $(y_n)_n$ tiene un punto de aglomeración para la topología débil que pertenece a $P_Y(x)$. Obsérvese que la sucesión $(y_n)_n$ tiene un punto de aglomeración y en la topología débil ya que está acotada; en efecto, $(y_n)_n$ es acotada gracias a la desigualdad

$$\|y_n-x\|\leq\|y_n-x_n\|+\|x_n-x\|=d(x_n,Y)+\|x_n-x\|,$$

y a que $(d(x_n,Y))_n$ converge a $d(x,Y)$. Ahora, concluimos que $y\in P_Y(x)$, ya que

$$\begin{aligned}\big|x^*(x-y)\big|&\leq\big|x^*(x-x_n)\big|+\big|x^*(x_n-y_n)\big|+\big|x^*(y_n-y)\big|\\&\leq\|x-x_n\|+\|x_n-y_n\|+\big|x^*(y_n-y)\big|\\&\leq\|x-x_n\|+d(x_n,Y)+\big|x^*(y_n-y)\big|,\end{aligned}$$

para cada $x^*\in B_{X^*}$. De aquí se sigue que $\big|x^*(x-y)\big|\leq d(x,Y)$, para cada $x^*\in B_{X^*}$; en consecuencia, $\|x-y\|=d(x,Y)$, con lo que $y\in P_Y(x)$ y la prueba termina. □

$\boxed{\textit{Funciones de la primera clase de Baire}}$ Sean X e Y dos espacios topológicos y $f : X \longrightarrow Y$ una función. Se dice que f es de la *primera clase de Baire* si existe una sucesión de funciones continuas $(f_n)_n$ de X en Y tal que $f(x) = \text{lím}_n f_n(x)$, para cada $x \in X$.

$\boxed{\textit{Selectores}}$ Dados dos espacios topológicos X e Y, un *selector* para una aplicación multivaluada $\psi : X \longrightarrow 2^Y$ es una aplicación $f : X \longrightarrow Y$ tal que $f(x) \in \psi(x)$, para cada $x \in X$.

Si X es un espacio de Banach finito-dimensional e $Y \subset X$ un subespacio cerrado, no se tiene, en general, que la proyección métrica P_Y tenga selectores continuos, [17]. Lo mejor que se puede decir es lo que se establece en el corolario siguiente.

Corolario 3.4.2. *Si $\big(X, \|\cdot\|\big)$ es un espacio de Banach reflexivo e Y un subespacio cerrado de X, la proyección métrica $P_Y : X \longrightarrow 2^Y$ tiene un selector que es de la primera clase de Baire.*

Demostración. Basta tener en cuenta que P_Y es *usco* desde la norma de X a la topología débil de Y, y utilizar el teorema 8 de [66]. $\square$

Ya en 1933, S. Mazur observó que, para un funcional lineal continuo $x^* \in X^*$, con $\|x^*\| = 1$, en el espacio de Banach real X, el hiperplano $F_1 := \big\{x \in X : x^*(x) = 1\big\}$ tiene un elemento de *norma mínima* si, y sólo si, x^* alcanza su supremo en la bola unidad cerrada.

Lema 3.4.3. *Sean $(X, \|\cdot\|)$ un espacio normado y $x^* \in X^*$, con $\|x^*\| = 1$. Se tienen las igualdades:*

$$\text{ínf}\big\{\|y\| : x^*(y) = 1\big\} = 1 = \sup\big\{x^*(x) : \|x\| = 1\big\} = \sup\big\{x^*(x) : \|x\| \leq 1\big\}.$$

Demostración. La segunda igualdad es bien conocida, y se sigue directamente de la definición de la norma. Para la primera igualdad, tomamos $y \in X$ con

$$1 = x^*(y) \leq \|x^*\| \, \|y\| = \|y\|.$$

Así, tenemos que $\text{ínf}\big\{\|y\| : x^*(y) = 1\big\} \geq 1$. Por otro lado, de nuevo por la definición de la norma, para cada $\varepsilon > 0$, existe $x_\varepsilon \in X$ tal que $\|x_\varepsilon\| = 1$ y $x^*(x_\varepsilon) \geq 1 - \varepsilon$. Consecuentemente,

$$x^*\left(\frac{x_\varepsilon}{x^*(x_\varepsilon)}\right) = 1, \qquad \left\|\frac{x_\varepsilon}{x^*(x_\varepsilon)}\right\| \leq \frac{1}{1-\varepsilon}.$$

Como $\varepsilon > 0$ es arbitrario, la última desigualdad prueba que $\text{ínf}\big\{\|y\| : x^*(y) = 1\big\} = 1$. $\square$

Proposición 3.4.4. *Sean X un espacio normado y $x^* \in X^*$, con $\|x^*\| = 1$. Las siguientes afirmaciones son equivalentes:*

(i) *x^* alcanza su supremo en B_X.*

(ii) *Para un (y en consecuencia, para cada) $\alpha \in \mathbb{R}$, el hiperplano $H_\alpha = \big\{x \in X : x^*(x) = \alpha\big\}$ es proximinal.*

Demostración. Para $x_0 \notin H_\alpha$, la sustitución

$$y = \frac{x_0 - z}{x^*(x_0) - \alpha}$$

conduce a que

$$\text{ínf}\Big\{\|x_0 - z\| : x^*(z) = \alpha\Big\} = \text{ínf}\Big\{\big|x^*(x_0) - \alpha\big|\,\|y\| : x^*(y) = 1\Big\}.$$

Esto es, la existencia de una mejor aproximación para x_0 en H_α equivale a la existencia de la mejor aproximación de 0 a H_1. El lema 3.4.3 permite ahora terminar la prueba. □

Obsérvese que, si X es reflexivo, entonces cada funcional $x^* \in X^*$ alcanza su supremo en B_X. Pero si tomamos, por ejemplo, $x^* = (1 - 1/n)_n \in \ell^\infty = (\ell^1)^*$, no va a existir ningún elemento $x = (\xi_n)_n \in \ell^1$ con

$$\sum_n |\xi_n| = 1 \quad \text{y} \quad x^*(x) = \xi_n\Big(1 - \frac{1}{n}\Big) = 1.$$

Consecuentemente, x^* no alcanza su supremo.

El teorema de Bishop-Phelps 3.3.4 asegura que, para cada conjunto convexo, cerrado y acotado C de un espacio de Banach real X, el conjunto de todos los $x^* \in X^*$ que alcanzan su máximo en C es denso en X^*. El siguiente notable resultado, de naturaleza similar, es más difícil de probar.

Teorema 3.4.5 (James). *Un conjunto débil cerrado C de un espacio de Banach X es débilmente compacto si, y sólo si, todos los $x^* \in X^*$ alcanzan su máximo en C.*

Demostración. Véase el apartado §6 en [51]. □

Supuesto que hemos demostrado el teorema de James, concluimos la sección con la siguiente caracterización de los espacios de Banach reflexivos.

Corolario 3.4.6. *Para un espacio de Banach X, las siguientes afirmaciones son equivalentes:*

(i) *Todos los subconjuntos débilmente cerrados de X son proximinales.*
(ii) *Todas las variedades afines cerradas de X son proximinales.*
(iii) *Todos los hiperplanos afines cerrados de X son proximinales.*
(iv) *X es reflexivo.*

Demostración. Claramente, se tiene la cadena de implicaciones *(i)* ⇒ *(ii)* ⇒ *(iii)*. Veamos pues *(iii)* ⇒ *(iv)*: de acuerdo con la proposición 3.4.4, si se da *(iii)*, entonces todo funcional $x^* \in X^*$ alcanza la norma en B_X; el teorema de James 3.4.5 nos dice que B_X es débilmente compacto y, por tanto, X es reflexivo. Para la implicación *(iv)* ⇒ *(i)* basta utilizar el mismo tipo de argumentos que ya hemos empleado en el principio de la demostración del apartado *(ii)* de la proposición 3.4.1. □

El teorema de James, 1950-1964. R. C. James anunció, en 1950, que un espacio de Banach X con base es reflexivo si tiene la propiedad de que cada $x^* \in X^*$ alcanza su supremo en la bola unidad cerrada de todas las normas equivalentes. El mismo año, V. Klee demostró este resultado sin requerir la existencia de base. En 1957, James encontró una prueba sofisticada de que si B_X es la bola unidad de un espacio de Banach separable X, entonces B_X es débilmente compacto (*i.e.*, X es reflexivo) si, y sólo si, cada $x^* \in X^*$ alcanza su supremo en B_X. En 1962, Klee conjeturó que esta propiedad podría caracterizar los conjuntos débil cerrados de un espacio de Banach (separable) que son débilmente compactos. Finalmente, en 1964, James publicó el siguiente resultado, del cual el teorema 3.4.5 es un caso particular:

> **Teorema (James, 1964).** *En un espacio localmente convexo quasi-completo E, un conjunto débilmente cerrado A es débilmente compacto si, y sólo si, todos los funcionales $x' \in E'$ alcanzan su supremo en A.*

3.5 Conexión con el proyecto de investigación: fronteras de James

Sea X el espacio de Banach de las funciones continuas $u : \overline{D} \longrightarrow \mathbb{R}$, definidas en el disco unidad cerrado de $\mathbb{C}$, con la propiedad de que la restricción de u al disco abierto D es armónica, dotado de la norma del supremo $\|\cdot\|_\infty$. Si $u \in X$, la fórmula de Poisson, [35, p. 259, Corollary 2.9], permite escribir

$$u(re^{i\alpha}) = \frac{1}{2\pi}\int_0^{2\pi} \frac{1-r^2}{1-2r\cos(\theta-\alpha)+r^2}\,u(e^{i\theta})\,d\theta, \tag{3.16}$$

para cada $0 \le r < 1$ y para todo $\alpha \in \mathbb{R}$, véase la figura 3.7.

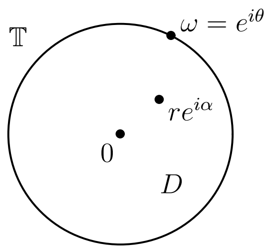


Figura 3.7: Los valores de u en $\mathbb{T}$ determinan sus valores en D

El núcleo de Poisson,

$$P_r(\theta) := \frac{1-r^2}{1-2r\cos\theta+r^2},$$

satisface que $P_r(\theta) > 0$, para cada $\theta \in \mathbb{R}$, y que $1 = \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} P_r(\theta)d\theta$. Si $(u_n)_n$ y u están en X, y suponemos que

$$(u_n)_n \text{ es acotada en } \|\cdot\|_\infty, \text{ y } \lim_n u_n(w) = u(w), \text{ para cada } |w| = 1,$$

entonces, una aplicación directa del teorema de Convergencia Dominada de Lebesgue sobre la ecuación (3.16), para u_n y u, permite concluir que

$$\lim_n u_n(z) = u(z), \qquad \text{para cada } |z| \leq 1.$$

De esta forma, la convergencia puntual de $(u_n)_n$ hacia u sobre la frontera $\mathbb{T}$ de $\overline{D}$, es un *test de convergencia* que implica (¡asegura!) la convergencia puntual de $(u_n)_n$ hacia u sobre todo el disco $\overline{D}$. Esta conexión entre la convergencia en la *frontera* y la convergencia en el *interior* no es particular de este ejemplo, y responde a un comportamiento más general que podemos describir en términos funcionales analíticos:

Si $u \in X$, el valor absoluto $|u|$ es una función continua en $\overline{D}$ y subarmónica en D. Consecuentemente,

$$\|u\|_\infty = \max_{z\in\overline{D}} |u(z)| = \max_{w\in\mathbb{T}} |u(w)|.$$

La igualdad anterior, junto con el hecho de que el problema de Dirichlet tiene solución en D, [35, p. 257, Theorem 2.4], nos permite escribir $(X, \|\cdot\|_\infty) = \big(C(\mathbb{T}), \|\cdot\|_\infty\big)$. Teniendo presente que los puntos extremales de la bola unidad del dual de $C(\mathbb{T})$ son las *deltas de Dirac*, $\{\pm\delta_w : w \in \mathbb{T}\}$, nuestro test de convergencia anterior se puede leer como sigue: *sean* $(u_n)_n$ *una sucesión acotada en* X *y* $u \in X$*. De la condición* $\lim_n x^*(u_n) = x^*(u)$*, para cada* $x^* \in \mathrm{Ext}(B_{X^*})$*, se sigue que* $\lim_n x^*(u_n) = x^*(u)$*, para cada* $x^* \in X^*$. Efectivamente, conocido que el dual de $\big(C(\mathbb{T}), \|\cdot\|_\infty\big)$ se describe como un espacio de medidas numerablemente aditivas (de Radon) vía el teorema de Riesz, [34, Theorem 7.3.5], cuando miramos $(u_n)_n$ en $C(\mathbb{T})$, nuestra afirmación anterior es una consecuencia directa del teorema de Convergencia Dominada de Lebesgue. La interpretación en X se obtiene como consecuencia de que las topologías débiles de X y $C(\mathbb{T})$ se identifican. Obsérvese que, una vez que sabemos que $\lim_n x^*(u_n) = x^*(u)$, para cada $x^* \in X^*$, se tiene, en particular, que $\lim_n u_n(z) = u(z)$, para todo z con $|z| \leq 1$, dado que cada delta de Dirac δ_z, $|z| \leq 1$, es un elemento de X^*.

El siguiente resultado es cierto en general.

Teorema 3.5.1 (Rainwater, [39, p. 155])**.** *Sea* X *un espacio de Banach y sean* $(x_n)_n$ *una sucesión en* X *y* $x \in X$*. Entonces,* $(x_n)_n$ *converge débilmente a* x *si, y sólo si,* $(x_n)_n$ *es acotada en norma y, para cada* $x^* \in \mathrm{Ext}(B_{X^*})$*, se tiene que* $\lim x^*(x_n) = x^*(x)$.

El teorema de Rainwater es una simple consecuencia del teorema de Representación Integral de Choquet que sigue, en combinación con el teorema de Convergencia Dominada de Lebesgue.

$\boxed{\textit{Probabilidades regulares de Borel}}$ Si H es un espacio compacto, representaremos por $\mathscr{P}(H)$ el conjunto de todas las *probabilidades regulares de Borel* en H, *i.e.*, medidas numerablemente aditivas, definidas en la σ-álgebra de Borel $\mathscr{B}(H)$ del espacio H, que son regulares, en el sentido de que

$$\mu(B) = \sup\{\mu(F) : F \subset B,\ F \text{ compacto}\}.$$

Tal y como se observó ya en la demostración del teorema 3.2.1, si $C(H)$ es el espacio de Banach de funciones continuas en H dotado de la norma del supremo, entonces $\mathscr{P}(H)$ es $\sigma(C(H)^*, C(H))$-compacto.

$\boxed{\textit{Baricentro}}$ Sean $E[\mathfrak{T}]$ un *e.l.c.*, H un subconjunto compacto de E y $\mu \in \mathscr{P}(H)$. Un *baricentro* de μ es un vector $x \in E$ que satisface la igualdad

$$x^*(x) = \int_H x^*(h)d\mu(h), \tag{3.17}$$

para cada $x^* \in E'$. Obsérvese que la parte derecha de la igualdad (3.17) está bien definida, ya que $x^*|_H$ es $\mathfrak{T}$-continua y acotada, y por tanto μ-integrable. Además, caso de que μ tenga un baricentro, éste ha de ser necesariamente único, gracias a que E' separa los puntos de E, véase el corolario 1.3.21. En lo que sigue, denotaremos por x_μ el baricentro de $\mu \in \mathscr{P}(H)$, en caso de que exista.

Teorema 3.5.2 (Teorema de representación Integral de Choquet, [39, p. 154])**.** *Sea K un subconjunto convexo, no vacío, compacto y metrizable de un espacio localmente convexo E. Entonces, cada punto de K es el baricentro de una medida de Borel regular de probabilidad concentrada en los puntos extremales de K. Más precisamente, si $x \in K$, entonces existe una medida de Borel regular de probabilidad μ definida en K para la que $\mu(\mathrm{Ext(K)}) = 1$, y para la que, dada cualquier función afín continua f definida en K, se tiene que*

$$f(x) = \int_K f(k)\, d\mu(k).$$

Veremos cómo los resultados anteriores concernientes a convergencia sobre extremales son casos particulares de otros potentes resultados que, a su vez, se siguen del teorema de James 3.4.5, algunos de los cuales exponemos a continuación. Una prueba de un resultado más general que el teorema de Rainwater se ofrece en el teorema 3.5.10 más adelante. Primero necesitamos fijar alguna notación y terminología.

$\boxed{\textit{Conjunto normante y fronteras de James}}$ Sean X un espacio de Banach, y B un subconjunto de B_{X^*}. B es un subconjunto *1-normante* (brevemente, *normante*) de B_{X^*} si, para cada $x \in X$, se tiene que $\|x\| = \sup\{|x^*(x)| : x^* \in B\}$. Se dice que B es una *frontera de James* para B_{X^*} si, para cada $x \in X$, se tiene que $\|x\| = \max\{|x^*(x)| : x^* \in B\}$, *i.e.*, si para cada $x \in X$, existe un elemento $b^* \in B$ tal que $|b^*(x)| = \|x\|$.

Toda frontera de James de B_{X^*} es un subconjunto normante. Sin embargo, el recíproco no es cierto: si X es un espacio de Banach, B_X es un subconjunto normante de $B_{X^{**}}$, pero, después del teorema de James, B_X no es una frontera de $B_{X^{**}}$ a menos que X sea reflexivo. La noción de frontera de James aparece asociada al propio teorema de James. Un ejemplo natural de frontera de James para la bola dual de un espacio $\big(C(K),\|\cdot\|_\infty\big)$ es el conjunto $\{\pm\delta_x : x\in K\}$ de las deltas de Dirac: *toda función real continua en un compacto alcanza su máximo*. Con las identificaciones evidentes, $\mathbb{T}$ es una frontera de James para el espacio $(X,\|\cdot\|_\infty)$ de las funciones continuas en $\overline{D}$ que son armónicas en D. De hecho, después del principio del máximo de Bauer 3.1.11, el ejemplo natural de frontera abstracta de James lo proporciona el conjunto de puntos extremales de B_{X^*}, $\operatorname{Ext}(B_{X^*})$. Hay, sin embargo, fronteras de James que son disjuntas del conjunto de puntos extremales: se comprueba fácilmente que, si Γ es un conjunto no numerable y consideramos el espacio $\big(\ell^1(\Gamma),\|\cdot\|_1\big)$, el conjunto

$$B := \Big\{(x_\gamma)_{\gamma\in\Gamma} : x_\gamma\in\{-1,0,1\} \text{ y } \{\gamma\in\Gamma : x_\gamma\neq 0\} \text{ es numerable}\Big\}$$

es una frontera de James para $B_{\ell^\infty(\Gamma)}$ que es disjunta de $\operatorname{Ext}\big(B_{\ell^\infty(\Gamma)}\big)$, dado que cada punto extremal $(e_\gamma)_{\gamma\in\Gamma}\in B_{\ell^\infty(\Gamma)}$ debe ser, necesariamente, -1 ó 1, para cada coordenada e_γ.

En esta sección prestaremos atención al siguiente problema que aparece en [38, Problem I.2] (véanse también [53, question V.2] y [54]):

> **Problema de la frontera de James:** *Sean X un espacio de Banach, B una frontera de B_{X^*} y H un subconjunto acotado de X. ¿Es cierto que H es $\sigma(X,B)$-compacto si, y sólo si, H es débilmente compacto?*

Obsérvese que, dados un espacio de Banach X y un subconjunto normante cualquiera B de B_{X^*}, no se cumple, en general, que todo subconjunto $\sigma(X,B)$-compacto y acotado de X sea también $\sigma(X,X^*)$-compacto. Por ejemplo, si Y es un espacio de Banach, B_{Y^*} es $\sigma(Y^*,B_Y)$-compacto, donde B_Y es, claramente, un subconjunto normante de $B_{Y^{**}}$; pero no es compacto respecto a la topología $\sigma(Y^*,B_{Y^{**}})$ a menos que Y^* sea reflexivo, después del teorema de James.

Respuestas positivas al problema de la frontera pueden considerarse como tests para obtener la compacidad débil de ciertos subconjuntos de un espacio de Banach. Nótese que, en ocasiones, la topología débil puede ser difícil de tratar si no se dispone de una caracterización adecuada del espacio dual (por ejemplo, en espacios de funciones Bochner integrables, $L^1(\mu,X)$, véase el capítulo 5, o en espacios de medidas numerablemente aditivas, $\operatorname{ca}(\Omega,\Sigma)$, véase [39, p. 86-104]), y es aquí cuando la topología de convergencia sobre una frontera razonable puede ser de gran utilidad en el estudio de la compacidad débil.

Obsérvese, por otro lado, que, cada vez que se tiene una respuesta positiva al problema en cuestión para una frontera $B\subset B_{X^*}$, se deduce, en particular, que si $(x_n)_n$ es acotada en norma y, para cada $x^*\in B$, se tiene que $\lim x^*(x_n)=x^*(x)$, entonces $(x_n)_n$ converge débilmente a x, véase el teorema 3.5.10.

Se conoce que el problema de la frontera tiene respuestas positivas en los siguientes casos:

(i) Cuando H es convexo, [101].

(ii) Si $B = \mathrm{Ext}(B_{X^*})$, [14].

(iii) Si X no contiene una copia isomorfa de $\ell^1(\Gamma)$, con $|\Gamma| = \mathfrak{c}$, [22, 28].

(iv) Cuando $X = C(K)$, dotado con su norma canónica $\|\cdot\|_\infty$, donde K es un espacio compacto arbitrario, [21].

El caso (i) puede ser obtenido a partir del teorema de James, y lo presentaremos a continuación como teorema 3.5.4. La prueba original para (ii) dada en [14] utiliza, entre otras cosas, profundos resultados establecidos en [12]. El caso (iii) se reduce al (i): si $\ell^1(\Gamma) \not\subset X$, $|\Gamma| = \mathfrak{c}$, y $C \subset B_{X^*}$ es 1-normante, se demuestra que, para cada conjunto $H \subset X$ acotado en norma y $\sigma(X,C)$-compacto, la envoltura convexa cerrada $\overline{\mathrm{co}(H)}^{\sigma(X,C)}$ es, de nuevo, $\sigma(X,C)$-compacta, [22, 28]; la clase de espacios de Banach satisfaciendo los requerimientos en (iii) es una amplia clase de espacios de Banach a la que pertenecen los espacios débilmente compactamente generados, los débilmente Lindelöf, los espacios con bola dual sucesionalmente compacta, etc. Las técnicas utilizadas en (iv) son distintas, y extienden de forma natural ideas clásicas de Grothendieck, [58]. Debemos hacer notar que es fácil demostrar que, para cada conjunto Γ, el problema de la frontera tiene también respuesta positiva para el espacio $\ell^1(\Gamma)$ dotado de su norma canónica, véanse [21, 28].

El resto de la sección está dedicado a probar que el problema de la frontera expuesto anteriormente tiene solución positiva si H es un subconjunto convexo y $\sigma(X,B)$-compacto; esto nos servirá para demostrar una versión general del teorema de Rainwater, del que deduciremos que, si H es $\sigma(X,B)$-sucesionalmente compacto, entonces H es débilmente compacto.

Proposición 3.5.3. *Sean X un espacio de Banach, $B \subset B_{X^*}$ un subconjunto normante y $H \subset X$ convexo y $\sigma(X,B)$-compacto. Entonces H es, necesariamente, acotado en norma.*

Demostración. Como consecuencia del teorema de la acotación uniforme se tiene el siguiente resultado, [95, theorem 2.9]: *sean X e Y dos espacios vectoriales topológicos, K un subconjunto compacto convexo de X y Γ una colección de aplicaciones lineales continuas de X en Y. Supongamos que las órbitas $\Gamma(x) = \{\Lambda x : \Lambda \in \Gamma\}$ son subconjuntos acotados de Y, para todo $x \in K$. Entonces, existe un conjunto acotado $B \subset Y$ tal que $\Lambda(K) \subset B$ para toda $\Lambda \in \Gamma$.* Si tomamos X como nuestro espacio de Banach, $Y = \mathbb{R}$ y $\Gamma = B$, se concluye, a partir del resultado anterior, que existe un $M > 0$ tal que

$$b^*(K) \subset [-M, M], \text{ para cada } b^* \in B.$$

La inclusión anterior nos permite concluir, utilizando que B es normante, que, para cada $x \in K$, se tiene que $\|x\| \leq M$, y la demostración termina. □

Teorema 3.5.4 (Simons)**.** *Sean H un subconjunto convexo de un espacio de Banach X y B una frontera de B_{X^*}. Entonces, H es $\sigma(X,B)$-compacto si, y sólo si, H es débilmente compacto.*

Demostración. Hay que demostrar que si H es $\sigma(X,B)$-compacto, entonces H es débilmente compacto. Puede suponerse, y así lo haremos a partir de ahora, que H es absolutamente convexo, debido a que la envoltura absolutamente convexa y $\sigma(X,B)$-cerrada de un conjunto convexo y $\sigma(X,B)$-compacto es también $\sigma(X,B)$-compacta (véase [71, §20.7(8)]).

Tomemos una sucesión arbitraria $(x_n)_n$ en H. La prueba se reduce a demostrar que el conjunto $\{x_n : n \in \mathbb{N}\}$ es débilmente relativamente compacto en X, ya que, en este caso, el teorema de Eberlein-Šmulian nos asegura la existencia de una subsucesión de $(x_n)_n$ que converge débilmente a un punto de H (al ser H débilmente cerrado). Esto prueba que H es débilmente secuencialmente compacto, y ahora, una nueva aplicación del teorema de Eberlein-Šmulian conduce a que H es débilmente compacto.

Para demostrar que $\{x_n : n \in \mathbb{N}\}$ es débilmente relativamente compacto, definimos el operador lineal continuo $S : \ell^1 \longrightarrow X$ dado por

$$S\big((\lambda_n)_n\big) := \sum_{n=1}^{\infty} \lambda_n x_n$$

(nótese que H es acotado). Si definimos $Y := \overline{\operatorname{span}(B)}^{\|\cdot\|}$, entonces $B_{X^*} = \overline{B_Y}^{\sigma(X^*,X)}$, ya que B es normante. Para comprobar que $\{x_n : n \in \mathbb{N}\}$ es débilmente relativamente compacto en X, basta ver que $S^*(B_Y)$ es débilmente relativamente compacto en ℓ^∞, donde $S^* : X^* \longrightarrow \ell^\infty$ es el operador adjunto de S. En efecto, si $\overline{S^*(B_Y)}^{\sigma(\ell^\infty,(\ell^\infty)^*)}$ es débilmente compacto, se cumple que

$$\overline{S^*(B_Y)}^{\sigma(\ell^\infty,\ell^1)} = \overline{S^*(B_Y)}^{\sigma(\ell^\infty,(\ell^\infty)^*)};$$

como, además, S^* es $\sigma(X^*,X)$-$\sigma(\ell^\infty,\ell^1)$-continua, tendríamos que

$$S^*(B_{X^*}) = S^*\left(\overline{B_Y}^{\sigma(X^*,X)}\right) \subset \overline{S^*(B_Y)}^{\sigma(\ell^\infty,\ell^1)} = \overline{S^*(B_Y)}^{\sigma(\ell^\infty,(\ell^\infty)^*)}.$$

Por tanto, $S^*(B_{X^*})$ es débilmente relativamente compacto, de donde se tiene que S^* es un operador débilmente compacto. S es también débilmente compacto, y de aquí deducimos que $S(B_{\ell^1})$ es un conjunto débilmente relativamente compacto en X que contiene a $\{x_n : n \in \mathbb{N}\}$.

Finalmente, comprobaremos que $S^*(B_Y)$ es débilmente relativamente compacto en ℓ^∞. Utilizando el teorema de James 3.4.5, basta ver que cada elemento μ de $B_{(\ell^\infty)^*}$ alcanza su supremo en $S^*(B_Y)$, es decir, que $z := \mu \circ S^* \in X^{**}$ alcanza su supremo en B_Y. Para ello, tomamos una red $(\lambda_n^i)_{i\in D}$ en B_{ℓ^1} que converge a μ en la topología $\sigma\big((\ell^\infty)^*, \ell^\infty\big)$. Para cada $b^* \in B$, $S^*(b^*) = b^* \circ S$ está en ℓ^∞, luego se cumple que

$$b^*\left(\sum_{n=1}^{\infty} \lambda_n^i x_n\right) = \sum_{n=1}^{\infty} \lambda_n^i b^*(x_n) \to z(b^*). \tag{3.18}$$

Al ser H absolutamente convexo y cerrado en norma, se cumple que $y_i = \sum_{n=1}^{\infty} \lambda_n^i x_n$ está en H, para cada $i \in D$. La $\sigma(X,B)$-compacidad de H permite asegurar que cierta subred $(y_{i_k})_k$ de $(y_i)_i$

verifica que $\sigma(X,B)$–$\lim_k y_{i_k} = y$, para cierto $y \in H$. Por lo tanto, si b^* es un elemento de B, se satisface que

$$b^*(x) = \lim_k b^*(y_{i_k}) = \lim_k \left(\sum_{n=1}^{\infty} \lambda_n^{i_k} x_n \right). \tag{3.19}$$

De las ecuaciones (3.18) y (3.19) deducimos que

$$z(b^*) = b^*(y), \quad \text{para cada } b^* \in B.$$

La igualdad anterior se extiende a los elementos b^* de Y y, en consecuencia,

$$\sup\{z(b^*) : b^* \in B_Y\} = \sup\{b^*(y) : b^* \in B_Y\} = \|y\| = b_y^*(y) = z(b_y^*),$$

para cierto $b_y^* \in B \subset B_Y$, donde hemos utilizado que B es una frontera de B_{X^*}. Por lo tanto, z alcanza su supremo en B_Y, y la prueba está concluida. □

Si tomamos como punto de partida el Problema de la Frontera, el teorema anterior sugiere que, para resolver el problema de la frontera, es suficiente establecer un teorema tipo Krein-Šmulian para topologías de la forma $\sigma(X,B)$. El teorema de Krein-Šmulian general 3.2.3 es válido, solamente, para topologías $\mathfrak{T}$ que son compatibles con el par dual $\langle X, X^* \rangle$ y, desafortunadamente, para una frontera de James $B \subset B_{X^*}$, la topología $\sigma(X,B)$ es, en principio, estrictamente más gruesa que la topología débil de X. Los resultados que siguen van encaminados a poner de manifiesto que, en una situación bastante razonable (cuando $(B_{X^*}, \sigma(X^*,X))$ es angélica), el teorema de Krein-Šmulian todavía es cierto para la topología $\sigma(X,B)$.

El lector especializado ya se habrá dado cuenta de que la demostración ofrecida del teorema de Krein-Šmulian para la topología débil 3.2.1, utiliza, sin llegar a darle nombre, la noción de *baricentro*, concepto íntimamente ligado al estudio de conjuntos convexos compactos, tal y como puede constatarse, por ejemplo, en las monografías [33, 39, 86].

Los siguientes resultados pueden encontrarse en [86, Proposition 1.1 y Proposition 1.2].

Teorema 3.5.5. *Sea H un subconjunto compacto de un espacio localmente convexo $E[\mathfrak{T}]$ tal que su envoltura convexa cerrada $\overline{\mathrm{co}(H)}$ es compacta. Entonces, cada $\mu \in \mathscr{P}(H)$ tiene un único baricentro en E.*

Teorema 3.5.6. *Sea H un subconjunto compacto de un espacio localmente convexo $E[\mathfrak{T}]$. Se tiene la igualdad*

$$\overline{\mathrm{co}(H)} = \{x_\mu : \mu \in \mathscr{P}(H),\ \mu \ \textit{tiene un baricentro}\}. \tag{3.20}$$

El siguiente lema muestra la relación entre la existencia de baricentros y el teorema de Krein-Šmulian (préstese atención otra vez a la demostración del teorema 3.2.1).

Lema 3.5.7. *Sea H un subconjunto compacto de un espacio localmente convexo $E[\mathfrak{T}]$. Entonces, $\overline{\mathrm{co}(H)}$ es $\mathfrak{T}$-compacto si, y sólo si, cada probabilidad $\mu \in \mathscr{P}(H)$ tiene un baricentro en E.*

Demostración. Si $\overline{\mathrm{co}(H)}$ es $\mathfrak{T}$-compacto, entonces cada $\mu \in \mathscr{P}(H)$ tiene un baricentro en E, después del teorema 3.5.5. Recíprocamente, supongamos que cada $\mu \in \mathscr{P}(H)$ tiene un baricentro en E y probemos que $\overline{\mathrm{co}(H)}$ es $\mathfrak{T}$-compacto. Dado que la aplicación $\varphi : \mathscr{P}(H) \longrightarrow E$ definida por $\varphi(\mu) = x_\mu$ es $\sigma\big(C(H)^*, C(H)\big)$-$\sigma(E,E')$-continua, obtenemos que $\varphi(\mathscr{P}(H))$ es $\sigma(E,E')$-compacto. Después de la igualdad (3.20), $\overline{\mathrm{co}(H)}$ es también $\sigma(E,E')$-compacto. La $\mathfrak{T}$-compacidad de $\overline{\mathrm{co}(H)}$ se razona de forma similar a como se concluyó la demostración del corolario 3.2.3. $\square$

Conjuntos sucesional y numerablemente compactos Un espacio topológico X se dice que es *numerablemente compacto* si cada cubrimiento abierto numerable del espacio tiene un subcubrimiento finito. Equivalentemente, X es numerablemente compacto si toda sucesión en X tiene un punto de aglomeración en X. El espacio X se dice *sucesionalmente compacto* si toda sucesión en X tiene una subsucesión convergente a un punto de X. Dado el espacio topológico X, un subconjunto A de X se dice que es numerablemente compacto o sucesionalmente compacto si, con la topología inducida, A es un espacio numerablemente compacto o sucesionalmente compacto, respectivamente.

Conjuntos relativamente sucesionalmente y numerablemente compactos Sea X un espacio topológico y $A \subset X$. A es *relativamente numerablemente compacto* (respectivamente, *relativamente sucesionalmente compacto*) si toda sucesión en A tiene un punto de aglomeración en X (respectivamente, una subsucesión convergente a un punto de X). Obsérvese que la definición dada para que A sea relativamente numerablemente compacto *no es equivalente a* que la clausura $\overline{A}$ sea numerablemente compacta, como puede verse en [51, Example 1.2.(9)].

Espacio angélico Un espacio topológico X se dice que es *angélico* si, para cada conjunto relativamente numerablemente compacto $A \subset X$, se tiene que:

(i) A es relativamente compacto.

(ii) Para cada $y \in \overline{A}$, existe una sucesión $(y_n)_n$ en A tal que $\lim_n y_n = y$.

Una buena referencia para el estudio sistemático de los espacios angélicos es la monografía [51]. En los espacios angélicos, los conceptos de subconjunto compacto, sucesionalmente compacto y numerablemente compacto coinciden (también las correspondientes nociones relativas). Entre otros, son espacios angélicos los espacios métricos, el espacio de funciones continuas $\big(C(K), \tau_p(K)\big)$ en un compacto dotado de su topología de convergencia puntual $\tau_p(K)$ sobre K, los espacios normados dotados de sus topologías débiles, etc.

El siguiente lema será utilizado posteriormente.

Lema 3.5.8. *Sean K e Y espacios topológicos y $f : K \longrightarrow Y$ una aplicación. Si K es un espacio compacto angélico, las siguientes afirmaciones son equivalentes:*

(i) f es continua.

(ii) f es sucesionalmente continua, i.e., *para cada sucesión $(x_n)_n$ que converge en K se tiene que $(f(x_n))_n$ converge en Y, y se da la igualdad $f(\lim_n x_n) = \lim_n f(x_n)$.*

Demostración. La implicación *(i)* $\Rightarrow$ *(ii)* es clara. Recíprocamente, supongamos que *(ii)* es cierto. Para demostrar que f es continua basta probar que, para cada subconjunto C de K, se tiene que $f\big(\overline{C}\big) \subset \overline{f(C)}$. Esto último se sigue de la continuidad sucesional de f y de la definición de espacio angélico. $\square$

Con las ideas que utilizamos para demostrar el teorema de Krein-Šmulian 3.2.1 y el lema anterior, podemos probar el siguiente resultado de [30].

Proposición 3.5.9. *Sea X un espacio de Banach con bola dual $\big(B_{X^*}, \sigma(X^*,X)\big)$ angélica. Si B es un subconjunto normante de B_{X^*} y $H \subset X$ es acotado en norma y $\sigma(X,B)$-compacto, entonces $\overline{\mathrm{co}(H)}^{\sigma(X,B)}$ es $\sigma(X,B)$-compacto.*

Demostración. La condición de que $B \subset B_{X^*}$ sea normante implica que su envoltura absolutamente convexa $D := \mathrm{aco}(B)$ sea $\sigma(X^*,X)$-densa en B_{X^*}. Demostramos esto por contradicción, suponiendo que D no es denso en B_{X^*}. En tal caso, podemos tomar un elemento x_0^* de B_{X^*} que no está en $\overline{D}^{\sigma(X^*,X)}$. El teorema de separación 1.3.26 nos permite separar estrictamente el compacto convexo $\{x_0^*\}$ del compacto absolutamente convexo $\overline{D}^{\sigma(X^*,X)}$, de modo que existen un elemento $x \in X = (X^*, \sigma(X^*,X))'$ y escalares $\alpha \in \mathbb{R}$ y $\varepsilon > 0$, tales que

$$\big|x^*(x)\big| \leq \alpha < \alpha + \varepsilon < \big|x_0^*(x)\big|, \text{ para cada } x^* \in D.$$

Al ser B normante, tenemos que $\|x\| < \big|x_0^*(x)\big|$, con $x_0^* \in B_{X^*}$, lo cual es absurdo.

Observemos, por otro lado, que se tiene la igualdad $\sigma(X,B) = \sigma(X,D)$. Para demostrar que $\overline{\mathrm{co}(H)}^{\sigma(X,D)}$ es compacto, tenemos que probar, por el lema 3.5.7, que cada $\mu \in \mathscr{P}\big(H, \sigma(X,D)\big)$ tiene un baricentro en $\big(X, \sigma(X,D)\big)$. Procedemos tal y como hicimos en el teorema 3.2.1, y para $\mu \in \mathscr{P}\big(H, \sigma(X,D)\big)$, definimos

$$T_\mu(x^*) = \int_H x^*|_H \, d\mu,$$

para cada $x^* \in X^*$. Obsérvese que la integral anterior está bien definida, puesto que cualquier $x^* \in B_{X^*} = \overline{D}^{\sigma(X^*,X)}$ y, dado que $\big(B_{X^*}, \sigma(X^*,X)\big)$ es angélica, existe una sucesión $(d_n^*)_n$ en D que converge para $\sigma(X^*,X)$ hacia x^*. Como cada $d_n^*|_H$ es $\sigma(X,D)$-continua, se tiene que x^* es medible respecto de la σ-álgebra de Borel $\mathscr{B}\big(H, \sigma(X,D)\big)$ y, dado que x^* es acotada sobre H y μ una probabilidad, concluimos que $x^*|_H$ es μ-integrable. Por un lado, $T_\mu \in X^{**}$ y, por otro, T_μ restringido a B_{X^*} es $\sigma(X^*,X)$-continua. Para ver esto último basta demostrar que T_μ es $\sigma(X^*,X)$-sucesionalmente continua, lo que se razona con ayuda del teorema de Convergencia Dominada de Lebesgue; el resto de la prueba continúa ahora como en el teorema 3.2.1, con los cambios obvios. Dejamos que el lector cuide los detalles para acabar la demostración. $\square$

Combinando el teorema 3.5.4 con la proposición 3.5.9, demostramos el siguiente resultado, el cual ofrece una solución positiva para el problema de la frontera en el caso de bola dual angélica, [30], y una versión general del teorema de Rainwater, [101].

Teorema 3.5.10. *El problema de la frontera tiene solución positiva para los espacios de Banach con bola dual angélica. Como consecuencia, si X es un espacio de Banach arbitrario y $B \subset B_{X^*}$ una frontera de James, se tiene que:*

(i) Si una sucesión acotada en norma $(x_n)_n$ en X converge hacia x para la topología $\sigma(X,B)$, entonces también converge débilmente.

(ii) Si $H \subset X$ es un conjunto acotado en norma y $\sigma(X,B)$-sucesionalmente compacto, entonces H es débilmente compacto.

Demostración. El hecho de que el problema de la frontera tenga solución para los espacios con bola dual angélica se sigue directamente del teorema 3.5.4 y la proposición 3.5.9.

Para la demostración de *(i)* razonamos como sigue: sea $Y := \overline{\operatorname{span}\{x_n : n \in \mathbb{N}\} \cup \{x\}}^{\|\cdot\|}$; entonces, Y es un espacio de Banach separable (su bola dual es, por tanto, metrizable para la topología débil y, en consecuencia, angélica). Si $i : Y \longrightarrow X$ es la inclusión e $i^* : X^* \longrightarrow Y^*$ el operador adjunto, entonces $i^*(B)$ es una frontera en B_{Y^*}. El conjunto $H = \{x_n : n \in \mathbb{N}\} \cup \{x\}$ es $\sigma(X,B)$-compacto, y también es $\sigma\big(Y, i^*(B)\big)$-compacto. Gracias a que el problema de la frontera tiene solución para espacios con bola dual angélica, se concluye que H es débilmente compacto en Y, o lo que es lo mismo, en X. Como la topología $\sigma(X,B)$ es Hausdorff y más gruesa que la topología débil de X, la sucesión $(x_n)_n$, que es débilmente relativamente compacta, tiene por único punto de aglomeración débil a x, lo que, necesariamente, implica que $(x_n)_n$ converge a x en la topología débil.

Demostremos ahora *(ii)*. Por el teorema de Eberlein-Šmulian, basta probar que cada sucesión $(x_n)_n$ en H tiene una subsucesión débilmente convergente a un punto de H. Como H es $\sigma(X,B)$-sucesionalmente compacto, existe una subsucesión $(x_{n_k})_k$ de $(x_n)_n$ que converge a un punto x de H en la topología $\sigma(X,B)$. Utilizando el apartado *(i)*, se concluye que $(x_{n_k})_k$ converge en la topología débil a x, y la prueba ha terminado. □

Acabamos la sección poniendo de manifiesto que la resolución positiva del problema de la frontera es, de hecho, equivalente a la validez de un resultado tipo Krein-Šmulian para la topología de convergencia puntual asociada a la frontera.

Proposición 3.5.11. *Sean X un espacio de Banach, $B \subset B_{X^*}$ una frontera de James para B_{X^*} y $H \subset X$ un conjunto acotado en norma y $\sigma(X,B)$-compacto. Las siguientes afirmaciones son equivalentes:*

(i) H es débilmente compacto.

(ii) $\overline{\operatorname{co}(H)}^{\sigma(X,B)}$ es $\sigma(X,B)$-compacto.

Demostración. La implicación *(ii)* $\Rightarrow$ *(i)* es clara: si $\overline{\operatorname{co}(H)}^{\sigma(X,B)}$ es $\sigma(X,B)$-compacto, se tiene que $\overline{\operatorname{co}(H)}^{\sigma(X,B)}$ es débilmente compacto, después del teorema 3.5.4; como H es un subconjunto débilmente cerrado de $\overline{\operatorname{co}(H)}^{\sigma(X,B)}$, también se tiene que H es débilmente compacto.

El recíproco *(i)* $\Rightarrow$ *(ii)* es como sigue. Si H es débilmente compacto, las topologías $\sigma(X,X^*)$ y $\sigma(X,B)$ coinciden en H y, por tanto, las σ-álgebras de Borel y las probabilidades regulares de Borel asociadas también coinciden. Dado que el teorema de Krein-Šmulian 3.2.1 se satisface para la topología débil, tenemos que $\overline{\mathrm{co}(H)}^{\sigma(X,X^*)}$ es débilmente compacto, y podemos utilizar el lema 3.5.7 para concluir que cada $\mu \in \mathscr{P}\big(H,\sigma(X,B)\big) = \mathscr{P}\big(H,\sigma(X,X^*)\big)$ tiene un baricentro $x_\mu \in X$. Ahora, el teorema 3.5.6 nos proporciona la igualdad

$$\overline{\mathrm{co}(H)}^{\sigma(X,B)} = \Big\{x_\mu : \mu \in \mathscr{P}\big(H,\sigma(X,B)\big)\Big\} = \Big\{x_\mu : \mu \in \mathscr{P}\big(H,\sigma(X,X^*)\big)\Big\} = \overline{\mathrm{co}(H)}^{\sigma(X,X^*)}.$$

Así, $\overline{\mathrm{co}(H)}^{\sigma(X,B)}$ es débilmente compacto y, en particular, $\sigma(X,B)$-compacto. $\square$

PARA SABER MÁS

- Los libros [33, 39, 86] son excelentes referencias para lecturas adicionales sobre el teorema de Krein-Milman. En los tres libros se expone el teorema de Choquet, que aparece recogido como teorema 3.5.2 en este capítulo. En [39, Chapter IX] se presenta el teorema de Rainwater 3.5.1 como consecuencia del teorema de Choquet. El libro [86] contiene numerosas aplicaciones del teorema de Krein-Milman. El survey [94] ofrece un buen resumen sobre cuestiones básicas relacionadas con el teorema de Krein-Milman.
- Lecturas recomendadas donde encontrar el Principio Variacional de Ekeland y el teorema de Bishop-Phelps son [38, 50] y [85]. El primer libro es una monografía dedicada al estudio de la diferenciabilidad y el renormamiento en espacios de Banach, mientras que el segundo es un texto de propósito general, dedicado a numerosas cuestiones en análisis funcional, que cuenta con una extensa colección de ejercicios. El tercer libro ofrece un compendio sobre diferenciabilidad en espacios de Banach, tanto de funciones, como de medidas: espacios de Asplund *versus* la propiedad de Radon-Nikodým en duales; véase el capítulo 5 más adelante.
- El teorema de James 3.4.5 es un resultado espectacular que combina una propiedad topológica con una propiedad analítica-lineal. Hay distintas demostraciones en el caso separable que son fáciles de entender y gozan de aprecio entre la comunidad matemática, basadas en lo que se conoce como la *desigualdad de Simons*, véanse [101] y [50, p. 84]. Las pruebas conocidas del caso no separable son mucho más complicadas, véase [51]. Es un sentir unánime entre especialistas, que una demostración sencilla del teorema de James sería muy bien apreciada entre la comunidad científica, y ayudaría a entender y resolver algunas cuestiones aún abiertas en temas relacionados, tales como el problema de la *frontera* expuesto en este capítulo. El libro [102] contiene significativas aplicaciones del teorema de James a la teoría general de la aproximación en espacios normados.

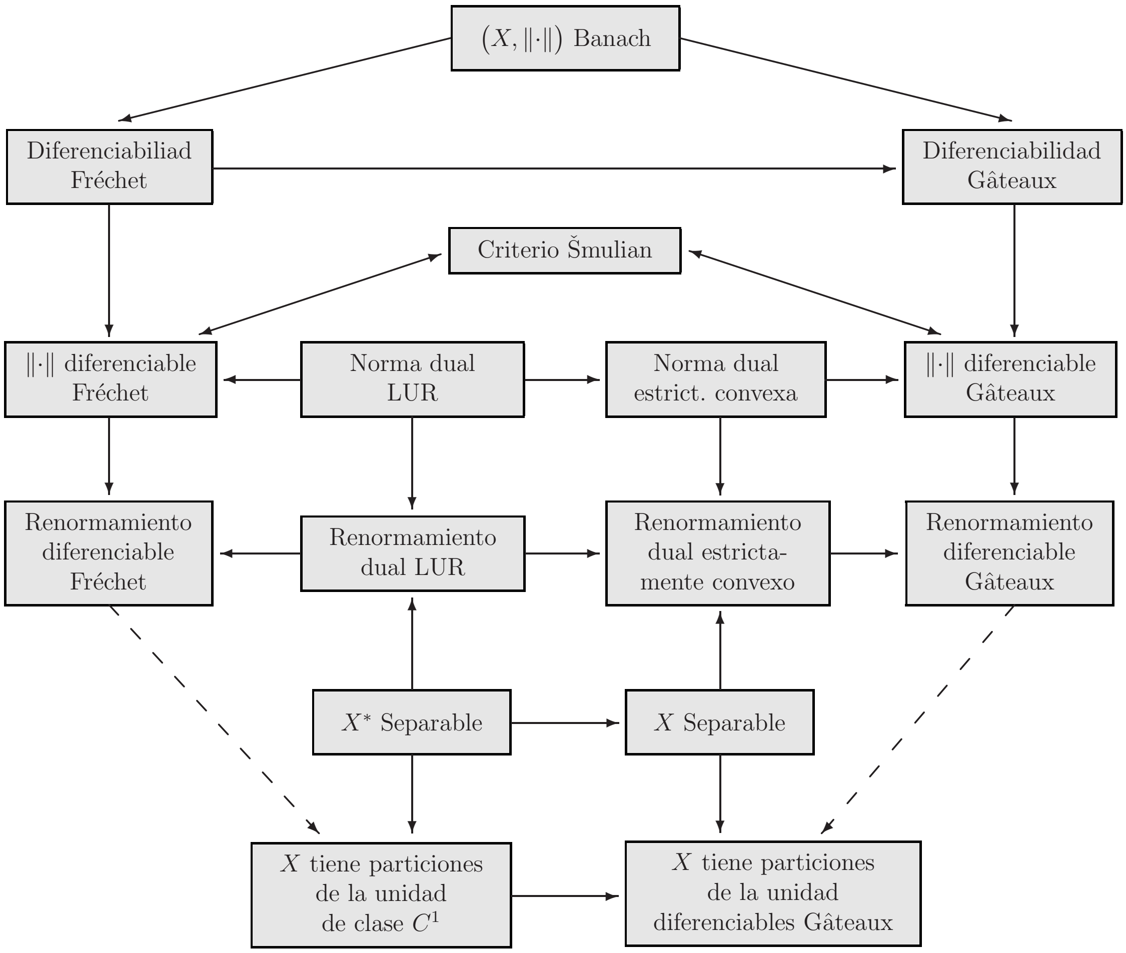
Capítulo 4

Derivadas de Gâteaux y Fréchet

«OBJETIVOS»

- Estudiar las nociones de diferenciabilidad Fréchet y Gâteaux, con especial atención a cuestiones de diferenciabilidad de la norma en espacios de Banach.
- Establecer el criterio de Šmulian con el que se caracteriza la diferenciabilidad de Fréchet y de Gâteaux de la norma de un espacio de Banach.
- Establecer la *dualidad* entre la diferenciabilidad de Gâteaux de una norma (respectivamente, norma dual) y la condición de que la norma dual (respectivamente, norma predual) sea estrictamente convexa.
- Demostrar que, si una norma dual es localmente uniformemente convexa, entonces su norma predual es diferenciable Fréchet.
- Establecer los resultados clásicos de renormamiento localmente uniformemente convexo (respectivamente, estrictamente convexo) en espacios de Banach separables (respectivamente, en duales de espacios separables). Deducir de lo anterior que espacios separables (respectivamente, con dual separable) se pueden renormar con normas diferenciables Gâteaux (respectivamente, Fréchet).
- Demostrar que, en espacios de Banach con dual separable, cada cubrimiento abierto del espacio tiene una partición de la unidad subordinada de clase C^1.

ESTE capítulo está dedicado al estudio de la diferenciabilidad y el renormamiento en espacios de Banach. La diferencial (derivada) de una función es una de las nociones centrales del Análisis matemático. Construir, en un espacio de Banach, una norma equivalente que sea diferenciable Fréchet, proporciona funciones meseta de clase C^1. Renormamiento y diferenciabilidad confluyen de forma natural. Empezamos recordando la definición de diferencial de Fréchet (ampliamente estudiada en los cursos de Cálculo Diferencial en varias variables) como contraposición a la noción de diferencial de Gâteaux (que únicamente implica la existencia de derivadas direccionales): además de estudiar las propiedades básicas de estos conceptos, mostramos sus di-



Cuadro 4.1: Esquema del capítulo *Derivadas de Gâteaux y Fréchet*

ferencias y recordamos su relación con la continuidad de las funciones involucradas, véanse los ejemplos 4.1.1, 4.1.4 y 4.1.5, y las figuras 4.1, 4.2 y 4.3. Estudiamos, en el teorema de Šmulian 4.1.19, una condición necesaria y suficiente para que una norma, en un espacio de Banach, sea diferenciable, y con ella obtenemos la dualidad entre condiciones de *rotundidad* y condiciones de suavidad en los teoremas 4.1.24 y 4.1.25. La segunda parte del capítulo está dedicada a demostrar los resultados clásicos de renormamiento para espacios separables y espacios con duales separables, corolario 4.2.3 y teorema 4.2.4, que utilizamos como herramienta para deducir de lo anterior que espacios separables (respectivamente, con dual separable) se pueden renormar con normas diferenciables Gâteaux (respectivamente, Fréchet), véase el teorema 4.2.8. Este último resultado nos da la clave para demostrar la existencia de particiones de la unidad de clase C^1 en espacios de Banach con dual separable, teorema 4.3.2.

4.1 Diferenciabilidad de Gâteaux y de Fréchet

En esta sección estudiamos las nociones de diferenciabilidad de Fréchet y diferenciabilidad Gâteaux para funciones arbitrarias entre espacios de Banach. En particular, estudiamos condiciones de diferenciabilidad para la propia norma en un espacio de Banach.

Diferenciabilidad de Gâteaux y de Fréchet Sean X, Y espacios normados y U un subconjunto abierto de X. Decimos que una aplicación $\phi : U \longrightarrow Y$ es *diferenciable Gâteaux* (respectivamente, *Fréchet*) en $x \in U$, si existe un operador lineal acotado $T : X \longrightarrow Y$ tal que

$$\phi(x+h) = \phi(x) + Th + r(h), \tag{4.1}$$

donde, para todo $h \in X$, $\lim_{t\to 0}\|r(th)\|/t = 0$ (respectivamente, $\lim_{\|h\|\to 0} r(h)/\|h\| = 0$, es decir, $r(h) = o(\|h\|)$).

Es claro que, si ϕ es diferenciable Fréchet en un punto x, entonces ϕ es diferenciable Gâteaux en dicho punto. Cuando ϕ es diferenciable Gâteaux en x, el operador acotado T proporcionado por la ecuación (4.1) es único. En efecto, si T_1 también satisface que $\phi(x+h) = \phi(x) + T_1h + r_1(h)$ y $\lim_{t\to 0}\|r_1(th)\|/t = 0$, entonces

$$(T - T_1)(h) = \frac{r_1(th) - r(th)}{t},$$

que converge a 0 cuando t tiende hacia 0, para cada $h \in -x + U$. Como U es abierto, obtenemos que $T = T_1$. Si ϕ es diferenciable Gâteaux (respectivamente, Fréchet) en x, el operador acotado T lo denotaremos por $\phi'(x)$ o $D\phi(x)$, y lo llamaremos la *diferencial* (a veces, la derivada) de Gâteaux (respectivamente, Fréchet) de ϕ en x.

Obsérvese que la diferencial de Gâteaux $D\phi(x)$, si existe, asigna, a cada $h \in X$, su *derivada direccional*

$$D_h\phi(x) := \lim_{t\to 0} \frac{\phi(x+th) - \phi(x)}{t}.$$

Ejemplo 4.1.1 (Una función continua con derivadas direccionales en todas las direcciones que no es diferenciable Gâteaux). Sea U un subconjunto abierto de R^n y $\phi : U \longrightarrow \mathbb{R}$. Es inmediato convencerse de que, si ϕ es diferenciable Gâteaux en $x = (x_1, \dots, x_n)$, entonces existen las derivadas parciales $\frac{\partial\phi}{\partial u_i}(x) = D_i\phi(x)$, para cada $i = 1, \dots, n$, y $D\phi(x) = \big(D_1\phi(x), \dots, D_n\phi(x)\big)$.

Por otro lado, si existen las derivadas parciales $D_i\phi(x)$, para cada $i = 1, \dots, n$, la función ϕ no tiene por qué ser diferenciable Gâteaux. Efectivamente, si consideramos $\phi(u) = \sqrt{|u_1u_2|}$, para $u = (u_1, u_2) \in \mathbb{R}^2$, es evidente que $D_1\phi(0) = D_2\phi(0) = 0$. Así, si ϕ tuviese diferencial de Gâteaux en 0, entonces debería ser $D\phi(0) = 0$, pero

$$\frac{r(th)}{t} = \frac{\phi(0+th) - \phi(0)}{t} = \frac{|t|}{t} = \operatorname{sign} t,$$

donde $h = (1, 1)$. Luego ϕ no tiene diferencial de Gâteaux en 0.

Existen, de hecho, funciones con derivadas direccionales en todas las direcciones que no son diferenciables Gâteaux. La función

$$\phi(u) = \begin{cases} 0 & \text{si } u = (u_1, u_2) = (0,0), \\ \dfrac{u_1^2 u_2}{u_1^2 + u_2^2} & \text{si } u = (u_1, u_2) \neq (0,0) \end{cases}$$

(véase la figura 4.1) es continua en $(0,0)$ y, para todo $h \in \mathbb{R}^2$, existe $D_h\phi(0) = f(h)$.

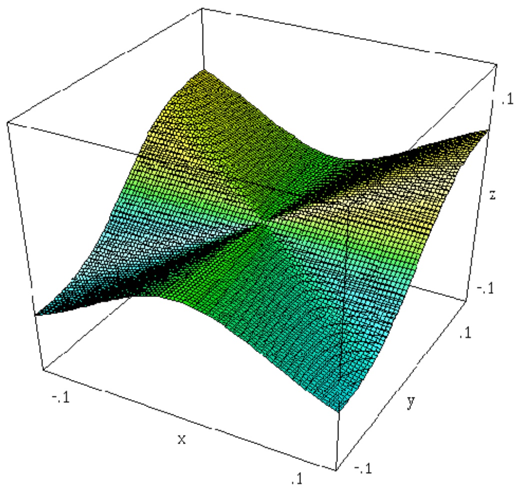


Figura 4.1: Una función continua que no es diferenciable Gâteaux: ejemplo 4.1.1

Sin embargo, la asociación $h \to D_h\phi(0) = f(h)$ no es lineal y, consecuentemente, ϕ no es diferenciable Gâteaux. $\square$

Lema 4.1.2. *Sean X e Y espacios normados, U un subconjunto abierto de X y $\phi : U \longrightarrow Y$ una aplicación. Si ϕ es diferenciable Fréchet en $x \in U$, entonces ϕ es continua en x. Más aún, existe $\delta > 0$ tal que, si $\|h\| < \delta$, entonces*

$$\big\|\phi(x+h) - \phi(x)\big\| \leq \big(\|D\phi(x)\| + 1\big)\|h\|. \tag{4.2}$$

Demostración. Si ponemos $T = D\phi(x)$, se tiene, por definición, que $\phi(x+h) = \phi(x) + Th + r(h)$, con $r(h) = o(\|h\|)$. Del hecho de que $r(h) = o(\|h\|)$, se sigue la existencia de un $\delta > 0$ tal que $\big\|r(h)\big\| \leq \|h\|$, si $\|h\| < \delta$. Entonces,

$$\big\|\phi(x+h) - \phi(x)\big\| \leq \|Th\| + \|h\| \leq \big(\|T\| + 1\big)\|h\|. \qquad \square$$

Lema 4.1.3. *Sean X e Y espacios normados, U un subconjunto abierto de X y $\phi : U \longrightarrow Y$ una aplicación. Si ϕ es diferenciable Gâteaux en $x \in U$, entonces, para cada $h \in -x+U$, existe $\delta = \delta(h) > 0$ tal que, si $|t| < \delta$, entonces*

$$\big\|\phi(x+th) - \phi(x)\big\| \leq \big(\|D\phi(x)\| + 1\big)\|h\|\,|t|\,. \tag{4.3}$$

Demostración. La desigualdad (4.3) se sigue de la definición de diferencial, análogamente a cómo se deducía, en el lema anterior, la desigualdad (4.2). □

Obsérvese que la desigualdad (4.3) no implica que ϕ sea continua en x, tal y como pone de manifiesto el ejemplo que presentamos a continuación.

Ejemplo 4.1.4 (Una función diferenciable Gâteaux que no es continua). Definamos

$$\phi(u) = \begin{cases} 0 & \text{si } u = (u_1, u_2), \ \text{con } u_1 = 0, \\ \dfrac{\|u\|^4}{u_1^2} & \text{si } u = (u_1, u_2), \ \text{con } u_1 \neq 0. \end{cases}$$

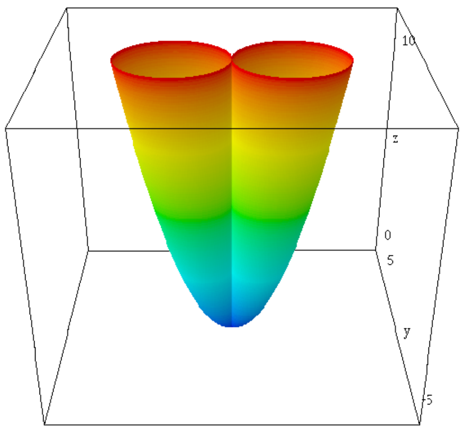


Figura 4.2: Una función diferenciable Gâteaux que no es continua: ejemplo 4.1.4

Tomemos ahora $s \in (0,\infty)$ y pongamos $x_s = (s, s^{1/4})$. Entonces se tiene que $\|x_s\|^2 = s^2 + s^{1/2}$ y, en consecuencia, $\lim_{s\to 0}\|x_s\| = 0$, mientras que $\lim_{s\to 0}\phi(x_s) = \lim_{s\to 0}\big(s + s^{-1/2}\big)^2 = \infty$. Por lo tanto, ϕ no es continua en 0. Veamos que, sin embargo, ϕ es diferenciable Gâteaux en 0. Para cada $x = (x_1, x_2) \in \mathbb{R}^2$,

$$\phi(tx) = \begin{cases} 0 & \text{si } x_1 = 0, \\ t^2\dfrac{\|x\|^4}{x_1^2} & \text{si } x_1 \neq 0; \end{cases}$$

luego $\lim_{t\to 0}\big[\phi(0+tx) - \phi(0)\big]/t = 0$, es decir, $D\phi(0) = 0$. Véase la figura 4.2 □

Ejemplo 4.1.5 (Una función continua, diferenciable Gâteaux, que no es diferenciable Fréchet). Sea $U = \{u \in \mathbb{R}^2 : \|u\| < 1\}$ y consideremos la función

$$\phi(u) = \begin{cases} 0 & \text{si } u = (u_1, u_2), \text{ con } u_1 = 0, \\ \dfrac{\|u\|^3}{|u_1| + u_2^2} & \text{si } u = (u_1, u_2), \text{ con } u_1 \neq 0. \end{cases}$$

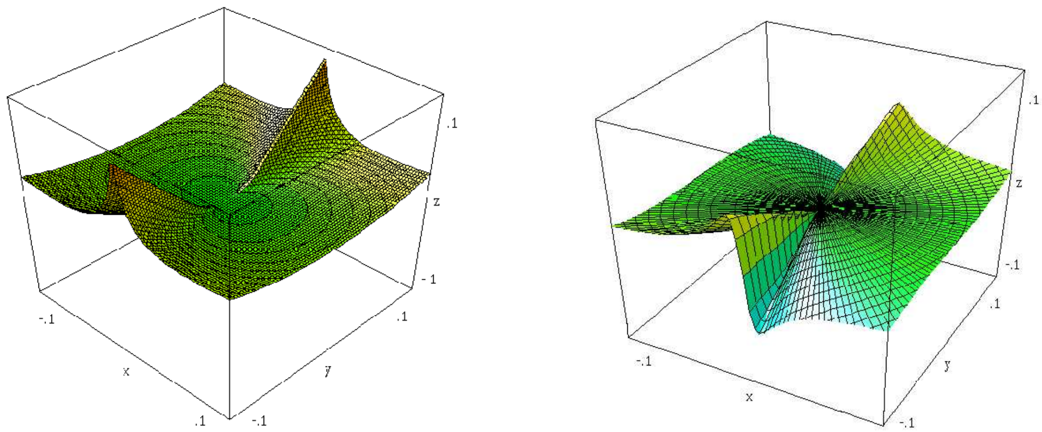


Figura 4.3: Ambas gráficas representan funciones continuas, diferenciables Gâteaux, que no son diferenciables Fréchet: la gráfica de la izquierda corresponde al ejemplo 4.1.5, mientras que la gráfica de la derecha es la de la función $f(u_1, u_2) = (u_1 u_2^2)/(u_1^2 + u_2^4)$, si $(u_1, u_2) \neq (0,0)$, y $f(0,0) = 0$

Veamos que ϕ es una función continua en 0. Para $u = (u_1, u_2) \in U$, $u_1 \neq 0$, se tiene que

$$\frac{u_1^2 + u_2^2}{|u_1| + u_2^2} < 1, \quad \text{luego} \quad \left(\frac{u_1^2 + u_2^2}{|u_1| + u_2^2}\right)^{3/2} < 1.$$

En consecuencia,

$$\phi(u) = \left(\frac{u_1^2 + u_2^2}{|u_1| + u_2^2}\right)^{3/2} \left(|u_1| + u_2^2\right)^{1/2} \leq \left(|u_1| + u_2^2\right)^{1/2}.$$

Como $\phi(u) \geq 0$, obtenemos que $\lim_{\|u\| \to 0} \phi(u) = 0$.

Comprobemos ahora que ϕ es diferenciable Gâteaux en 0. Fijamos $u = (u_1, u_2)$, con $u_1 \neq 0$. Entonces,

$$\frac{\phi(0 + tu) - \phi(0)}{t} = \frac{\phi(tu)}{t} = \frac{|t|^3}{t\,|t|} \frac{\left(u_1^2 + u_2^2\right)^{3/2}}{|u_1| + |t|\, u_2^2} = t \frac{\left(u_1^2 + u_2^2\right)^{3/2}}{|u_1| + |t|\, u_2^2},$$

y dado que

$$\lim_{t\to 0} t\frac{\left(u_1^2+u_2^2\right)^{3/2}}{|u_1|+|t|\,u_2^2}=0,$$

deducimos que ϕ es diferenciable Gâteaux en 0, con $D\phi(0)=0$.

Veamos, por último, que ϕ no es diferenciable Fréchet en 0. Si tomamos $u_t=(t^2,t)$, la igualdad $\|u_t\|=\left(t^4+t^2\right)^{1/2}$ nos asegura que $\lim\limits_{t\to 0}\|u_t\|=0$. Por otro lado tenemos que

$$\phi(u_t)=\frac{\|u_t\|^3}{2t^2}\geq\frac{\|u_t\|^3}{2\|u_t\|^2}=\frac{\|u_t\|}{2}.$$

Así, se tiene que

$$\frac{\big\|r(u_t)\big\|}{\|u_t\|}=\frac{\big\|\phi(u_t)\big\|}{\|u_t\|}\geq\frac{\|u_t\|}{2\|u_t\|}=\frac{1}{2},$$

y por tanto, ϕ no es diferenciable Fréchet en 0. □

La siguiente proposición pone de manifiesto cuáles son las últimas razones por las que se pueden construir ejemplos, como los anteriores, que separan las nociones de diferenciabilidad Gâteaux y Fréchet.

Proposición 4.1.6. *Sean X e Y espacios normados, U un subconjunto abierto de X, $\phi : U \longrightarrow Y$ una aplicación y $x\in U$. Las siguientes afirmaciones son equivalentes:*

(i) ϕ es diferenciable Fréchet en x.

(ii) ϕ es diferenciable Gâteaux en x, y se cumple que

$$D_v\big(\phi(x)\big)=\lim_{t\to 0}\frac{\phi(x+tv)-\phi(x)}{t}\tag{4.4}$$

uniformemente en $S_X=\big\{v\in X:\|v\|=1\big\}$.

Demostración. Veamos que si ϕ es diferenciable Fréchet, entonces el límite (4.4) es uniforme en S_X. Según la definición dada por (4.1), se tiene que

$$\phi(x+h)=\phi(x)+Th+r(h),$$

donde $r(h)=o(\|h\|)$, *i.e.*, para cada $\varepsilon>0$, existe $\delta>0$ tal que la condición $\|h\|<\delta$ implica que $\big\|r(h)\big\|<\varepsilon\|h\|$. Si $|t|<\delta$, entonces, para cada $v\in S_X$, se cumple que $\big\|r(tv)\big\|<\varepsilon\,|t|$; luego $\lim_{t\to 0} r(tv)/t=0$ uniformemente en $v\in S_X$. Como

$$r(tv)=\phi(x+tv)-\phi(x)-tD_v\phi(x),$$

se concluye que el límite en (4.4) es uniforme en S_X. La implicación *(ii)* $\Rightarrow$ *(i)* también es sencilla: si se verifica *(ii)*, la diferencial de Gâteaux para ϕ satisface claramente la condición exigida para ser diferencial de Fréchet. □

Proposición 4.1.7. *Si $\phi, \psi : U \longrightarrow Y$ son diferenciables Gâteaux (respectivamente, Fréchet) en $x \in U$, donde $U \subset X$ es abierto, se verifican:*

(i) $\alpha\phi + \beta\psi$ *tiene diferencial de Gâteaux (respectivamente, Fréchet) en x, para cualesquiera* $\alpha, \beta \in \mathbb{R}$, *y* $D(\alpha\phi + \beta\psi)(x) = \alpha D\phi(x) + \beta D\psi(x)$.

(ii) Si ϕ es una función real, entonces $\phi\psi$ es diferenciable Gâteaux (respectivamente, Fréchet) en x y $D(\phi\psi)(x) = D\phi(x)\psi(x) + \phi(x)D\psi(x)$.

Proposición 4.1.8 (Regla de la cadena)**.** *Sean X, Y, Z espacios normados, $U \subset X$ y $V \subset Y$ abiertos, $\phi : U \longrightarrow V$, $\psi : V \longrightarrow Z$ aplicaciones, $x \in U$ e $y = \phi(x)$. Supongamos que ϕ es diferenciable Fréchet (respectivamente, Gâteaux) en x y ψ es diferenciable Fréchet en y. Entonces, la composición $\psi \circ \phi$ tiene diferencial de Fréchet (respectivamente, Gâteaux) en x, y se tiene la igualdad*

$$D(\psi \circ \phi)(x) = D\psi(y) \circ D\phi(x).$$

Demostración. Como ψ es diferenciable Fréchet en y, sabemos que

$$\psi(y+k) = \psi(y) + D\psi(y)k + r_1(k),$$

donde $r_1(k) = o(\|k\|)$, para $k \in Y$. Al ser ϕ diferenciable Fréchet (respectivamente, Gâteaux) en x, tenemos que ϕ satisface

$$\phi(x+h) = \phi(x) + D\phi(x)h + r(h),$$

donde $\lim_{\|h\|\to 0} r(h)/\|h\| = 0$ (respectivamente, $\lim_{t\to 0} \|r(th)\|/t = 0$), si $h \in X$. Sea $h \in -x + U$ tal que $\phi(x+h) \in V$, y pongamos $k = \phi(x+h) - \phi(x)$, $T = D\psi(y)$, $S = D\phi(x)$. Entonces,

$$\begin{aligned}(\psi \circ \phi)(x+h) - (\psi \circ \phi)(x) &= \psi\big(\phi(x+h)\big) - \psi\big(\phi(x)\big) \\ &= T\big(\phi(x+h) - \phi(x)\big) + r_1(k) \\ &= T\big(Sh + r(h)\big) + r_1(k) = TSh + Tr(h) + r_1(k).\end{aligned} \tag{4.5}$$

Supongamos que ϕ es diferenciable Fréchet en x: Para probar que $\psi \circ \phi$ es diferenciable Fréchet en x y que $D(\psi \circ \phi)(x) = TS$, debemos ver que $Tr(h) + r_1(k) = o(\|h\|)$ en la ecuación (4.5). Fijamos $\varepsilon > 0$. Como $r_1(k) = o(\|k\|)$, existe $\eta > 0$ tal que $\|r_1(k)\|/\|k\| < \varepsilon$, si $\|k\| < \eta$. La definición de diferenciabilidad de Fréchet y el lema 4.1.2 nos proporcionan un $\delta > 0$ para ϕ tal que $\|r(h)\|/\|h\| < \varepsilon$ y $\big\|\phi(x+h) - \phi(x)\big\| \le \big(\|D\phi(x)\| + 1\big)\|h\|$, siempre que $\|h\| < \delta$. Entonces, por (4.5), obtenemos que, para $\|h\| < \min\big\{\delta, \eta/(\|S\|+1)\big\}$, se tiene la siguiente desigualdad, la cual nos permite concluir la demostración de este caso:

$$\begin{aligned}\big\|(\psi \circ \phi)(x+h) - (\psi \circ \phi)(x) - TSh\big\| &\le \|T\|\,\big\|r(h)\big\| + \big\|r_1(k)\big\| < \|T\|\varepsilon\|h\| + \varepsilon\|k\| \\ &\le \|T\|\varepsilon\|h\| + \big(\|S\|+1\big)\varepsilon\|h\| \\ &= \big(\|T\| + \|S\| + 1\big)\varepsilon\|h\|.\end{aligned}$$

Supongamos que ϕ es diferenciable Gâteaux en x: Probemos que $\psi\circ\phi$ es diferenciable Gâteaux en x y que $D(\psi\circ\phi)(x)=TS$. Tomamos $\varepsilon>0$. Al igual que en el caso anterior, como $r_1(k)=o(\|k\|)$, existe $\eta>0$ tal que $\|r_1(k)\|/\|k\|<\varepsilon$, si $\|k\|<\eta$. Fijamos $h\in X$. Al ser ϕ diferenciable Gâteaux en x, existe $\delta>0$ tal que $\|r(th)\|/|t|<\varepsilon\|h\|$ y

$$\big\|\phi(x+th)-\phi(x)\big\|\leq\big(\|D\phi(x)\|+1\big)\|h\|\,|t|,$$

si $|t|<\delta$, véase el lema 4.1.3. Sustituyendo h por th en la igualdad (4.5), obtenemos que, para $|t|<\text{mín}\big\{\delta,\eta/\big[(\|S\|+1)\|h\|\big]\big\}$, se da la desigualdad

$$\begin{aligned}\big\|(\psi\circ\phi)(x+th)-(\psi\circ\phi)(x)-TS(th)\big\| &< \|T\|\varepsilon\|h\|\,|t|+\varepsilon\|k\|\\ &\leq \|T\|\varepsilon\|h\|\,|t|+\big(\|S\|+1\big)\varepsilon\|h\|\,|t|\\ &= \big(\|T\|+\|S\|+1\big)\varepsilon\|h\|\,|t|,\end{aligned}$$

y la prueba termina. □

$\boxed{\textit{Incremento } \Delta^2\phi}$ Sean X e Y espacios normados, U un subconjunto abierto de X, $\phi:U\longrightarrow Y$ una aplicación y $x\in U$. En lo que sigue, para $h\in -x+U$, utilizaremos la notación

$$\Delta^2\phi(x,h)=\phi(x+h)+\phi(x-h)-2\phi(x).$$

Lema 4.1.9. *Sean X e Y espacios normados, U un subconjunto abierto de X, $\phi:U\longrightarrow Y$ una aplicación y $x\in U$. Si ϕ es diferenciable Fréchet (respectivamente, Gâteaux) en x, entonces*

$$\Delta^2\phi(x,h)=o(\|h\|)\quad\big(\textit{respectivamente, para cada } h\in -x+U \textit{ fijo, } \Delta^2\phi(x,th)=o(t)\big).$$

Demostración. Si ϕ es diferenciable Fréchet en x, obtenemos que

$$\phi(x+h)-\phi(x)=D\phi(x)h+o(\|h\|),\qquad \phi(x-h)-\phi(x)=-D\phi(x)h+o(\|h\|).$$

Sumando las igualdades anteriores, se tiene lo deseado. En el caso de diferenciabilidad Gâteaux, la prueba se razona de forma similar. □

STOP Para cualquier norma $\|\cdot\|$ en X, si tomamos como $\phi:X\longrightarrow\mathbb{R}$ la propia norma, esto es, $\phi(x):=\|x\|$, para $x\in X$, se tiene que $\Delta^2\phi(0,th)=2|t|\,\|h\|$, y por tanto, la norma $\|\cdot\|$ no tiene diferencial de Gâteaux en 0.

Lema 4.1.10. *Sea $\big(X,\|\cdot\|\big)$ un espacio normado. La norma, como aplicación $\phi(x):=\|x\|$, es diferenciable Fréchet (res. Gâteaux) en $x\neq 0$ si, y sólo si,*

$$\Delta^2\phi(x,h)=o(\|h\|)\quad\big(\textit{respectivamente, para cada } h \textit{ fijo, } \Delta^2\phi(x,th)=o(t)\big).$$

Si ϕ es diferenciable Gâteaux en x, entonces $D\phi(x)=x^*$, donde $x^*\in X^*$, $x^*(x)=\|x\|$ y $\|x^*\|=1$.

Demostración. Sea $x^* \in X^*$, $x^*(x) = \|x\|$ y $\|x^*\| = 1$. Vamos a probar que

$$0 \leq \|x+h\| - \|x\| - x^*(h) \leq \Delta^2\phi(x,h). \tag{4.6}$$

Pero tenemos

$$\begin{aligned} x^*(h) = x^*(x+h) - x^*(x) &\leq \|x+h\| - x^*(x) = \|x+h\| - \|x\| \\ &= \Delta^2\phi(x,h) - \|x-h\| + \|x\| \\ &\leq \Delta^2\phi(x,h) - x^*(x-h) + \|x\| = \Delta^2\phi(x,h) + x^*(h), \end{aligned} \tag{4.7}$$

lo cual implica (4.6). Ahora, por (4.6), podemos deducir que ϕ es diferenciable Fréchet (respectivamente, Gâteaux) en x si se satisface

$$\Delta^2\phi(x,h) = o(\|h\|) \quad \big(\text{respectivamente, para } h \text{ fijo, } \Delta^2\phi(x,th) = o(t)\big).$$

En este caso, $D\phi(x) = x^*$. □

Ejemplo 4.1.11 (Condición suficiente para la diferenciabilidad Fréchet). Sea $\varphi : U \longrightarrow \mathbb{R}$ una aplicación definida en un abierto $U \subset \mathbb{R}^n$. Si se supone que, en un entorno abierto $U_x \subset U$ del punto $x \in U$, existen las derivadas parciales $D_1\phi(u), D_2\phi(u), \ldots, D_{n-1}\phi(u)$ y son continuas en x, y que también existe $D_n\phi(x)$, entonces la aplicación φ es diferenciable Fréchet en x y se tiene que, [2, Teorema 12.11],

$$D\varphi(x) = \big(D_1\phi(x), D_2\phi(x), \ldots, D_n\phi(x)\big).$$

Más en general, [20, Proposición 3.7.2], si reemplazamos $\mathbb{R}$ por un espacio normado Y, las mismas hipótesis sobre φ implican también que φ es diferenciable Fréchet, y la aplicación dada por

$$D\varphi(x)h = \sum_{j=1}^{n} h_j D_j\varphi(x),$$

para $h \in \mathbb{R}^n$, es la diferencial de φ en x. □

Ejemplo 4.1.12 (Una norma diferenciable Fréchet en todo $x \neq 0$). Sea $X = \big(H, \langle\cdot,\cdot\rangle\big)$ un espacio de Hilbert. Como, para cada $x, h \in X$, se tiene que

$$\|x+h\|^2 - \|x\|^2 = 2\langle x,h\rangle + \|h\|^2,$$

de la definición de diferencial de Fréchet obtenemos que $\varphi(x) := \|x\|^2$ es diferenciable Fréchet y que $D\varphi(x) = 2\langle x,\cdot\rangle$. Consecuentemente, la norma $\phi(x) := \|x\| = \sqrt{\varphi}(x)$ es, en virtud de la regla de la cadena 4.1.8, diferenciable Fréchet para cada $x \neq 0$, y la expresión

$$D\phi(x) = D\sqrt{\varphi}(x) = \frac{1}{2}\frac{D\varphi}{\sqrt{\varphi}}(x) = \frac{\langle x,\cdot\rangle}{\|x\|}$$

proporciona su diferencial de Fréchet en x. □

Ejemplo 4.1.13 (Una norma no diferenciable Fréchet en ningún punto). La norma canónica de ℓ^1, $\left\|(x_n)_n\right\|_1 = \sum_n |x_n|$, no es diferenciable Fréchet en ningún punto $x \in \ell^1$. En efecto, si $\{e_n\}_{n\in\mathbb{N}}$ es la base canónica de ℓ^1, entonces, para cada $x \in \ell^1$,

$$\sup_{n\in\mathbb{N}}\left\{ \|x+te_n\|_1 + \|x-te_n\|_1 - 2\|x\|_1 \right\} \geq 2\,|t|\,.$$

En virtud del lema 4.1.10, se tiene que $\|\cdot\|_1$ no es diferenciable Fréchet en x. □

Ejemplo 4.1.14 (Una norma diferenciable Gâteaux que no es diferenciable Fréchet). Veamos que $\|\cdot\|_1$ en ℓ^1 tiene diferencial de Gâteaux en $x = (x_n)_{n=1}^{\infty}$ si $x_n \neq 0$, para todo $n \in \mathbb{N}$. Para cualesquiera $a, b \in \mathbb{R}$ se tiene que

$$|a+b| + |a-b| - 2\,|a| = \begin{cases} 0 & \text{si } |a| \geq |b|\,, \\ 2\big(|b| - |a|\big) & \text{si } |a| < |b|\,. \end{cases}$$

Sean $y = (y_n)_{n=1}^{\infty} \in \ell^1$ y $\varepsilon > 0$. Elegimos $N \in \mathbb{N}$ tal que $\sum_{n=N+1}^{\infty} |y_n| < \varepsilon$, y sea $\delta > 0$ tal que $\delta\,|y_n| \leq |x_n|$, para cada $n = 1, \ldots, N$. Entonces, si $|t| < \delta$, tenemos que

$$\|x+ty\|_1 + \|x-ty\|_1 - 2\|x\|_1 \leq 2 \sum_{n=N+1}^{\infty} |ty_n| = 2\,|t| \sum_{n=N+1}^{\infty} |y_n| < 2\varepsilon\,|t|\,.$$

De nuevo, en virtud del lema 4.1.10, la norma $\|\cdot\|_1$ tiene diferencial de Gâteaux en x. □

Proposición 4.1.15. *Sea X un espacio de Banach de dimensión finita. Si una norma en X es diferenciable Gâteaux en algún punto $x \in X$, entonces es diferenciable Fréchet en ese punto.*

Demostración. Supongamos que $\|\cdot\|$ no tiene diferencial de Fréchet en $x \in X$. Entonces, existen $\delta > 0$ y sucesiones $(y_n)_n$ en B_X y $(t_n)_n$ en $\mathbb{R}^+$, tales que $\lim_{n\to\infty} t_n = 0$ y

$$\frac{\|x+t_ny_n\| + \|x-t_ny_n\| - 2\|x\|}{t_n} \geq \delta. \tag{4.8}$$

Como la bola B_X es compacta, podemos suponer, sin pérdida de generalidad, que existe $y \in B_X$ tal que $\lim_{n\to\infty}\|y_n - y\| = 0$. Teniendo en cuenta que

$$\frac{\Big|\|x \pm t_ny_n\| - \|x \mp t_ny\|\Big|}{t_n} \leq \|y_n - y\|,$$

de la desigualdad (4.8) concluimos que existe $N \in \mathbb{N}$ tal que, para cada $n > N$,

$$\frac{\|x+t_ny\| + \|x-t_ny\| - 2\|x\|}{t_n} > \frac{\delta}{2},$$

lo cual significa que $\|\cdot\|$ no es diferenciable Gâteaux en x, y la prueba termina. □

$\boxed{\textit{Soporte de una función}}$ Si X es un espacio topológico y $\varphi : X \longrightarrow \mathbb{R}$ una función, el *soporte* de φ es el conjunto

$$\operatorname{sop}(\varphi) := \overline{\{x \in X : \varphi(x) \neq 0\}}.$$

$\boxed{\textit{Función meseta}}$ Sea X un espacio de Banach. Una función $\theta : X \longrightarrow \mathbb{R}$ se dice que es una *función meseta* si θ no es idénticamente nula y $\operatorname{sop}(\theta)$ es un conjunto acotado en X.

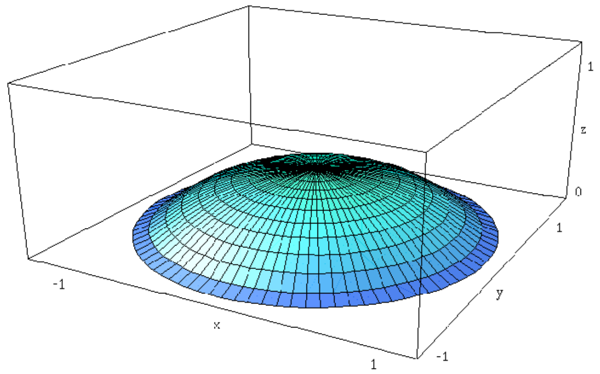


Figura 4.4: Función meseta en $\mathbb{R}^n$

Obsérvese que, en espacios de dimensión finita, esto es, en $X = \mathbb{R}^n$, existen *muchas* funciones meseta de clase C^∞. Ciertamente, si fijamos, por ejemplo, la norma euclídea en $\mathbb{R}^n$, la función

$$\theta(u) = \begin{cases} 0 & \text{si } \|u\|_2 \geq 1, \\ e^{\gamma \frac{1}{\|u\|_2^2 - 1}} & \text{si } \|u\|_2 < 1, \end{cases}$$

para $\gamma > 0$, es de clase C^∞ en $\mathbb{R}^n$ y su soporte está incluido en la bola unidad de $\mathbb{R}^n$. De hecho, las funciones de clase infinito con soporte compacto en $\mathbb{R}^n$ son tan abundantes como para poder tener con ellas el lema de separación tipo Urysohn que sigue: *si K y F son dos partes disjuntas de $\mathbb{R}^n$, K compacto y F cerrado, entonces existe una función ψ de clase C^∞, con soporte compacto, que satisface las propiedades siguientes:*

(i) $0 \leq \psi(x) \leq 1$, *para todo* $x \in \mathbb{R}^n$.

(ii) $\psi = 0$ *en un entorno de* F.

(iii) $\psi = 1$ *en un entorno de* K.

Gracias a este lema de separación, se puede garantizar la existencia de las llamadas *particiones de la unidad* de clase C^∞ en $\mathbb{R}^n$: *se denomina partición de la unidad de clase C^∞ subordinada a un recubrimiento abierto $\{V_i\}_{i\in I}$ de un abierto $U \subset \mathbb{R}^n$, a una familia de* *funciones* $\{\psi_i\}_{i\in I}$ *definidas en* U*, e indicada en el mismo conjunto de índices, que tiene las siguientes propiedades:*

(i) Para cada $i \in I$*, la función* ψ_i *es de clase* C^∞ *en* U*, tiene soporte contenido en* V_i *y* $\psi_i(x) \geq 0$*, para cada* $x \in U$*.*

(ii) Sobre cada compacto $K \subset U$*, sólo un número finito de* ψ_i *no son idénticamente nulas.*

(iii) Para cada $x \in U$*, se tiene* $\sum_{i\in I} \psi_i(x) = 1$*.*

Si U es un abierto de $\mathbb{R}^n$, entonces cada cubrimiento abierto de U tiene una partición de la unidad de clase C^∞ en U subordinada a dicho cubrimiento, véase [95, p. 152-153].

En dimensión infinita, la existencia de funciones meseta diferenciables no está garantizada.

Ejemplo 4.1.16 (Espacios de Banach sin funciones meseta diferenciables). En el espacio ℓ^1 (respectivamente, $\ell^1(\Gamma)$, con Γ no numerable), no existe una función meseta diferenciable Fréchet (respectivamente, continua y diferenciable Gâteaux) en todos los puntos.

Supongamos que θ es una función meseta continua en un espacio de Banach $\big(X, \|\cdot\|\big)$ arbitrario. Definimos

$$\varphi(x) = \begin{cases} \frac{1}{\theta^2(x)} - \|x\| & \text{si } \theta(x) \neq 0, \\ \infty & \text{si } \theta(x) = 0. \end{cases}$$

Es claro que φ es acotada inferiormente, y se puede comprobar de manera sencilla que es inferiormente semicontinua, siendo $\varphi(x) < \infty$ para cada $x \in X$ con $\theta(x) \neq 0$. Hacemos uso ahora del Principio Variacional de Ekeland, teorema 3.3.1, el cual nos asegura la existencia de $y \in X$, con $\theta(y) \neq 0$, satisfaciendo

$$\varphi(x) \geq \varphi(y) - \frac{\|x-y\|}{3}, \quad \text{para cada } x \in X.$$

Tomamos $x = y \pm h$ y sumamos las expresiones obtenidas, concluyendo entonces que

$$\varphi(y+h) + \varphi(y-h) \geq 2\varphi(y) - 2\frac{\|h\|}{3}.$$

Esto implica que

$$\Delta^2 \frac{1}{\theta^2}(y,h) \geq \|y+h\| + \|y-h\| - 2\|y\| - \frac{2}{3}\|h\|. \tag{4.9}$$

Caso $X = \ell^1$*:* Escribamos $y = (y_n)_n$ en ℓ^1, y tomemos una sucesión $(t_n)_n$ en $\mathbb{R}^+$ de forma que $\lim_{n\to\infty} t_n = 0$ y $t_n > 2|y_n|$. Sea $h_n = t_n e_n$, donde $\{e_n\}_n$ es la base canónica de ℓ^1. Como

$$\begin{aligned} \|y+h_n\| + \|y-h_n\| - 2\|y\| &= |y_n + t_n| + |y_n - t_n| - 2|y_n| \\ &= t_n + y_n + t_n - y_n - 2|y_n| \\ &= 2\big(t_n - |y_n|\big) > 2\left(t_n - \frac{t_n}{2}\right) = t_n = \|h_n\|, \end{aligned}$$

de (4.9) se deduce que $\Delta^2\theta^{-2}(y,h_n) \geq \|h_n\|/3$. Además, al ser $\|h_n\| = t_n$ y $\lim_{n\to\infty} t_n = 0$, en virtud del lema 4.1.9 obtenemos que θ^{-2} no tiene diferencial de Fréchet en y. Finalmente, como t^{-1} es diferenciable Fréchet para $t \neq 0$, deducimos que θ^2 no tiene diferencial de Fréchet en y, y así, podemos concluir que θ no tiene diferencial de Fréchet en y.

Caso $X = \ell^1(\Gamma)$: Dado que el soporte de y es numerable, existe $\alpha \in \Gamma \setminus \operatorname{sop}(y)$. Definimos

$$h(\gamma) = \begin{cases} 1 & \text{si } \gamma = \alpha, \\ 0 & \text{si } \gamma \neq \alpha. \end{cases}$$

Como $\|y \pm th\| = \|y\| + |t|$, de (4.9) se deduce que $\Delta\theta^{-2}(y,th) \geq 4|t|/3$. En virtud de la la regla de la cadena 4.1.8, obtenemos que θ^{-2} no es diferenciable Gâteaux en y. $\square$

Tenemos el siguiente resultado positivo.

Proposición 4.1.17. *Si un espacio de Banach X admite una norma equivalente que es diferenciable Fréchet (respectivamente, Gâteaux) en cada $x \neq 0$, entonces X tiene una función meseta que es diferenciable Fréchet (respectivamente, Gâteaux) en todo $x \in X$.*

Demostración. Supongamos que $\|\cdot\|$ es una norma en X, y sea $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ una función tal que $\varphi'(t)$ existe para todo $t \in \mathbb{R}$ y $\varphi(t) = 0$ si $|t| > a$, para algún $a \in \mathbb{R}$; por ejemplo,

$$\varphi(t) = \begin{cases} (t^2-1)^2 & \text{si } |t| \leq 1, \\ 0 & \text{si } |t| > 1. \end{cases}$$

Entonces, $\theta(x) = \varphi\left(\|x\|^2\right)$ es diferenciable Fréchet (respectivamente, Gâteaux) para cada $x \in X$, si $\|\cdot\|$ es diferenciable Fréchet (respectivamente, Gâteaux) para cada $x \neq 0$, véanse las proposiciones 4.1.7 y 4.1.8. $\square$

Función ε-soportada Sean X un espacio de Banach, $\varphi : X \longrightarrow \mathbb{R} \cup \{\infty\}$ una función para la cual $\operatorname{Dom}(\varphi) := \{x \in X : \varphi(x) < \infty\}$, y $\varepsilon > 0$. Decimos que φ está *ε-soportada* en $v \in \operatorname{Dom}(\varphi)$, si existen $x^* \in X^*$ y $\delta > 0$ tales que

$$\varphi(x) \geq \varphi(v) + x^*(x-v) - \varepsilon\|v-x\|,$$

para todo $x \in B(v,\delta) = \{x \in X : \|v-x\| < \delta\}$.

Teorema 4.1.18. *Sea X un espacio de Banach para el cual existe una función $\theta : X \longrightarrow [0,\infty)$, diferenciable Fréchet, tal que $\theta(0) > 0$ y $\theta(x) = 0$ si $x \notin B_X$. Sea $\varphi : X \longrightarrow \mathbb{R} \cup \{\infty\}$ una función inferiormente semicontinua. Entonces, para cada $\varepsilon > 0$, el conjunto de puntos donde φ está ε-soportada es denso en* $\operatorname{Dom}(\varphi)$.

Demostración. Sean $u \in \mathrm{Dom}(\varphi)$ y $\alpha > 0$. Dado $\varepsilon > 0$, encontraremos un $x_\varepsilon \in B(u,\alpha)$ tal que φ está $\varepsilon-$soportada en x_ε. Como φ es inferiormente semicontinua, existe $\beta > 0$ de forma que φ es acotada inferiormente en $B(u,\beta)$. Definimos

$$\psi(x) = \frac{1}{\theta\left(\frac{x-u}{\beta}\right)}.$$

Por construcción, tenemos que

$$u \in \mathrm{Dom}(\psi) \subset B(u,\beta)$$

y ψ tiene diferencial de Fréchet en $\mathrm{Dom}(\psi)$. Usando el principio variacional de Ekeland 3.3.1 para $\varphi+\psi$, deducimos la existencia de $x_\varepsilon \in \mathrm{Dom}(\varphi+\psi) \subset B(u,\beta)$ tal que, para cada $x \in X$,

$$\varphi(x) + \psi(x) \geq \varphi(x_\varepsilon) + \psi(x_\varepsilon) - \varepsilon\|x - x_\varepsilon\|. \tag{4.10}$$

Como ψ es diferenciable Fréchet, existe $\eta > 0$ de forma que, para cada $x \in B(x_\varepsilon,\eta)$, se tiene que

$$\psi(x_\varepsilon) + x^*(x - x_\varepsilon) + \varepsilon\|x - x_\varepsilon\| \geq \psi(x), \tag{4.11}$$

donde $x^* = D\psi(x_\varepsilon)$. De las desigualdades anteriores (4.10) y (4.11) obtenemos finalmente que

$$\varphi(x) \geq \varphi(x_\varepsilon) - x^*(x - x_\varepsilon) - 2\varepsilon\|x - x_\varepsilon\|,$$

para cada $x \in B(x_\varepsilon,\eta)$, y la prueba termina. □

STOP El corolario 4.3.1 nos asegura que si X^* es separable, entonces existe una función θ que satisface las propiedades del teorema anterior.

Teorema 4.1.19 (Šmulian)**.** *Sea $\|\cdot\|$ una norma en un espacio de Banach X, con norma dual también denotada por $\|\cdot\|$ en X^*, y esferas respectivas S_X y S_{X^*}. Entonces:*

(i) La norma $\|\cdot\|$ es diferenciable Fréchet en $x \in S_X$ si, y sólo si, las condiciones $x_n^*, y_n^* \in S_{X^*}$, $n \in \mathbb{N}$, con $\lim_n x_n^*(x) = \lim_n y_n^*(x) = 1$, implican que

$$\lim_n \|x_n^* - y_n^*\| = 0.$$

(ii) La norma $\|\cdot\|$ es diferenciable Gâteaux en $x \in S_X$ si, y sólo si, las condiciones $x_n^*, y_n^* \in S_{X^*}$, $n \in \mathbb{N}$, con $\lim_n x_n^*(x) = \lim_n y_n^*(x) = 1$, implican que

$$\sigma(X^*,X) - \lim_n (x_n^* - y_n^*) = 0.$$

Demostración. Veamos primero la demostración de *(i)*. Supongamos que $\|\cdot\|$ tiene diferencial de Fréchet en $x \in S_X$. Sean $x_n^*, y_n^* \in S_{X^*}$, $n \in \mathbb{N}$, con $\lim_n x_n^*(x) = \lim_n y_n^*(x) = 1$, y tomemos $\varepsilon > 0$. Por el lema 4.1.10, existe $\delta > 0$ tal que, si $\|h\| \leq \delta$, entonces

$$\|x+h\| + \|x-h\| < 2 + \varepsilon\|h\|.$$

Por lo tanto, si $\|h\| \leq \delta$, se tiene que

$$x_n^*(x+h) + y_n^*(x-h) \leq \|x+h\| + \|x-h\| < 2 + \varepsilon\|h\|,$$

de donde concluimos la desigualdad

$$(x_n^* - y_n^*)\,(h) < 2 - x_n^*(x) - y_n^*(x) + \varepsilon\|h\|, \tag{4.12}$$

para cada $n \in \mathbb{N}$. Tomemos ahora $n_0 \in \mathbb{N}$ tal que $2 - x_n^*(x) - y_n^*(x) < \varepsilon\delta$ si $n \geq n_0$. Entonces, por (4.12) obtenemos que, para cada $n \geq n_0$,

$$(x_n^* - y_n^*)\,(h) < \varepsilon\big(\delta + \|h\|\big) \leq 2\varepsilon\delta,$$

para $\|h\| \leq \delta$. Sea $y \in S_X$ y pongamos $h = \pm\delta y$. Entonces, si $n \geq n_0$, $\pm\delta(x_n^* - y_n^*)(y) < 2\varepsilon\delta$, es decir, $\pm(x_n^* - y_n^*)(y) < 2\varepsilon$; en consecuencia,

$$\big|(x_n^* - y_n^*)(y)\big| < 2\varepsilon.$$

Así, $\|x_n^* - y_n^*\| < 2\varepsilon$ si $n \geq n_0$.

Recíprocamente, supongamos que $\|\cdot\|$ no tiene diferencial de Fréchet en x. Por el lema 4.1.10, existen $\varepsilon > 0$ y $h_n \in X$, $n \in \mathbb{N}$, con $h_n \neq 0$ y $\lim_n \|h_n\| = 0$, tales que

$$\|x + h_n\| + \|x - h_n\| \geq 2 + \varepsilon\|h_n\|. \tag{4.13}$$

Elegimos $x_n^*, y_n^* \in S_{X^*}$, $n \in \mathbb{N}$, de forma que

$$x_n^*(x+h_n) \geq \|x+h_n\| - \frac{\|h_n\|}{n} \quad \text{e} \quad y_n^*(x-h_n) \geq \|x-h_n\| - \frac{\|h_n\|}{n}. \tag{4.14}$$

Como

$$1 \geq x_n^*(x) = x_n^*(x+h_n) - x_n^*(h_n) \geq \|x+h_n\| - \frac{\|h_n\|}{n} - x_n^*(h_n)$$

y $\lim_n \big(\|x+h_n\| - \|h_n\|/n - x_n^*(h_n)\big) = 1$, obtenemos que $\lim_n x_n^*(x) = 1$. De forma análoga se tiene que $\lim_n y_n^*(x) = 1$. Ahora, de las desigualdades (4.13) y (4.14) se deduce que

$$x_n^*(x+h_n) + y_n^*(x-h_n) \geq 2 + \left(\varepsilon - \frac{2}{n}\right)\|h_n\|.$$

Entonces,

$$(x_n^* - y_n^*)(h_n) \geq 2 - x_n^*(x) - y_n^*(x) + \left(\varepsilon - \frac{2}{n}\right)\|h_n\| \geq \left(\varepsilon - \frac{2}{n}\right)\|h_n\|$$

y, en consecuencia, $\|x_n^* - y_n^*\| \geq \varepsilon - 2/n$.

Veamos ahora la prueba de *(ii)*. Supongamos que $\|\cdot\|$ tiene diferencial de Gâteaux en $x \in S_X$. Fijamos $y \in S_X$ y $\varepsilon > 0$. Por el lema 4.1.10, existe $\delta > 0$ tal que

$$\|x + ty\| + \|x - ty\| < 2 + \varepsilon\delta,$$

si $0 < t \leq \delta$. Sean $x_n^*, y_n^* \in S_{X^*}$, $n \in \mathbb{N}$, con $\text{lím}_n x_n^*(x) = \text{lím}_n y_n^*(x) = 1$. Entonces,

$$x_n^*(x+ty) + y_n^*(x-ty) \leq \|x+ty\| + \|x-ty\| < 2 + \varepsilon\delta,$$

de donde se sigue que

$$t(x_n^* - y_n^*)(y) < 2 - x_n^*(x) - y_n^*(x) + \varepsilon\delta.$$

Sea ahora $n_0 \in \mathbb{N}$ tal que, si $n \geq n_0$, $2 - x_n^*(x) - y_n^*(x) < \varepsilon\delta$. Entonces, $\delta(x_n^* - y_n^*)(y) < 2\varepsilon\delta$, es decir, $(x_n^* - y_n^*)(y) < 2\varepsilon$. Cambiando los papeles de x_n^* y de y_n^* obtenemos que $-(x_n^* - y_n^*)(y) < 2\varepsilon$, esto es, $\left|(x_n^* - y_n^*)(y)\right| < 2\varepsilon$, para $n \geq n_0$, lo que prueba que $\sigma(X^*, X) - \text{lím}_n(x_n^* - y_n^*) = 0$.

Supongamos ahora que $\|\cdot\|$ no es diferenciable Gâteaux en $x \in S_X$. Entonces, el lema 4.1.10 nos asegura la existencia de $\varepsilon > 0$, de $y \in S_X$ y de $(t_n)_n$ en $\mathbb{R}^+$, con $\text{lím}_n t_n = 0$, tales que

$$\|x + t_n y\| + \|x - t_n y\| \geq 2 + \varepsilon t_n. \tag{4.15}$$

Elegimos x_n^*, y_n^* en S_{X^*}, $n \in \mathbb{N}$, tales que

$$x_n^*(x+t_n y) \geq \|x+t_n y\| - \frac{t_n}{n} \quad \text{e} \quad y_n^*(x-t_n y) \geq \|x-t_n y\| - \frac{t_n}{n}. \tag{4.16}$$

Como

$$1 \geq x_n^*(x) = x_n^*(x+t_n y) - x_n^*(t_n y) \geq \|x+t_n y\| - \frac{t_n}{n} - x_n^*(t_n y) \geq 1 - t_n\|y\| - \frac{t_n}{n} - t_n\|y\| \to 1,$$

obtenemos que $\text{lím}_n x_n^*(x) = 1$. De manera análoga, $\text{lím}_n y_n^*(x) = 1$. Ahora, por las desigualdades (4.15) y (4.16), sabemos que

$$x_n^*(x+t_n y) + y_n^*(x-t_n y) \geq 2 + \left(\varepsilon - \frac{2}{n}\right) t_n.$$

Esto implica que

$$t_n(x_n^* - y_n^*)(y) \geq 2 - x_n^*(x) - y_n^*(x) + \left(\varepsilon - \frac{2}{n}\right) t_n \geq \left(\varepsilon - \frac{2}{n}\right) t_n,$$

de donde podemos concluir que

$$(x_n^* - y_n^*)(y) \geq \varepsilon - \frac{2}{n},$$

lo cual termina la demostración. □

Corolario 4.1.20. *En un espacio de Banach X, la norma $\|\cdot\|$ tiene diferencial de Gâteaux en $x \in X$, $x \neq 0$, si, y sólo si, existe un único $x^* \in S_{X^*}$ tal que $x^*(x) = \|x\|$. En este caso, la diferencial de Gâteaux en x es x^*.*

Demostración. Si $\|\cdot\|$ es diferenciable Gâteaux en $x \in X$, $x \neq 0$, en virtud del teorema de Šmulian 4.1.19 deducimos la existencia de un único $x^* \in X^*$, con $\|x^*\| = 1$, tal que $x^*(x) = \|x\|$.

Recíprocamente, razonando por reducción al absurdo, supongamos que la norma $\|\cdot\|$ no tiene diferencial de Gâteaux en en $x \in S_X$. Por el teorema de Šmulian 4.1.19, existen $y \in X$, $x_n^*, y_n^* \in S_{X^*}$ para $n \in \mathbb{N}$, y $\varepsilon > 0$ tales que

$$\lim_n x_n^*(x) = \lim_n y_n^*(x) = 1 \quad \text{ y } \quad (x_n^* - y_n^*)(y) \geq \varepsilon.$$

Sean $Y = \operatorname{span}\{x,y\}$, $z_n^* = x_n^*|_Y$ y $w_n^* = y_n^*|_Y$. Entonces, $z_n^*, w_n^* \in B_{Y^*}$, $\lim_n z_n^*(x) = \lim_n w_n^*(x) = 1$ y $(z_n^* - w_n^*)(y) \geq \varepsilon$. Como $\dim Y^* = \dim Y = 2 < \infty$, obtenemos que B_{Y^*} es compacta. Sin pérdida de generalidad, podemos suponer que $\lim_n \|z_n^* - z^*\| = \lim_n \|w_n^* - w^*\| = 0$, para ciertos $z^*, w^* \in B_{Y^*}$, en cuyo caso, $(z^* - w^*)(y) \geq \varepsilon$ y $z^*(x) = w^*(x) = 1$. Al ser $\|x\| = 1$, se tiene que $\|z^*\| = \|w^*\| = 1$, y por el teorema de Hahn-Banach, existen $x^*, y^* \in X^*$ tales que $\|x^*\| = \|y^*\| = 1$, $x^*|_Y = z^*$ e $y^*|_Y = w^*$. Entonces, $(x^* - y^*)(y) \geq \varepsilon$, luego $x^* \neq y^*$ y la prueba termina. $\square$

Si tomamos $x \in S_X$ y definimos $D_x = \{x^* \in S_{X^*} : x^*(x) = 1\}$, en virtud del teorema de Hahn-Banach se tiene que $D_x \neq \emptyset$ para todo $x \in S_X$. Así, por el corolario 4.1.20 anterior, la aplicación $D : S_X \longrightarrow S_{X^*}$ tal que, para cada $x \in S_X$, $D(x) = D_x$, es univaluada en el punto x si, y sólo si, la norma es diferenciable Gâteaux en x, en cuyo caso, $D(x)$ es la diferencial de Gâteaux de la norma.

Corolario 4.1.21. *Sea X un espacio de Banach y sea G (respectivamente, F) el conjunto de todos los puntos de S_X en los cuales la norma $\|\cdot\|$ es diferenciable Gâteaux (respectivamente, Fréchet). Entonces, cualquier selector $\partial : S_X \longrightarrow S_{X^*}$ de la aplicación $D : S_X \longrightarrow S_{X^*}$ es:*

(i) *$\|\cdot\|$-$\sigma(X^*,X)$-continuo en los puntos de G.*

(ii) *$\|\cdot\|$-$\|\cdot\|$-continuo en los puntos de F.*

Demostración. La demostración de *(i)* es como sigue. Sea $x \in G$. En virtud del teorema de Šmulian 4.1.19, para cada $u \in X$ y cada $\varepsilon > 0$, existe $\delta_u > 0$ tal que

$$\left|(\partial x - x^*)(u)\right| < \varepsilon, \quad \text{si } \|x^*\| = 1 \text{ y } \left|(\partial x - x^*)(x)\right| < \delta_u.$$

Tomamos un entorno U de ∂x en la topología $\sigma(X^*,X)$. Entonces, existen $\eta > 0$ y $u_1, \ldots, u_k \in X$ de forma que

$$\left\{x^* \in X^* : \left|(\partial x - x^*)(u_i)\right| < \eta,\ i = 1, \ldots, k\right\} \subset U.$$

Elegimos $\delta > 0$ tal que, si $\left|(\partial x - x^*)(x)\right| < \delta$, con $\|x^*\| = 1$, entonces

$$\left|(\partial x - x^*)(u_i)\right| < \eta \ \text{ para } i = 1, \ldots, k,$$

es decir, de modo que si $\big|(\partial x - x^*)(x)\big| < \delta$, con $\|x^*\| = 1$, entonces $x^* \in U$. Sea ahora $y \in S_X$ con $\|x-y\| < \delta$. Entonces,

$$\big|(\partial y - \partial x)(x)\big| = \big|\partial y(x) - \partial y(y)\big| \leq \|\partial y\|\|x-y\| < \delta,$$

y por tanto, $\partial y \in U$.

La demostración de *(ii)* es similar. Sea $(x_n)_n$ una sucesión en S_X tal que $\lim_n \|x_n - x\| = 0$, con $x \in F$. Se tiene que

$$\big|\partial x_n(x - x_n)\big| \leq \|\partial x_n\|\|x - x_n\| \to 0,$$

de donde podemos deducir que $\partial x_n(x) = \partial x_n(x_n) + \partial x_n(x - x_n) \to 1$. El teorema de Šmulian 4.1.19 para ∂x_n y ∂x nos permite concluir que $\lim_n \|\partial x_n - \partial x\| = 0$, y así queda terminada la prueba. $\square$

Corolario 4.1.22. *Si X es un espacio normado separable tal que la norma es diferenciable Fréchet para cada $x \neq 0$, entonces X^* también es separable.*

Demostración. Sea $\overline{\{x_n : n \in \mathbb{N}\}}^{\|\cdot\|} = S_X$. Probemos que $\overline{\{\partial x_n : n \in \mathbb{N}\}}^{\|\cdot\|} = S_{X^*}$. Para ello, tomemos $x^* \in S_{X^*}$ y $\varepsilon > 0$. En virtud del teorema de Bishop-Phelps 3.3.4, existen $y^* \in X^*$ y $x \in S_X$ tales que $y^*(x) = \|y^*\| \neq 0$ y $\|x^* - y^*\| < \varepsilon$. Consideremos $z^* = y^*/\|y^*\|$. Tenemos entonces que

$$\|y^* - z^*\| = \|y^*\|\left|1 - \frac{1}{\|y^*\|}\right| = \big|\|y^*\| - 1\big| = \big|\|y^*\| - \|x^*\|\big| \leq \|x^* - y^*\| < \varepsilon$$

y, en consecuencia, $\|x^* - z^*\| \leq \|x^* - y^*\| + \|y^* - z^*\| < 2\varepsilon$. Sea ahora $\lim_k \|x_{n_k} - x\| = 0$. Como $z^*(x) = \|z^*\| = 1$, el corolario 4.1.21 anterior nos asegura que $\lim_k \|\partial x_{n_k} - z^*\| = 0$. Así, podemos concluir que $\limsup_k \|\partial x_{n_k} - x^*\| \leq 2\varepsilon$. $\square$

Corolario 4.1.23. *Sea X un espacio de Banach. Si la norma de X^* es diferenciable Fréchet para cada $x^* \in X^*$, $x^* \neq 0$, entonces X es reflexivo.*

Demostración. Tomemos $x^{**} \in S_{X^{**}}$ y $\varepsilon > 0$. Procedemos como en el corolario anterior. En virtud del teorema de Bishop-Phelps, podemos encontrar $y^{**} \in S_{X^{**}}$ con $y^{**}(y^*) = 1$, para algún $y^* \in S_{X^*}$, y $\|x^{**} - y^{**}\| < 2\varepsilon$. Sean $(x_n)_n$ una sucesión en S_X tal que $\lim_n y^*(x_n) = 1$, y j la inclusión canónica de X en X^{**}, es decir, $jx(x^*) = x^*(x)$, para todo $x \in X$ y $x^* \in X^*$. Entonces, $\lim_n jx_n(y^*) = 1$. El teorema de Šmulian 4.1.19 nos asegura que $\lim_n \|jx_n - y^{**}\| = 0$, luego, $\limsup_n \|jx_n - x^{**}\| \leq 2\varepsilon$. Como X es completo, obtenemos que $x^{**} \in jX$. Es decir, X es reflexivo. $\square$

Teorema 4.1.24. *Sea X un espacio de Banach. Se tienen las siguientes propiedades:*

(i) *Si la norma en X^* es diferenciable Gâteaux en cada $x^* \in X^*$, $x^* \neq 0$, entonces la norma de X es estrictamente convexa.*

(ii) *Si la norma en X^* es estrictamente convexa, entonces la norma de X es diferenciable Gâteaux en cada $x \in X$, $x \neq 0$.*

Demostración. Para demostrar *(i)* tenemos que ver que, si $x, y, (x+y)/2 \in S_X$, entonces $x = y$. Elegimos $x^* \in S_{X^*}$ tal que $x^*\big((x+y)/2\big) = 1$. Al ser $x^*(x), x^*(y) \leq 1$, se tiene que $x^*(x) = x^*(y) = 1$. Como la norma en X^* es diferenciable Gâteaux en x^*, en virtud del corolario 4.1.20 obtenemos que $x = y$.

Demostremos ahora *(ii)*. Tomamos $x \in S_X$, y sean $x^*, y^* \in S_{X^*}$ con $x^*(x) = y^*(x) = 1$. Entonces, $2 \geq \|x^* + y^*\| \geq x^*(x) + y^*(x) = 2$, luego $\big\|(x^* + y^*)/2\big\| = 1$. Como la norma de X^* es estrictamente convexa, obtenemos que $x^* = y^*$. El corolario 4.1.20 nos asegura entonces que la norma de X es diferenciable Gâteaux en x. □

$\boxed{\textit{Norma localmente uniformemente convexa}}$ Una norma $\|\cdot\|$ en un espacio X se dice que es *localmente uniformemente convexa* (rotunda) si las condiciones

$$\|x_n\| = \|x\| = \lim_{n\to\infty}\left\|\frac{x_n + x}{2}\right\| = 1$$

implican que

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

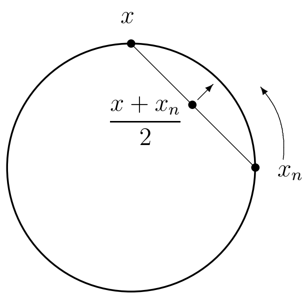


Figura 4.5: Norma localmente uniformemente convexa: en dimensión finita, los conceptos de norma uniformemente convexa, localmente uniformemente convexa y estrictamente convexa, coinciden

Teorema 4.1.25. *Si la norma de X^* es localmente uniformemente convexa, entonces la norma en X es diferenciable Fréchet para cada $x \neq 0$.*

Demostración. Para la demostración utilizamos de nuevo el teorema de Šmulian 4.1.19. Sean $x \in S_X$ y $x_n^*, y_n^* \in S_{X^*}$ con $\lim_n x_n^*(x) = \lim_n y_n^*(x) = 1$. Vamos a probar que $\lim_n \|x_n^* - y_n^*\| = 0$. Elegimos $x^* \in S_{X^*}$ tal que $x^*(x) = 1$. Entonces,

$$2 \geq \|x^* + x_n^*\| \geq (x^* + x_n^*)(x) \to 2.$$

Como la norma en X^* es localmente uniformemente convexa, obtenemos que $\lim_n \|x_n^* - x^*\| = 0$. De la misma manera, deducimos que $\lim_n \|y_n^* - x^*\| = 0$, y por tanto, $\lim_n \|x_n^* - y_n^*\| = 0$. □

El concepto de diferencial, 1877-1925. La derivada fue considerada inicialmente como una razón de variación, una velocidad o una pendiente. Este concepto ha ido evolucionando hasta convertirse en uno de los más importantes del Análisis. El uso de las derivadas parciales comenzó con L. Euler, A. Clairaut y J. d'Alembert, entre otros, a mediados del siglo XVIII, siendo Euler el que desarrolló el Cálculo en Derivadas Parciales en conexión con ciertos problemas de hidrodinámica. El cálculo diferencial en espacios de funciones lo inició V. Volterra en 1877 para estudiar problemas de cálculo de variaciones. En 1911, M. Fréchet enunció que una función real de dos variables era diferenciable en (a,b) si, y sólo si, su gráfica tenía plano tangente en ese punto. Más tarde, admitió que la expresión analítica adecuada era

$$f(a+h,b+k)-f(a,b)=D_1f(a,b)h+D_2f(a,b)k+\varepsilon(h,k)\big\|(h,k)\big\|,$$

donde $\varepsilon(h,k)$ tendía a cero cuando (h,k) se aproximaba a $(0,0)$. En un trabajo publicado en Abril de ese mismo año, Fréchet manifestó que, si bien al principio había creído que su definición de diferencial de una función ordinaria era nueva, con posterioridad había tenido noticias de una definición ligeramente diferente, pero equivalente, debida a W. H. Young (1909). A su vez, dicha definición había sido dada con anterioridad por O. Stolz, en 1893, y por J. Pierpont, en 1905. En 1913, R. Gâteaux introdujo la diferencial que lleva su nombre, que únicamente exige la existencia de todas las derivadas direccionales. El trabajo de Gâteaux no se dio a conocer hasta 1922, año en el que fue publicado por P. Levy. En 1925, Fréchet introdujo formalmente la noción de diferencial de funciones con valores en un espacio normado.

4.2 Renormamiento convexo

LAS proposición 4.1.17 y los teoremas 4.1.24 y 4.1.25 ponen de manifiesto la relación entre funciones mesetas diferenciables en un espacio de Banach y la existencia de renormamientos *buenos*. Nos ocupamos aquí de resultados clásicos de renormamiento que nos permiten obtener particiones de la unidad diferenciables en espacios de dimensión infinita.

Lema 4.2.1 (Argumentos de convexidad)**.** *Sea X un espacio vectorial y sea $\big(\|\cdot\|_n\big)_{n\in\mathbb{N}}$ una sucesión de seminormas en X tales que, para cada $x\in X$, la sucesión $\big(\|x\|_n\big)_n$ es acotada. Sea $(a_n)_n$ una sucesión de números no negativos tal que $\sum_{n=1}^{\infty}a_n<\infty$. Entonces $\|\cdot\|$, definida mediante*

$$\|x\|:=\left(\sum_{n=1}^{\infty}a_n\|x\|_n^2\right)^{1/2},$$

es una seminorma en X. Para cada $n\in\mathbb{N}$ y $x,y\in X$, se tiene que

$$\big(\|x\|_n-\|y\|_n\big)^2\le 2\big(\|x\|_n^2+\|y\|_n^2\big)-\|x+y\|_n^2\le\frac{1}{a_n}\Big(2\big(\|x\|^2+\|y\|^2\big)-\|x+y\|^2\Big).$$

Si $(x_k)_k$ es una sucesión en X e $y \in X$ son tales que

$$\lim_{k\to\infty} \|x_k + y\| = 2 \lim_{k\to\infty} \|x_k\| = 2\|y\|, \tag{4.17}$$

entonces

$$\lim_{k\to\infty} \|x_k + y\|_n = 2 \lim_{k\to\infty} \|x_k\|_n = 2\|y\|_n, \quad n = 1, 2, \ldots \tag{4.18}$$

En particular, si $\|x+y\| = 2\|x\| = 2\|y\|$, entonces

$$\|x+y\|_n = 2\|x\|_n = 2\|y\|_n, \quad n = 1, 2, \ldots \tag{4.19}$$

Demostración. Si $(\alpha_n)_{n=1}^{\infty}, (\beta_n)_{n=1}^{\infty} \in \ell^2$, tenemos que

$$\left(\sum_{n=1}^{\infty} (\alpha_n + \beta_n)^2\right)^{1/2} = \big\|(\alpha_n + \beta_n)_n\big\|_2 \leq \big\|(\alpha_n)_n\big\|_2 + \big\|(\beta_n)_n\big\|_2 = \left(\sum_{n=1}^{\infty} \alpha_n^2\right)^{1/2} + \left(\sum_{n=1}^{\infty} \beta_n^2\right)^{1/2}.$$

Entonces,

$$\begin{aligned}
\|x+y\| &= \left(\sum_{n=1}^{\infty} a_n \|x+y\|_n^2\right)^{1/2} \leq \left(\sum_{n=1}^{\infty} a_n \big(\|x\|_n + \|y\|_n\big)^2\right)^{1/2} \\
&= \left(\sum_{n=1}^{\infty} \big(\sqrt{a_n}\|x\|_n + \sqrt{a_n}\|y\|_n\big)^2\right)^{1/2} \leq \left(\sum_{n=1}^{\infty} a_n \|x\|_n^2\right)^{1/2} + \left(\sum_{n=1}^{\infty} a_n \|y\|_n^2\right)^{1/2} \\
&= \|x\| + \|y\|.
\end{aligned}$$

Por otro lado,

$$\begin{aligned}
2\big(\|x\|_n^2 + \|y\|_n^2\big) - \|x+y\|_n^2 &\geq 2\big(\|x\|_n^2 + \|y\|_n^2\big) - \big(\|x\|_n + \|y\|_n\big)^2 \\
&= \big(\|x\|_n - \|y\|_n\big)^2 \geq 0,
\end{aligned}$$

de donde deducimos que

$$\begin{aligned}
2\big(\|x\|^2 + \|y\|^2\big) - \|x+y\|^2 &= \sum_{n=1}^{\infty} a_n \Big(2\big(\|x\|_n^2 + \|y\|_n^2\big) - \|x+y\|_n^2\Big) \\
&\geq a_n \Big(2\big(\|x\|_n^2 + \|y\|_n^2\big) - \|x+y\|_n^2\Big) \\
&\geq a_n \big(\|x\|_n - \|y\|_n\big)^2 \geq 0, \quad n = 1, 2, \ldots
\end{aligned} \tag{4.20}$$

Sea ahora $(x_k)_k$ una sucesión satisfaciendo (4.17). Entonces, la desigualdad (4.20) nos asegura que

$$\begin{aligned}
0 = \lim_{k\to\infty} \Big(2\big(\|x_k\|^2 + \|y\|^2\big) - \|x_k + y\|^2\Big) &\geq a_n \limsup_{k\to\infty} \Big(2\big(\|x_k\|_n^2 + \|y\|_n^2\big) - \|x_k + y\|_n^2\Big) \\
\geq a_n \liminf_{k\to\infty} \Big(2\big(\|x_k\|_n^2 + \|y\|_n^2\big) - \|x_k + y\|_n^2\Big) &\geq 0, \quad n = 1, 2, \ldots
\end{aligned}$$

Utilizando de nuevo (4.20), se tiene que

$$0 = \lim_{k\to\infty}\Big(2\big(\|x_k\|_n^2 + \|y\|_n^2\big) - \|x_k + y\|_n^2\Big) \geq \limsup_{k\to\infty}\big(\|x_k\|_n - \|y\|_n\big)^2 \geq \liminf_{k\to\infty}\big(\|x_k\|_n - \|y\|_n\big)^2 \geq 0,$$

de donde se deduce (4.18). Tomando en (4.18) $x_k = x$, para todo $k \in \mathbb{N}$, obtenemos (4.19), y la prueba está completa. $\square$

Norma euclídea Una norma $\|\cdot\|$ en el espacio vectorial X se dice *euclídea* si, para cualesquiera $x, y \in X$, se satisface la Ley del Paralelogramo (véase la figura 1.5), es decir,

$$2\big(\|x\|^2 + \|y\|^2\big) = \|x+y\|^2 + \|x-y\|^2. \tag{4.21}$$

Las normas asociadas a productos escalares son euclídeas. En virtud del teorema de Jordan-von Neumann, [72, Theorem 21.4], el recíproco también es cierto: *si una norma es euclídea, entonces existe un producto escalar (necesariamente único) $\langle\cdot,\cdot\rangle$ tal que $\|x\| = \sqrt{\langle x, x\rangle}$, para todo $x \in X$.*

Corolario 4.2.2. *Sea X un espacio vectorial con dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$. Si $\|\cdot\|_2$ es una norma euclídea, entonces la norma dada por*

$$\|x\| = \big(\|x\|_1^2 + \|x\|_2^2\big)^{1/2}, \quad x \in X, \tag{4.22}$$

es estrictamente convexa.

Demostración. Sean $x, y \in X$ tales que $\|x\| = \|y\| = \big\|(x+y)/2\big\|$. Por el lema 4.2.1, se tiene que

$$\|x\|_2 = \|y\|_2 = \left\|\frac{x+y}{2}\right\|_2. \tag{4.23}$$

Dado que cada norma euclídea satisface la Ley del Paralelogramo,

$$2\|x\|_2^2 + 2\|y\|_2^2 = \|x+y\|_2^2 + \|x-y\|_2^2. \tag{4.24}$$

A partir de las igualdades (4.23) y (4.24) obtenemos que $\|x-y\|_2^2 = 0$, y como $\|\cdot\|$ es una norma, concluimos que $x = y$. $\square$

Obsérvese que, si en el corolario anterior la norma $\|\cdot\|_2$ se supone sólo estrictamente convexa, entonces la ecuación (4.22) todavía produce una norma estrictamente convexa (más en general, la suma de una función convexa con una función estrictamente convexa es estrictamente convexa). Para razonar esto, se sustituye la Identidad del Paralelogramo por la siguiente propiedad de las normas estrictamente convexas, [71, §26.1.(1)]:

Si $\|x+y\|_2 = \|x\|_2 + \|y\|_2$ e $y \neq 0$, entonces $x = \alpha y$, para algún α no negativo.

Esta propiedad es la clave para demostrar que un espacio X tiene un renormamiento estrictamente convexo si, y sólo si, existe un operador lineal, continuo e inyectivo de X sobre un espacio Y, que tiene norma estrictamente convexa, [71, §26.9.(3)].

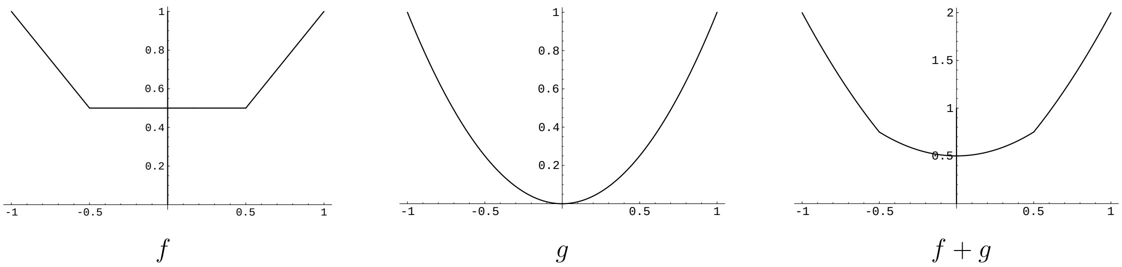


Figura 4.6: La suma de una función convexa f con una función estrictamente convexa g es estrictamente convexa

Corolario 4.2.3. *Si X es un espacio normado separable, entonces el dual X^* tiene una norma equivalente estrictamente convexa. Concretamente, si $D = \{x_n : n \in \mathbb{N}\}$ es un subconjunto denso para la norma en S_X, entonces*

$$|||x^*||| := \left(\|x^*\|^2 + \sum_{n=1}^{\infty} \frac{1}{2^n} x^*(x_n)^2 \right)^{1/2}, \quad x^* \in X^*,$$

es una norma equivalente, estrictamente convexa, en X^*.

Demostración. Si para cada $x^* \in X^*$ escribimos

$$\|x^*\|_2 := \left(\sum_{n=1}^{\infty} \frac{1}{2^n} x^*(x_n)^2 \right)^{1/2},$$

se prueba fácilmente que $\|\cdot\|_2$ es una seminorma en X^*; como D separa puntos de X^*, se concluye que, de hecho, $\|\cdot\|_2$ es una norma. Dado que la norma canónica de ℓ^2 satisface la identidad del paralelogramo, se deduce que $\|\cdot\|_2$ en X^* también es una norma euclídea en X^*. Por otro lado, para cada $x^* \in X^*$ se tiene la desigualdad $\|x^*\|_2 \leq \|x^*\|$. Esto nos dice que $|||\cdot|||$ es equivalente a la norma dual $\|\cdot\|$. El corolario 4.2.2 nos permite concluir que $|||\cdot|||$ es estrictamente convexa, y la demostración termina. $\square$

Obsérvese que un razonamiento similar al del corolario 4.2.3 permite demostrar el siguiente lema de renormamiento de Clarkson: *todo espacio normado separable $(X, \|\cdot\|)$ tiene una norma equivalente estrictamente convexa.* Para ello, basta fijar un conjunto numerable $D^* = \{x_n^*, n \in \mathbb{N}\} \subset X^*$ separando los puntos de X, véase el lema 2.4.8, y definir

$$|||x^*||| := \left(\|x\|^2 + \sum_{n=1}^{\infty} \frac{1}{2^n} x_n^*(x)^2 \right)^{1/2}, \quad x \in X.$$

La fórmula anterior nos da una norma estrictamente convexa en X equivalente a $\|\cdot\|$.

El lema de Clarkson comentado anteriormente puede mejorarse hasta el punto de obtener renormamientos localmente uniformemente convexos para espacios separables.

Teorema 4.2.4. *Si $(X, \|\cdot\|)$ es un espacio de Banach separable, entonces X admite una norma equivalente localmente uniformemente convexa. En concreto, si $D = \{e_n : n \in \mathbb{N}\}$ es denso para la norma en S_X, tomamos $D^* := \{x_n^* \in S_{X^*} : x_n^*(e_n) \geq 1/2,\ n \in \mathbb{N}\}$ y, para cada $n \in \mathbb{N}$, definimos $p_n(x) = \operatorname{ínf}_{\lambda \in \mathbb{R}} \|x - \lambda e_n\|$, $x \in X$, entonces la fórmula*

$$|||x||| := \left(\|x\|^2 + \sum_{n=1}^{\infty} \frac{1}{2^n} p_n(x)^2 + \sum_{n=1}^{\infty} \frac{1}{2^n} x_n^*(x)^2 \right)^{1/2},$$

para $x \in X$, es una norma localmente uniformemente convexa en X, equivalente a $\|\cdot\|$.

Demostración. Observemos en primer lugar que, por el lema 2.4.8, D^* separa los puntos de X. Consideremos ahora

$$\|x\|_2 = \left(\sum_{n=1}^{\infty} \frac{1}{2^n} x_n^*(x)^2 \right)^{1/2}.$$

Como D^* separa puntos de X, obtenemos que $\|\cdot\|_2$ es una norma en X, que es euclídea dado que la norma canónica de ℓ^2 lo es. Obsérvese por otro lado que, para cada $n \in \mathbb{N}$, p_n es una seminorma que satisface $p_n(x) \leq \|x\|$. Consecuentemente, $|||\cdot|||$ es una norma, equivalente a $\|\cdot\|$, que es estrictamente convexa gracias al corolario 4.2.2; probamos ahora que $|||\cdot|||$ es, de hecho, localmente uniformemente convexa. Sean $x_k, x \in X$, $k \in \mathbb{N}$, verificando

$$|||x_k||| = |||x||| = 1 \quad \text{y} \quad \lim_{k \to \infty} \left|\left|\left| \frac{x + x_k}{2} \right|\right|\right| = 1. \tag{4.25}$$

Por los argumentos de convexidad utilizados en el lema 4.2.1, tenemos que $\operatorname{lím}_{k \to \infty} p_n(x_k) = p_n(x)$, para cada $n = 1, 2, \ldots$ Demostremos ahora que $\{x_k : k \in \mathbb{N}\}$ es $\|\cdot\|$-totalmente acotado: veremos que, para cada $\varepsilon > 0$, existen $y_1, y_2, \ldots, y_m \in X$ tales que

$$\{x_k : k \in \mathbb{N}\} \subset \bigcup_{i=1}^{m} \{y \in X : \|y - y_i\| < 2\varepsilon\}. \tag{4.26}$$

Para empezar, obsérvese que la desigualdad $\|x\| \leq |||x||| = 1$ es evidente. Como $\overline{D}^{\|\cdot\|} = S_X$, existe $e_n \neq 0$ tal que $\|x - e_n\| < \varepsilon$, y de la desigualdad $p_n(x) \leq \|x - e_n\|$, concluimos que $p_n(x) < \varepsilon$. Ahora, la igualdad $\operatorname{lím}_k p_n(x_k) = p_n(x)$ implica la existencia de $k_\varepsilon \in \mathbb{N}$ tal que, para cada $k > k_\varepsilon$, se tiene que $p_n(x_k) < \varepsilon$. Entonces, si $k > k_\varepsilon$, existe $\lambda_k \in \mathbb{R}$ de forma que $\|x_k - \lambda_k e_n\| < \varepsilon$. Así,

$$|\lambda_k| = \frac{1}{\|e_n\|} \|\lambda_k e_n - x_k + x_k\| \leq \frac{1}{\|e_n\|} \left(\|\lambda_k e_n - x_k\| + \|x_k\| \right) \leq \frac{\varepsilon + 1}{\|e_n\|}.$$

Como el intervalo $[-(\varepsilon+1)/\|e_n\|, (\varepsilon+1)/\|e_n\|]$ es compacto, existen $\mu_1, \mu_2, \ldots, \mu_{m_\varepsilon}$ tales que, si $|\lambda| < (\varepsilon+1)/\|e_n\|$, entonces podemos encontrar $m \in \{1, \ldots, m_\varepsilon\}$ para el cual $|\lambda - \mu_m| < \varepsilon/\|e_n\|$. Es fácil ver ahora que $\{x_k : k = 1, \ldots, k_\varepsilon\} \cup \{\mu_m e_n : m = 1, \ldots, m_\varepsilon\}$ se puede tomar como el conjunto finito de los $y \in X$ que satisfacen la inclusión (4.26). Efectivamente, si tomamos x_k con

$k > k_\varepsilon$, entonces $\|x_k - \lambda_k e_n\| < \varepsilon$. Dado que $|\lambda_k| < (\varepsilon+1)/\|e_n\|$, encontramos un μ_m tal que $|\lambda_k - \mu_m| < \varepsilon/\|e_n\|$. En consecuencia,

$$\|x_k - \mu_m e_n\| \leq \|x_k - \lambda_k e_n\| + |\lambda_k - \mu_m|\,\|e_n\| < 2\varepsilon.$$

Al ser $\big(X, |||\cdot|||\big)$ completo (pues $\|\cdot\|$ y $|||\cdot|||$ son equivalentes), la sucesión $(x_k)_k$ está contenida en un subconjunto compacto de $\big(X, |||\cdot|||\big)$, y por lo tanto, para probar que $(x_k)_k$ converge hacia x, es suficiente demostrar que x es el único punto de aglomeración de $(x_k)_k$. Si y es un punto de aglomeración de $(x_k)_k$, existe una subsucesión $(x_{k_j})_j$ tal que $\lim_{j\to\infty} |||x_{k_j} - y||| = 0$. A partir de la igualdad (4.25) obtenemos que

$$|||x||| = |||y||| = \left|\left|\left|\frac{x+y}{2}\right|\right|\right| = 1,$$

y como $|||\cdot|||$ es estrictamente convexa, deducimos que $x = y$, y la prueba termina. $\square$

STOP Es conocido que cada espacio normado X se puede sumergir, de forma densa, en un espacio de Banach, *su complección*, véase la página 58. Cambiando un espacio normado por su complección, podemos convencernos de que el teorema 4.2.4 también es válido para espacios normados separables.

Lema 4.2.5. *Si $\big(X, \|\cdot\|\big)$ es un espacio normado, la norma dual en X^* es inferiormente semicontinua para la topología $\sigma(X^*, X)$.*

Demostración. Por definición, la norma dual se obtiene como

$$\|x^*\| := \sup\big\{|x^*(x)| : x \in B_X\big\}.$$

Para cada $x \in B_X$, la función $\hat{x} : X^* \longrightarrow \mathbb{R}$ dada por $\hat{x}(x^*) := x^*(x)$ es $\sigma(X^*, X)$-continua. Supremos de familias de funciones continuas son inferiormente semicontinuas, y así la prueba termina. $\square$

Lema 4.2.6. *Sea $\big(X, \|\cdot\|\big)$ un espacio de Banach y sea $|||\cdot|||$ una norma equivalente, en X^*, a la norma dual e inferiormente semicontinua en la topología $\sigma(X^*, X)$. Entonces, la fórmula*

$$|x| := \sup\big\{|x^*(x)| : x^* \in X^*,\ |||x^*||| \leq 1\big\}$$

es una norma equivalente en X, y $\big(X, |\cdot|\big)^* = \big(X^*, |||\cdot|||\big)$.

Demostración. Dado que $|||\cdot|||$ es $\sigma(X^*, X)$-inferiormente semicontinua, el conjunto

$$B = \big\{x^* \in X^* : |||x^*||| \leq 1\big\}$$

es $\sigma(X^*, X)$-cerrado. Consideremos ahora el par dual $\langle X, X^*\rangle$. Recordemos que el polar de B es

$$B^\circ = \Big\{x \in X : \sup_{x^* \in B} |x^*(x)| \leq 1\Big\}.$$

Como B es $\sigma(X^*,X)$-cerrado y absolutamente convexo, en virtud del teorema del bipolar 1.3.37 obtenemos que $(B^\circ)^\circ = B$. Definimos

$$|x^*|^* := \sup\big\{|x^*(x)| : x \in X,\, |x| \le 1\big\}, \quad x^* \in X^*.$$

Para concluir la prueba, demostraremos que $|x^*|^* = |||x^*|||$, para todo $x^* \in X^*$. Por la definición de $|\cdot|$, deducimos que $|x| = \sup\big\{|x^*(x)| : x^* \in B\big\}$. Entonces, $B^\circ = \big\{x \in X : |x| \le 1\big\}$, y por tanto,

$$\begin{aligned} B^{\circ\circ} &= \Big\{x^* \in X^* : \sup_{x \in B^\circ} |x^*(x)| \le 1\Big\} = \Big\{x^* \in X^* : \sup\big\{|x^*(x)| : x \in X,\, |x| \le 1\big\} \le 1\Big\} \\ &= \big\{x^* \in X^* : |x^*|^* \le 1\big\}. \end{aligned}$$

Como $B = B^{\circ\circ}$, obtenemos la igualdad

$$\big\{x^* \in X^* : |||x^*||| \le 1\big\} = \big\{x^* \in X^* : |x^*|^* \le 1\big\}.$$

Esto implica que, para cada $x^* \in X^*$, $x^* \neq 0$, se tiene que

$$\left|\left|\left| \frac{x^*}{|x^*|^*} \right|\right|\right| \le 1 \quad \text{y} \quad \left| \frac{x^*}{|||x^*|||} \right|^* \le 1,$$

y por tanto, $|x^*|^* = |||x^*|||$, para cada $x^* \in X^*$. Veamos por último que la norma $|\cdot|$ en X es equivalente a la norma original $\|\cdot\|$. Por ser $|||\cdot|||$ equivalente a la norma dual en X^*, existen constantes $\alpha, \beta > 0$ tales que

$$\alpha B_{X^*} \subset B \subset \beta B_{X^*}.$$

Tomando polares en $\langle X, X^*\rangle$, las inclusiones anteriores proporcionan, véase la proposición 1.3.36,

$$\frac{1}{\alpha} B_X \supset B^\circ \supset \frac{1}{\beta} B_X,$$

lo cual demuestra que $|\cdot|$ y $\|\cdot\|$ son equivalentes. □

Lema 4.2.7. *Sean X un espacio de Banach, $e^* \in S_{X^*}$ y $p(x^*) = \text{ínf}\big\{\|x^* - \lambda e^*\| : \lambda \in \mathbb{R}\big\}$. Entonces, p es $\sigma(X^*,X)$-inferiormente semicontinua en aB_{X^*}, para cada $a > 0$.*

Demostración. Fijamos $x^* \in B_{X^*}$ y $\varepsilon > 0$. Sea $\{\lambda_i : i = 1,\dots,n\}$ una colección finita de puntos en $\mathbb{R}$ tal que, para cada $\lambda \in [-2a, 2a]$, existe λ_i con $|\lambda - \lambda_i| < \varepsilon/3$. Para cada $i = 1,\dots,n$, podemos encontrar $z_i \in S_X$ de forma que

$$(x^* - \lambda_i e^*)(z_i) > \|x^* - \lambda_i e^*\| - \frac{\varepsilon}{3}.$$

Tomamos $V = \big\{y^* \in X^* : (x^* - y^*)(z_i) < \varepsilon/3,\, i = 1,\dots,n\big\}$. Fijamos $y^* \in V \cap aB_{X^*}$, y sea $\lambda \in \mathbb{R}$ tal que $p(y^*) = \|y^* - \lambda e^*\|$. Obsérvese además que $p(y^*) \le \|y^*\| \le a$. Tenemos entonces que

$$|\lambda| = \|\lambda e^*\| \le \|y^*\| + \|y^* - \lambda e^*\| = \|y^*\| + p(y^*) \le 2a.$$

Para el λ anterior, existe un λ_i de modo que $|\lambda-\lambda_i|<\varepsilon/3$. En consecuencia,

$$\begin{aligned} p(y^*) &= \|y^*-\lambda e^*\| \geq \|y^*-\lambda_i e^*\| - \|\lambda e^* - \lambda_i e^*\| \geq (y^*-\lambda_i e^*)(z_i) - |\lambda-\lambda_i| \\ &> (y^*-\lambda_i e^*)(z_i) - \frac{\varepsilon}{3} = (x^*-\lambda_i e^*)(z_i) + (y^*-x^*)(z_i) - \frac{\varepsilon}{3} \geq \|x^*-\lambda_i e^*\| - \frac{\varepsilon}{3} - \frac{\varepsilon}{3} - \frac{\varepsilon}{3} \\ &\geq p(x^*) - \varepsilon. \end{aligned}$$

Esto prueba que p es $\sigma(X^*,X)$-inferiormente semicontinua. □

Teorema 4.2.8. *Sea X un espacio de Banach.*

(i) Si X es separable, entonces X admite una norma equivalente diferenciable Gâteaux en cada $x \in X$ no nulo.

(ii) Si X^* es separable, entonces X admite una norma equivalente diferenciable Fréchet para cada $x \in X$ no nulo.

Demostración. Las demostraciones de los dos apartados son similares. Veamos primero *(i)*. Introducimos la norma equivalente en X^*, $|||\cdot|||$, que utilizamos en el corolario 4.2.3,

$$|||x^*||| = \left(\|x^*\|^2 + \sum_{n=1}^{\infty} \frac{1}{2^n} x^*(x_n)^2 \right)^{1/2}, \quad x^* \in X^*.$$

Es claro que cada $x^* \in X^*$ es $\sigma(X^*,X)$-continua, y por lo tanto, también lo es $(x^*)^2$. Además, como la serie $\sum_{n=1}^{\infty} 2^{-n} x^*(x_n)^2$ converge uniformemente en cada bola aB_{X^*}, para $a>0$, obtenemos que $\|x^*\|_2 = \left(\sum_{n=1}^{\infty} 2^{-n} x^*(x_n)^2\right)^{1/2}$ es $\sigma(X^*,X)$-continua en aB_{X^*}, $a>0$. Ahora bien, por el lema 4.2.5, $\|\cdot\|$ es $\sigma(X^*,X)$-inferiormente semicontinua. Utilizando las propiedades sobre funciones inferiormente semicontinuas recogidas en la página 112, $|||\cdot|||$ es $\sigma(X^*,X)$-inferiormente semicontinua en cada bola aB_{X^*}, $a>0$. Probemos ahora que $|||\cdot|||$ es, de hecho, $\sigma(X^*,X)$-inferiormente semicontinua en X^*. Fijamos $y^* \in X^*$ y $\varepsilon>0$. Al ser $|||\cdot|||$ una norma equivalente en X^*, existe un $a>0$ tal que $B = \{x^* \in X^* : |||x^*||| \leq |||y^*|||\} \subset aB_{X^*}$. Como $|||\cdot|||$ es $\sigma(X^*,X)$-inferiormente semicontinua en aB_{X^*}, podemos deducir la existencia de un entorno V de 0 en X^*, para la topología $\sigma(X^*,X)$, tal que $|||x^*||| > |||y^*||| - \varepsilon$, si $x^* \in V \cap aB_{X^*}$. Si $x^* \notin aB_{X^*}$, obtenemos que $x^* \notin B$, luego $|||x^*||| > |||y^*|||$. Entonces, $|||x^*||| > |||y^*||| - \varepsilon$ para todo $x^* \in V$, y por tanto, $|||\cdot|||$ es, de hecho, $\sigma(X^*,X)$-inferiormente semicontinua en X^*. Tomemos ahora

$$|x| := \sup\{|x^*(x)| : x^* \in X^*, |||x^*||| \leq 1\}, \quad x \in X.$$

En virtud del lema 4.2.6 $(X,|\cdot|)^* = (X^*, |||\cdot|||)$. Como $|||\cdot|||$ es estrictamente convexa, por el teorema 4.1.24 $|\cdot|$ es diferenciable Gâteaux, para cada $x \in X$ no nulo.

Veamos la demostración de *(ii)*. Introducimos la norma equivalente $|||\cdot|||$ dada en el teorema 4.2.4, ahora para X^* en lugar de para X, es decir,

$$|||x^*||| = \left(\|x^*\|^2 + \sum_{n=1}^{\infty} \frac{1}{2^n} p_n(x^*)^2 + \sum_{n=1}^{\infty} \frac{1}{2^n} x^*(x_n)^2 \right)^{1/2}, \quad x^* \in X^*,$$

donde

$$x_n \in S_X \quad \text{y} \quad p_n(x^*) = \inf_{\lambda\in\mathbb{R}} \|x^* - \lambda e_n^*\|, \quad n = 1,2,\ldots,$$

siendo $e_n^* \in S_{X^*}$ con $e_n^*(x_n) \geq 1/2$, $n = 1,2,\ldots$, y $\{e_n^* : n \in \mathbb{N}\}$ $\|\cdot\|$-denso en S_{X^*}. Por el citado teorema 4.2.4 $|||\cdot|||$ es una norma equivalente localmente uniformemente convexa. Con un razonamiento análogo al expuesto en el apartado *(i)*, obtenemos que $\|\cdot\|$ es $\sigma(X^*,X)$-inferiormente semicontinua, y que $\sum_{n=1}^{\infty} 2^{-n} x^*(x_n)^2$ es $\sigma(X^*,X)$-continua. Además, por las propiedades de estabilidad para funciones semicontinuas y el lema 4.2.7, $\sum_{n=1}^{\infty} 2^{-n} p_n(x^*)^2$ es $\sigma(X^*,X)$-inferiormente semicontinua en cada bola aB_{X^*}, $a > 0$. Así, la norma $|||\cdot|||$ es $\sigma(X^*,X)$-inferiormente semicontinua en cada bola aB_{X^*}, con $a > 0$. Razonando tal y como hicimos en *(i)*, obtenemos que $|||\cdot|||$ es $\sigma(X^*,X)$-inferiormente semicontinua en X. Tomamos de nuevo

$$|x| = \sup\big\{|x^*(x)| : x^* \in X^*,\ |||x^*||| \leq 1\big\}.$$

En virtud del lema 4.2.6, $\big(X,|\cdot|\big)^* = \big(X^*,|||\cdot|||\big)$. Como $|||\cdot|||$ es localmente uniformemente convexa, por el teorema 4.1.25 $|\cdot|$ es diferenciable Fréchet, para cada $x \in X$ no nulo. $\square$

Renormamiento en espacios de Banach. El corolario 4.2.3 fue establecido, en 1936, por J. A. Clarkson, el teorema 4.2.4 por M. Kadec en 1959, el teorema 4.2.8, apartado *(i)*, por M. Day en 1975, y el apartado *(ii)*, independientemente por Kadec y V. Klee en 1965. Utilizando que todos los espacios de Banach separables admiten una norma localmente uniformemente convexa, teorema 4.2.4, Kadec demostró, en 1967, que todos los espacios de Banach separables de dimensión infinita son homeomorfos, resolviendo así uno de los problemas planteados en el *Scottish Book*.

4.3 Particiones de la unidad

DEDICAMOS esta sección al estudio de la existencia de particiones de la unidad formadas por funciones diferenciables. Empezamos por demostrar el siguiente resultado que es consecuencia de los resultados establecidos en la sección anterior.

Corolario 4.3.1. *Sea X un espacio de Banach.*

(i) *Si X^* es separable, entonces, para cada $x \in X$ y cada entorno U_x de x, existe una función meseta $\theta_x : X \longrightarrow [0,1]$ tal que $\theta_x(x) > 0$, $\operatorname{sop}\theta_x \subset U_x$ y θ_x es diferenciable Fréchet, con diferencial continua.*

(ii) *Si X es separable, entonces, para cada $x \in X$ y cada entorno U_x de x, existe una función meseta continua $\theta_x : X \longrightarrow [0,1]$ tal que $\theta_x(x) > 0$, $\operatorname{sop}\theta_x \subset U_x$, θ_x es diferenciable Gâteaux y $D\theta_x : \big(X,\|\cdot\|\big) \longrightarrow \big(X^*,\sigma(X^*,X)\big)$ es continua.*

Demostración. Veamos *(i)*. En virtud del teorema 4.2.8 anterior, existe una norma equivalente $|||\cdot|||$ en X, que es diferenciable Fréchet en cada punto no nulo. Vamos a representar por ∂ la diferencial de Fréchet de la norma $|||\cdot|||$. Fijamos $\varepsilon > 0$, y definimos

$$\varphi(t) := \begin{cases} (\varepsilon^2 - t^2)^2 & \text{si } |t| \le \varepsilon, \\ 0 & \text{si } |t| > \varepsilon. \end{cases}$$

Claramente, φ' es continua en $\mathbb{R}$. Entonces, la función $\theta(y) := \varphi\big(|||y|||\big)$ es, por la regla de la cadena 4.1.8, diferenciable Fréchet, siendo

$$D\theta(y) = \begin{cases} -4\big(\varepsilon^2 - |||y|||^2\big)\,|||y|||\,\partial y & \text{si } |||y||| \le \varepsilon, \\ 0 & \text{si } |||y||| > \varepsilon. \end{cases}$$

El corolario 4.1.21 nos asegura ahora que $D\theta : X \longrightarrow X^*$ es $\|\cdot\|$-$\|\cdot\|$-continua. Además, $\operatorname{sop}\theta$ está contenido en la bola de centro cero y radio ε. Tomamos un $x \in X$ arbitrario. Entonces, la función $\theta_x(y) = \theta(x+y)$ satisface las condiciones buscadas, para $U_x = \big\{y \in X : |||x-y||| \le \varepsilon\big\}$.

La demostración de *(ii)* se deduce de forma análoga, de nuevo, a partir del teorema 4.2.8 anterior, de la proposición 4.1.8 y del corolario 4.1.21. □

Teorema 4.3.2 (Asplund)**.** *Sea X un espacio de Banach tal que X^* es separable. Entonces, para cada cubrimiento abierto $\mathscr{U}$ de X, existe una partición de la unidad de clase C^1 que está subordinada a $\mathscr{U}$.*

Antes de ver la demostración de este teorema, vamos a probar un lema previo.

Lema 4.3.3 (Lindelöf)**.** *Sea X un espacio métrico separable, y sea $\mathscr{U}$ una familia de conjuntos abiertos de X que recubre X. Entonces, existe un subrecubrimiento numerable $\{U_n \in \mathscr{U} : n \in \mathbb{N}\}$ del espacio X.*

Demostración. Sea $\{x_n : n \in \mathbb{N}\}$ un conjunto numerable denso en X, y sea $\{B_m : m \in \mathbb{N}\}$ una reordenación de $B(x_n,q) = \big\{y \in X : d(x_n,y) < q\big\}$, para $n \in \mathbb{N}$ y $q \in \mathbb{Q}^+$. Definimos una aplicación $k : X \longrightarrow \mathbb{N}$ de forma que $x \in B_{k(x)}$ y existe $U \in \mathscr{U}$ con $B_{k(x)} \subset U$. Representamos por $\mathbb{M} := k(X)$. Claramente,

$$\bigcup_{m \in \mathbb{M}} B_m = X, \tag{4.27}$$

ya que, si $x \in X$, entonces $x \in B_m$, para $m = k(x)$. Veamos ahora que, para cada $m \in \mathbb{M}$, existe un $U_m \in \mathscr{U}$ con $B_m \subset U_m$. En efecto, como $m \in \mathbb{M}$, existe $z \in X$ de forma que $m = k(z)$; luego podemos encontrar $U_m \in \mathscr{U}$ tal que $B_m \subset U_m$. Por (4.27), obtenemos finalmente que $\bigcup_{m\in\mathbb{M}} U_m = X$. □

Demostración del teorema 4.3.2. Sea $\mathscr{U}$ un recubrimiento por abiertos de X. En virtud del corolario 4.3.1, para cada $x \in X$, existen $U_x \in \mathscr{U}$ y una función meseta θ_x de clase C^1, tales que

$\theta_x(y) \geq 0$ para todo $y \in X$, $\theta_x(x) > 0$ y sop $\theta_x \subset U_x$. Pongamos $W_x := \{y \in X : \theta_x(y) > 0\}$. Entonces, $W_x \subset U_x$. Como θ_x es continua, se tiene que W_x es abierto. Además, dado que X^* es separable, por el lema 1.3.48 sabemos que X también es separable. Utilizamos ahora el lema 4.3.3, el cual nos asegura la existencia de $x_1, x_2, \ldots$ de forma que $X = \bigcup_{n=1}^{\infty} W_{x_n}$. Sea

$$V_n = \Big\{y \in X : \theta_{x_k}(y) < \frac{1}{n},\ k = 1,2,\ldots,n-1\Big\} \cap W_{x_n}, \quad \text{con } V_1 = W_{x_1}.$$

Fijado $x \in X$, podemos encontrar $N = N(x) \in \mathbb{N}$ tal que $\theta_{x_1}(x) = \ldots = \theta_{x_{N-1}}(x) = 0$ y $\theta_{x_N}(x) > 0$. Claramente, $x \in V_N$. Tomamos $a(x) \in \big(0, \theta_{x_N}(x)\big)$, y sea

$$G_x := \{y \in X : \theta_{x_N}(y) > a(x)\}.$$

Entonces, se tiene que

$$V_n \cap G_x = \emptyset, \tag{4.28}$$

para $n > \max\{a^{-1}(x), N\}$; en efecto, si $y \in V_n$, como $n > N$, obtenemos que $\theta_{x_N}(y) < n^{-1} < a(x)$. Luego $y \notin G_x$. Sean ahora $\varphi_n \in C^1(\mathbb{R})$, $n = 0,1,2,\ldots$, tales que

- $\varphi_0(t) > 0$ si $t > 0$, y $\varphi_0(t) = 0$ si $t \leq 0$,
- $\varphi_n(t) > 0$ si $t \in (0,1/n)$, y $\varphi_n(t) = 0$ si $t \notin (0,1/n)$, $n = 1,2,\ldots$

Por ejemplo, podemos considerar las funciones

$$\varphi_0(t) = \begin{cases} t^2 & \text{si } t > 0 \\ 0 & \text{si } t \leq 0, \end{cases} \qquad \varphi_n(t) = \begin{cases} t^2(n^{-2} - t)^2 & \text{si } t \in (0,1/n) \\ 0 & \text{si } t \notin (0,1/n), \end{cases}$$

para $n = 1,2,\ldots$ Tomamos entonces

$$f_1(t_1) := \varphi_0(t_1), \qquad f_n(t_1,t_2,\ldots,t_n) := \varphi_n(t_1)\ldots\varphi_n(t_{n-1})\varphi_0(t_n),\ n = 2,3,\ldots,$$

y $g_n(y) := f_n\big(\theta_{x_1}(y), \theta_{x_2}(y), \ldots, \theta_{x_n}(y)\big)$. Es claro que

$$V_n \supset \{y \in X : g_n(y) > 0\}. \tag{4.29}$$

Como $x \in G_x$, las relaciones (4.28) y (4.29) anteriores nos aseguran que $\{n \in \mathbb{N} : g_n(x) > 0\}$ es un conjunto finito para cada $x \in X$. Definimos

$$g(x) := \sum_{n=1}^{\infty} g_n(x).$$

Además, al ser $g(x) \geq g_N(x) > 0$, obtenemos que $g(x) > 0$, y esto, para cada $x \in X$. Sea, finalmente, $h_n(x) = g_n(x)/g(x)$. Entonces, para cada $x \in X$, tendremos que

$$\sum_{n=1}^{\infty} h_n(x) = 1$$

y, para cada $y \in G_x$, $h_n(y) = 0$ si $n > \max\{a^{-1}(x), N\}$, lo que concluye la demostración. □

4.4 Conexión con el proyecto de investigación: espacios de Asplund

Teorema 4.4.1. *Sean Y un espacio normado con norma estrictamente convexa, X un espacio normado y $T : X \longrightarrow Y$ una aplicación lineal, continua e inyectiva. Entonces, X admite una norma equivalente estrictamente convexa.*

De hecho, basta tomar $|||x||| := \|x\|_X + \|Tx\|_Y$, para todo $x \in X$, pues se puede comprobar que $|||\cdot|||$ es una norma equivalente y estrictamente convexa sobre X. Este resultado puede demostrarse fácilmente utilizando las técnicas de la sección 4.2, como ha sido comentado en la observación de la página 161.

Este hecho fue utilizado por Lindenstrauss, [73], por ejemplo, para probar que en cualquier espacio de Banach reflexivo X existe un renormamiento estrictamente convexo, al ser capaz de construir un operador lineal continuo e inyectivo $T : X \longrightarrow c_0(I)$, para algún conjunto de índices I. Sin embargo, Dashiell y Lindenstrauss, [37], dieron ejemplos de espacios de Banach X con norma estrictamente convexa y sin ningún operador $T : X \longrightarrow c_0(I)$ lineal, continuo e inyectivo. Nuestro grupo de investigación ha trabajado, y trabaja, en problemas de transferencia *no lineal* de buenas normas, como las estrictamente convexas; es decir, en el estudio de las aplicaciones, no lineales en general, $\phi : X \longrightarrow Y$, que nos permitan transferir una norma estrictamente convexa en Y a una norma estrictamente convexa en X. Esta investigación se ha centrado, principalmente, en la transferencia de normas localmente uniformemente convexas (brevemente, LUR), propiedad que añade, a la convexidad estricta, la igualdad de las topologías débil y de la norma en la esfera unidad del espacio considerado (lo que se conoce como propiedad de Kadec). Los resultados para estas normas son muy satisfactorios, véase [79], y resulta ahora un problema de enorme interés analizar en profundidad este estudio para comprender la transferencia de normas Kadec y de normas estrictamente convexas por separado.

El éxito conseguido para normas LUR se fundamenta en la siguiente caracterización, véanse [78], [89]:

Teorema 4.4.2. *Sean X un espacio normado y F un subespacio normante de su dual, X^*. Entonces X admite una norma equivalente $\sigma(X,F)$-inferiormente semicontinua y LUR si, y sólo si, para cada $\varepsilon > 0$, podemos escribir*

$$X = \bigcup_{n=1}^{\infty} X_{n,\varepsilon}$$

de forma que, para todo $x \in X_{n,\varepsilon}$, exista un semiespacio H $\sigma(X,F)$-abierto conteniendo a X, con

$$\operatorname{diam}(H \cap X_{n,\varepsilon}) < \varepsilon.$$

Si se cambia renormamiento estrictamente convexo por renormamiento LUR, el teorema 4.4.1 no es cierto: basta observar que el operador $T : \ell^\infty \longrightarrow \ell^2$ dado por $T\big((x_n)_n\big) := (2^{-n}x_n)_n$ es continuo e inyectivo y, sin embargo, ℓ^∞ no admite una norma equivalente LUR. La propiedad que debe tener T para que se pueda transferir la norma LUR en Y a una norma LUR en X fue analizada en [78], y se corresponde a la siguiente definición:

Definición 4.4.3. *Una aplicación ϕ definida de un espacio métrico (X,d) en otro espacio métrico (Y,ρ) se dice que es co-σ-continua si, para todo $\varepsilon > 0$, podemos escribir*

$$X = \bigcup_{n=1}^{\infty} X_{n,\varepsilon},$$

y encontrar un $\delta_n(x) > 0$, para cada $x \in X_{n,\varepsilon}$, de forma que $d(x,y) < \varepsilon$, siempre que $y \in X_{n,\varepsilon}$ y $\rho\big(\phi(x),\phi(y)\big) < \delta_n(x)$.

Así, si tenemos un operador $T : X \longrightarrow Y$ lineal, continuo y co-σ-continuo entre espacios normados, si Y admite una norma equivalente LUR, entonces X también la tiene. El operador $T : X \longrightarrow c_0(I)$ construido por Lindenstrauss para X reflexivo, e incluso cualquier operador lineal, continuo e inyectivo $T : X \longrightarrow c_0(I)$ para X débilmente compactamente generado (Amir-Lindenstrauss, [1]) o X débilmente numerablemente K-determinado (Vašák, [105]), es, automáticamente, co-σ-continuo. Para el caso no lineal precisamos del siguiente tipo de aplicaciones:

Definición 4.4.4. *Una aplicación $\phi : X \longrightarrow (Y,\rho)$, donde X es un espacio localmente convexo e (Y,ρ) es un espacio métrico, es σ-*slicely*-continua si, para todo $\varepsilon > 0$, podemos escribir*

$$X = \bigcup_{n=1}^{\infty} X_{n,\varepsilon}$$

de forma que, para cada $x \in X_{n,\varepsilon}$, existe un semiespacio abierto H en X, con $x \in H$ y tal que

$$\operatorname{diam}\big(\phi(H \cap X_{n,\varepsilon})\big) < \varepsilon.$$

Esto nos permite llegar al siguiente resultado:

Teorema 4.4.5. *Sean X un espacio normado y F un subespacio normante en el dual X^*. Entonces, X admite una norma equivalente $\sigma(X,F)$-inferiormente semicontinua y LUR si, y sólo si, existe un espacio métrico (Y,ρ) y una aplicación $\phi : X \longrightarrow Y$ que sea σ-*slicely*-continua para $\sigma(X,F)$ y co-σ-continua para la norma.*

Este teorema fundamenta un nuevo marco para el estudio de renormamientos del tipo LUR. Tal y como hemos manifestado anteriormente, resulta ahora de enorme interés analizar por separado las normas estrictamente convexas y las normas Kadec. En nuestro proyecto BFM2002-01719 se pormenorizan los aspectos relevantes para estas investigaciones.

Para enfatizar la no linealidad de nuestra noción de aplicación σ-*slicely*-continua, digamos, por ejemplo, que la aplicación de dualidad que asigna, a cada $x \in X$ (X espacio normado), un

$x^* \in S_{X^*}$ tal que $x^*(x) = \|x\|$, es, cuando la norma $\|\cdot\|$ es diferenciable Fréchet, una aplicación σ-*slicely*-continua, gracias al criterio de Šmulian, véase [79]. Esto nos sitúa en predisposición de analizar la conexión, si la hay, entre renormamiento diferenciable Fréchet y renormamiento LUR, cuestión ésta del máximo interés entre los especialistas, [38, 61, 111], y que se ha desvelado cierta cuando existe una norma dual LUR: si un espacio de Banach X admite una norma dual LUR en X^*, entonces X tiene una norma LUR equivalente, (Haydon, comunicación privada), admitiendo estos espacios particiones de la unidad de clase C^1, como lo hacían los considerados por Asplund en el teorema 4.3.2, véanse [4] o [38, Chapter VIII, theorems 3.2, 3.12].

PARA SABER MÁS

- Los libros [38] y [50] son dos buenas referencias para complementar las cuestiones sobre renormamiento y diferenciabilidad estudiadas en este capítulo. El segundo libro es un texto de propósito general, que proporciona una introducción al Análisis Funcional lineal y no lineal en espacios de Banach. El primer libro es una monografía autocontenida orientada, tanto a estudiantes, como a especialistas en espacios de Banach, que contiene todos los resultados expuestos en este capítulo y otros muchos que conectan renormamientos y diferenciabilidad en espacios de Banach.
- Dos conceptos importantes en espacios de Banach, cuyas definiciones están asociadas a la noción de diferenciabilidad, son los de *espacio de Asplund* y *espacio débil de Asplund*. Un espacio de Banach X se dice que es un espacio de Asplund (respectivamente, débil de Asplund) si cada función continua convexa definida en un subconjunto abierto Ω de X es diferenciable Fréchet (respectivamente, Gâteaux) en algún subconjunto G_δ denso de Ω. X es un espacio de Asplund si, y sólo si, cada subespacio $Y \subset X$ separable tiene dual separable, véase [15, Theorem 4.2.13]. Esta caracterización conecta los espacios de Asplund con la propiedad de Radon-Nikodým, véase la sección 5.4. Los libros [49] y [85] son buenas referencias para lecturas avanzadas sobre los espacios de Asplund y débil de Asplund.

Capítulo 5

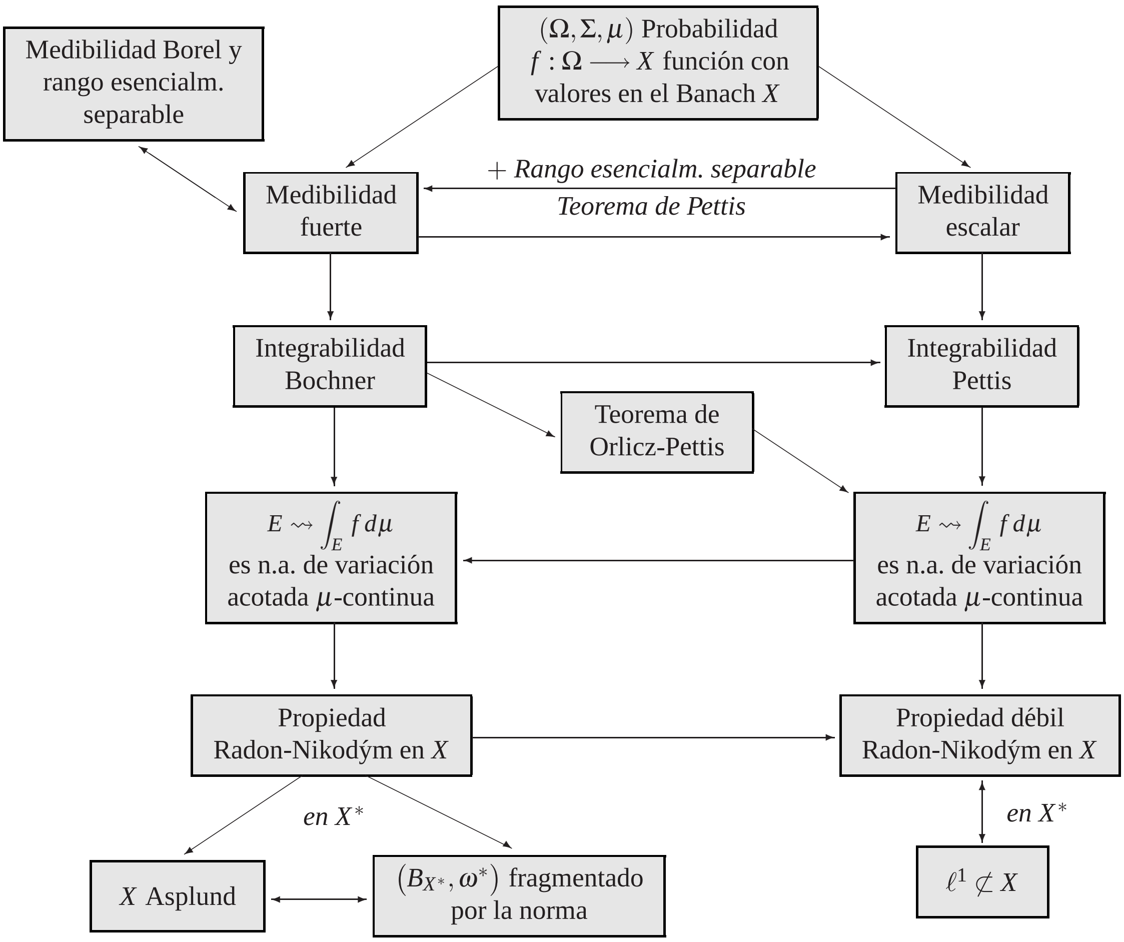
Integración en espacios de Banach

«OBJETIVOS»

- Estudiar, para funciones con valores en espacios de Banach, las nociones de medibilidad fuerte y medibilidad escalar, así como la relación existente entre ellas. Distinguir entre ambos conceptos, y estudiar y entender su relación con la noción clásica de medibilidad Borel.
- Introducir y estudiar propiedades básicas de la integral de Bochner. Aprender, mediante ejemplos notables, algunas técnicas asociadas al concepto de integral de Bochner, que le confieren carácter de herramienta valiosa en el contexto de los espacios de Banach: demostración del teorema de Krein-Šmulian y del teorema de Orlicz-Pettis vía la integral de Bochner.
- Conocer las propiedades básicas de la integral de Pettis como ampliación y en contraposición a la integral de Bochner.
- Introducir al alumno en las técnicas de integración vectorial como preámbulo a posibles líneas de investigación clásicas y actuales.

DESDE los orígenes de la teoría de los espacios de Banach, se planteó el problema de determinar en qué medida se podían extender conceptos y resultados clásicos del Análisis, de una o varias variables reales, al nuevo ámbito infinito-dimensional. Naturalmente, en este sentido, la Teoría de la Integración no quedó al margen, y ya en 1927, Graves, [57], puso de manifiesto que el caso de funciones con valores en espacios de Banach presenta diferencias significativas con el caso de funciones escalares. En concreto, Graves dio un ejemplo de una función $f : [0,1] \longrightarrow \ell^\infty\big([0,1]\big)$ que es integrable Riemann pero que no es fuertemente medible, es decir, límite en casi todo punto de una sucesión de funciones simples.

Durante la década de los 30, surgieron diferentes extensiones de la teoría de integración de Lebesgue al caso de funciones vectoriales, destacando las debidas a Bochner, [9], y Pettis, [84]. Desde el principio quedó claro que, aunque dichas nociones de integral coincidían con la integral



Cuadro 5.1: Esquema del capítulo *Integración en espacios de Banach*

de Lebesgue para funciones escalares, diferían entre sí en el caso general. Para más información sobre otras teorías estudiadas en la misma época nos remitimos a [63].

Generalmente, la integral de funciones con valores en espacios de Banach más utilizada ha sido la de Bochner, véase la monografía de Diestel y Uhl [40], que es restrictiva en el sentido de que cualquier función integrable Bochner debe ser fuertemente medible y, en consecuencia, existen funciones vectoriales integrables Riemann que no son integrables Bochner. La integral de Pettis, que no precisa medibilidad fuerte y es más general que la de Bochner, no fue estudiada en profundidad hasta finales de los 70. La memoria de Talagrand [103] y los artículos expositorios [80, 81] son buenas referencias para la integral de Pettis.

En este capítulo estudiamos las nociones de medibilidad fuerte y medibilidad escalar de funciones con valores en espacios de Banach, mostrando que son distintas en general, y ligándolas entre sí a través del clásico teorema de medibilidad de Pettis 5.1.4. Introducimos la noción de integral de Bochner, para la que ponemos de manifiesto que el Teorema de Convergencia Dominada 5.2.3 y algunos otros resultados conocidos sobre funciones escalares integrables se extienden sin dificultad. Estudiamos la integral indefinida asociada a una función integrable Bochner, demostrando que es una medida numerablemente aditiva de variación acotada y absolutamente continua respecto de la medida escalar utilizada para integrar, teorema 5.2.4. Mostramos, mediante un ejemplo, que el teorema de Radon-Nikodým no se puede extender a los espacios de Banach generales, ejemplo 5.2.10. Utilizamos la integral de Bochner para probar el teorema de Krein-Šmulian 5.2.8 (compárese esta demostración con la ofrecida en el teorema 3.2.1), y para demostrar el teorema de Orlicz-Pettis 5.2.12. El capítulo se completa con una rápida introducción a las integrales de Dunford y de Pettis, donde nos preocupamos de establecer, con ayuda del teorema de Orlicz-Pettis, que una función integrable Dunford es integrable Pettis si, y sólo si, su integral indefinida asociada es numerablemente aditiva, proposición 5.3.4 y teorema 5.3.6. Concluimos el capítulo con una breve exposición sobre la Propiedad de Radon-Nikodým y la Propiedad Débil de Radon-Nikodým, que conectan con líneas actuales de investigación en topología, renormamiento, diferenciación e integración vectorial. El cuadro 5.1 resume en un esquema la interconexión existente entre las distintas partes de este capítulo.

Primeras definiciones. En todo este capítulo, (Ω,Σ,μ) será un espacio completo de medida finita, y $\big(X,\|\cdot\|\big)$ denotará un espacio de Banach real.

$\boxed{\textit{Espacio de medida completo}}$ Sean Ω un conjunto no vacío, Σ una σ-álgebra de conjuntos y μ una medida finita definida en Σ. Decimos que el espacio de medida (Ω,Σ,μ) es un *espacio completo* si cada vez que $E \in \Sigma$, con $\mu(E) = 0$ y $A \subset E$, se tiene que $A \in \Sigma$. Todo espacio de medida (Ω,Σ,μ) puede completarse, [34, p. 36].

$\boxed{\textit{Funciones medibles}}$ Si (Ω,Σ) es un espacio medible, una función $f : \Omega \longrightarrow \mathbb{R}$ se dice *medible* si, para cada conjunto de Borel B en $\mathbb{R}$, se tiene que $f^{-1}(B) \in \Sigma$. La familia de las funciones reales medibles es estable por las siguientes operaciones: suma y producto de funciones, supremos de familias numerables y límites puntuales de sucesiones. Si (Ω,Σ,μ) es un espacio de medida completo, entonces el límite en casi todo punto respecto a μ de una sucesión de funciones medibles es de nuevo medible, [34, Capítulo 1].

$\boxed{\textit{Para casi todo punto}}$ Si (Ω,Σ,μ) es un espacio de medida y $\big[P(w)\big]$ es una cierta propiedad que se puede definir en términos de $w \in \Omega$, diremos que $[P]$ es cierta *μ-para casi todo punto en* Ω (brevemente, μ-p.c.t. $w \in \Omega$ o μ-p.c.t.p.) si $\big[P(w)\big]$ se satisface para todos los puntos de Ω, salvo quizás, en un conjunto de μ medida cero. Cuando μ se dé por supuesta, escribiremos simplemente $[P]$ p.c.t. $w \in \Omega$, o $[P]$ p.c.t.p.

$\boxed{\textit{Funciones integrables}}$ Si (Ω,Σ,μ) es un espacio de medida, representaremos por $\mathscr{L}^1(\mu)$ el espacio de las funciones reales μ-integrables Lebesgue en Ω. Para $f \in \mathscr{L}^1(\mu)$, escribiremos $\|f\|_1 := \int_\Omega |f|\,d\mu$. Los detalles sobre la integral y los espacios de funciones integrables asociados pueden consultarse en [96] y [34].

5.1 Medibilidad en espacios de Banach

Los conceptos que introducimos en las definiciones que siguen serán utilizados con profusión, tanto al estudiar la integrabilidad Bochner como la integrabilidad Pettis, en las secciones 5.2 y 5.3. El paralelismo de las nociones vectoriales con sus contrapartidas escalares es total.

Función simple Una función $s : \Omega \longrightarrow X$ se dice *simple* si es de la forma

$$s = \sum_{i=1}^{n} \alpha_i \chi_{A_i},$$

donde $\alpha_i \in X$ y $A_i \in \Sigma$.

Función μ-medible Se dice que una función $f : \Omega \longrightarrow X$ es *μ-medible* si existe una sucesión $(s_n)_n$ de funciones simples de Ω en X tal que

$$\lim_n \|s_n - f\| = 0, \quad \mu \text{ p.c.t. } w \in \Omega.$$

Función débilmente medible Se dice que $f : \Omega \longrightarrow X$ es *débilmente medible* si, para cada $x^* \in X^*$, la función $x^* f$ es medible.

Tenemos las siguientes propiedades inmediatas de las funciones medibles:

- El conjunto de las funciones μ-medibles es un espacio vectorial.
- Para cada función simple $s = \sum_{i=1}^{n} \beta_i \chi_{B_i}$, se puede encontrar una partición $\{A_i\}_{i=1}^{m}$ de Ω en Σ de forma que $s = \sum_{i=1}^{m} \alpha_i \chi_{A_i}$. De este modo se tiene que $s(w) = \alpha_i$ si, y sólo si, $w \in A_i$. En consecuencia, tomando normas, tenemos que

$$\|s\| = \sum_{i=1}^{m} \|\alpha_i\| \chi_{A_i},$$

  y por tanto, esta función es una función escalar simple.
- Si g es una función vectorial μ-medible, entonces $\|g\|$ es medible. Efectivamente, si $(s_n)_n$ es una sucesión de funciones simples tales que

$$\lim_n \|s_n - g\| = 0, \quad \text{p.c.t. } w \in \Omega,$$

  se tiene que

$$0 \leq \limsup_n \big| \|s_n\| - \|g\| \big| \leq \lim \|s_n - g\| = 0, \quad \text{p.c.t. } w \in \Omega,$$

  y por lo tanto, la función $\|g\|$ es una función escalar medible.

Cambiando el valor absoluto $|\cdot|$ por la norma $\|\cdot\|$ en la prueba clásica del teorema de Egoroff, se puede obtener la versión que sigue para funciones vectoriales.

Teorema 5.1.1 (Egoroff)**.** *Sea $f_n : \Omega \longrightarrow X$, $n \in \mathbb{N}$, una sucesión de funciones μ-medibles, y sea $f : \Omega \longrightarrow X$ una función tal que $(f_n)_n$ converge hacia f μ-p.c.t.p. Entonces, para cada $\varepsilon > 0$, existe $E \subset \Omega$ con $\mu(E) < \varepsilon$ y tal que $(f_n)_n$ converge hacia f uniformemente en $\Omega \setminus E$.*

Demostración. Para todo $n \in \mathbb{N}$ se tiene que

$$\lim_m \|f_n - f_m\| = \|f_n - f\|, \quad \text{p.c.t. } w \in \Omega.$$

Como $\|f_n - f_m\|$ es medible, la función $\|f_n - f\|$ es medible para todo $n \in \mathbb{N}$. Sea

$$\begin{aligned} D_{mn} &= \Big\{ w \in \Omega : \big\|f_r(w) - f(w)\big\| \geq \frac{1}{m}, \ \text{ para algún } r \geq n \Big\} \\ &= \bigcup_{r=n}^{\infty} \Big\{ w \in \Omega : \big\|f_r(w) - f(w)\big\| \geq \frac{1}{m} \Big\} \in \Sigma. \end{aligned}$$

Para m fijo, $\bigcap_{n=1}^{\infty} D_{mn}$ es un conjunto nulo, puesto que está contenido en el conjunto de puntos donde $(f_n)_n$ no converge a f. Para cada $m \in \mathbb{N}$,

$$D_{mn} \searrow \bigcap_{n=1}^{\infty} D_{mn},$$

y así, como el espacio de medida es finito, la σ-aditividad de μ implica que

$$\lim_n \mu(D_{mn}) = 0, \quad \text{para todo } m \in \mathbb{N}.$$

Ahora, dado $\varepsilon > 0$, para cada $m \in \mathbb{N}$, existe $n(m)$ de forma que $\mu\big(D_{m,n(m)}\big) < \varepsilon/2^m$. Definimos $E := \bigcup_{m=1}^{\infty} D_{m,n(m)}$. Claramente, $E \in \Sigma$ satisface que $\mu(E) \leq \varepsilon$, y $(f_n)_n$ converge uniformemente hacia f en $\Omega \setminus E$. □

Como consecuencia del teorema de Egoroff se tiene el siguiente resultado:

Proposición 5.1.2. *El conjunto de las funciones μ-medibles con valores en un espacio de Banach es cerrado por límites de sucesiones convergentes en casi todo punto.*

Demostración. Sean $f : \Omega \longrightarrow X$ una función y $f_n : \Omega \longrightarrow X$, $n \in \mathbb{N}$, una sucesión de funciones μ-medibles tales que $(f_n)_n$ converge μ-p.c.t.p. en Ω hacia f. Para cada $n \in \mathbb{N}$, existe una sucesión de funciones simples $(f_{n,m})_m$ tal que $(f_{n,m})_m$ converge hacia f_n μ-p.c.t.p. Fijado $n \in \mathbb{N}$, podemos aplicar el teorema de Egoroff y obtener $m(n) \in \mathbb{N}$ y $E_n \in \Sigma$, con $\mu(E_n) \leq 1/2^n$, tales que

$$\big\|f_{n,m(n)}(w) - f_n(w)\big\| \leq \frac{1}{n}, \quad \text{para todo } \ w \in \Omega \setminus E_n.$$

Definamos

$$E := \bigcap_{k=1}^{\infty} \Big(\bigcup_{n \geq k} E_n \Big) \in \Sigma.$$

Claramente, $\mu(E) = 0$. Sea ahora $w \in \Omega \setminus E$. Entonces, existe $k \in \mathbb{N}$ tal que $w \in \Omega \setminus E_n$, para cada $n \geq k$. Así, si $n \geq k$, se tiene que

$$\left\| f_{n,m(n)}(w) - f_n(w) \right\| \leq \frac{1}{n}.$$

Como $(f_n)_n$ converge μ-p.c.t.p. hacia f, resulta también que $\left(f_{n,m(n)}\right)_n$ converge μ-p.c.t.p. hacia f y, en consecuencia, f es μ-medible, lo que concluye la prueba. $\square$

El siguiente teorema es crucial en el estudio que se hará de la integrabilidad Bochner. Establecemos primero un pequeño resultado que supone una ligera mejora del lema 2.4.8.

Lema 5.1.3. *Si $\left(X, \|\cdot\|\right)$ es un espacio normado, sea $D = \{x_n : n \in \mathbb{N}\}$ un subconjunto denso de una parte $A \subset X$. Si, para cada $n \in \mathbb{N}$, elegimos $x_n^* \in B_{X^*}$ de forma que $x_n^*(x_n) = \|x_n\|$, entonces $D^* := \{x_n^* : n \in \mathbb{N}\}$ es normante para A,* i.e.*, para cada $x \in A$, se tiene que*

$$\|x\| = \sup\left\{ \left|x_n^*(x)\right| : n \in \mathbb{N} \right\}. \tag{5.1}$$

Demostración. Dado $\varepsilon > 0$, tomamos x_n tal que $\|x - x_n\| < \varepsilon$. Entonces,

$$x_n^*(x) = x_n^*(x_n) + x_n^*(x - x_n) \geq \|x_n\| - \|x - x_n\| \geq \|x\| - 2\|x - x_n\| \geq \|x\| - 2\varepsilon,$$

lo que prueba (5.1). $\square$

Teorema 5.1.4 (Teorema de medibilidad de Pettis)**.** *Para una función $f : \Omega \longrightarrow X$, las siguientes afirmaciones son equivalentes:*

(i) f es μ-medible.

(ii) (a) Existe $E \in \Sigma$ con $\mu(E) = 0$ y tal que $f(\Omega \setminus E)$ es separable.

(b) Para cada $x^* \in X^*$, la función $x^* f$ es medible.

Demostración. Probemos *(i)* $\Rightarrow$ *(ii)* Supongamos que f es μ-medible. Por la continuidad y linealidad de cada $x^* \in X^*$, resulta evidente que la función $x^* f$ es límite p.c.t.p. de una sucesión de funciones simples y medibles, y que por lo tanto, $x^* f$ es medible. Para ver la parte (a), supongamos que $(s_n)_n$ es una sucesión de funciones simples, y sea $E \in \Sigma$, con $\mu(E) = 0$, y tal que $(s_n)_n$ converge hacia f en $\Omega \setminus E$. Entonces, $f(\Omega \setminus E) \subset \overline{\bigcup_n s_n(\Omega \setminus E)}$, que es separable gracias a que las funciones s_n son simples.

Veamos *(ii)* $\Rightarrow$ *(i)* Supongamos ahora que $f : \Omega \longrightarrow X$ cumple las condiciones de *(ii)*. Sea $E \in \Sigma$ tal que $\mu(E) = 0$ y $f(\Omega \setminus E)$ es separable. Sea $\{x_n : n \in \mathbb{N}\} \subset f(\Omega \setminus E)$ un conjunto denso numerable. Por el lema 5.1.3, si elegimos $x_n^* \in X^*$ con

$$x_n^*(x_n) = \|x_n\| \quad \text{y} \quad \|x_n^*\| = 1,$$

se tiene que, para $w \in \Omega \setminus E$,

$$\big\| f(w) \big\| = \sup \Big\{ \big| x_n^* f(w) \big| : n \in \mathbb{N} \Big\},$$

y por tanto, $\|f\|$ es una función medible. Análogamente se obtiene que, para cada $n \in \mathbb{N}$, la función $g_n(w) = \big\| f(w) - x_n \big\|$ es una función escalar medible. Dado $\varepsilon > 0$, consideramos el conjunto

$$B_n = \Big\{ w \in \Omega : \big\| f(w) - x_n \big\| < \varepsilon \Big\} \in \Sigma.$$

Así, si definimos

$$g(w) = \begin{cases} x_n & \text{si } w \in B_n \setminus \bigcup_{m<n} B_m, \\ 0 & \text{en el resto,} \end{cases}$$

se tiene que

$$\big\| f(w) - g(w) \big\| < \varepsilon, \qquad \text{para cada } \ w \in \Omega \setminus E.$$

En consecuencia, tomando $A_n := B_n \setminus \bigcup_{m<n} B_m$, queda claro que

$$g = \sum_{n=1}^{\infty} x_n \chi_{A_n}$$

es una función medible (por ser límite de simples) que aproxima uniformemente, en casi todo punto, a f. Se tiene así que, para cada $m \in \mathbb{N}$, existe g_m de la forma anterior tal que

$$\big\| f(w) - g_m(w) \big\| < \frac{1}{m}, \qquad \text{p.c.t.p. en } \Omega.$$

En consecuencia, f es límite en casi todo punto de una sucesión de funciones medibles y, dado que el conjunto de las funciones μ-medibles es cerrado para este tipo de límites después de la proposición 5.1.2, resulta que f es medible. □

El resto de la sección está dedicado a obtener algunas consecuencias del teorema de medibilidad de Pettis.

Corolario 5.1.5. *Una función $f : \Omega \longrightarrow X$ es medible si, y sólo si, es límite uniforme en casi todo punto de una sucesión de funciones medibles que toman una cantidad numerable de valores.*

Demostración. La condición suficiente es consecuencia inmediata de la proposición 5.1.2. La condición necesaria es parte de la demostración del teorema anterior, donde se ha obtenido que, para cada $n \in \mathbb{N}$, existe g_n medible, tomando una cantidad numerable de valores, tal que

$$\big\| g_n(w) - f(w) \big\| < \frac{1}{n}, \qquad \text{p.c.t.p. en } \Omega. \qquad \square$$

Modificando adecuadamente los valores de las funciones $(g_n)_n$, puede encontrarse una sucesión $(f_n)_n$ de funciones medibles numerablemente valuadas, con $(f_n)_n$ convergiendo hacia f casi uniformemente y con $\|f_n\| \leq \|f\|$, véase [42, Prop. 14, p. 99].

Repasando la demostración del teorema anterior, se observa que es posible obtener una condición suficiente de medibilidad más débil que la que hemos dado en el teorema de Pettis.

Corolario 5.1.6. *Sean $f : \Omega \longrightarrow X$ una función μ-esencialmente separable valuada y $B \subset B_{X^*}$ un conjunto normante para X, tales que $x^* f$ es medible para cada $x^* \in B$. Entonces, f es μ-medible.*

Demostración. Basta repetir la demostración que se hace del teorema de medibilidad de Pettis, teniendo en cuenta que el lema 5.1.3 puede mejorarse de la siguiente forma: si para cada $n \in \mathbb{N}$ elegimos $(x^*_{n,m})_m$ en B tal que $\|x_n\| = \sup\big\{|x^*_{n,m}(x_n)| : m \in \mathbb{N}\big\}$, entonces $D^* := \{x^*_{n,m} : n, m \in \mathbb{N}\}$ es normante para el conjunto $A \subset X$ allí considerado. □

Los conceptos de medibilidad y medibilidad débil no coinciden en general:

Ejemplo 5.1.7 (Una función débilmente medible que no es medible). Consideremos el espacio de Hilbert $\ell^2\big([0,1]\big)$, y sea

$$\begin{aligned} f:\ [0,1] &\to \ell^2\big([0,1]\big) \\ t &\to e_t \end{aligned}$$

donde $\big\{e_t : t \in [0,1]\big\}$ es la base ortonormal canónica del espacio de Hilbert $\ell^2\big([0,1]\big)$. Entonces, para cada $x^* \in \big(\ell^2([0,1])\big)^* = \ell^2\big([0,1]\big)$, se tiene que

$$\begin{aligned} x^*:\ [0,1] &\to \mathbb{R} \\ t &\to x^*(t) \end{aligned}$$

es tal que

$$\sum_{t \in [0,1]} \big|x^*(t)\big|^2 < \infty,$$

y por tanto, sólo hay una cantidad numerable de $x^*(t) \neq 0$. Así, como $x^* f(t) = x^*(t)$, resulta que $x^* f$ es nula salvo, a lo más, en un conjunto numerable, que es de medida de Lebesgue nula y, en consecuencia, $x^* f$ es medible. Por otra parte, $f\big([0,1] \setminus E\big)$ es separable si, y sólo si, $[0,1] \setminus E$ es numerable, en cuyo caso, E tiene medida de Lebesgue 1, y así, por el teorema de medibilidad de Pettis, f no puede ser medible. □

La construcción anterior se puede hacer más en general en espacios con un *carácter de densidad* grande y en los que existen *bases débiles*. Por ejemplo, si X es un espacio de Banach reflexivo con carácter de densidad, al menos, el cardinal de $[0,1]$ (*i.e.*, cualquier conjunto denso de X tiene cardinal, al menos, $\mathfrak{c}$), entonces existe una función $f : [0,1] \longrightarrow X$ escalarmente medible pero no μ-medible, véase [92].

Proposición 5.1.8. *Sean f y g funciones μ-medibles. Si para cada $x^* \in X^*$ se tiene que*

$$x^* f = x^* g = 0, \quad \textit{p.c.t.p. en } \Omega \text{ (el conjunto de medida nula depende de } x^*\text{)},$$

entonces $f = g$ p.c.t.p. en Ω.

Demostración. Por el teorema de medibilidad de Pettis 5.1.4 podemos suponer, sin pérdida de generalidad, que X es separable. Gracias al lema 5.1.3, podemos tomar un conjunto numerable $D^* := \{x_n^* : n \in \mathbb{N}\} \subset B_{X^*}$ normante para X. Para cada $n \in \mathbb{N}$, sea $E_n \in \Sigma$, con $\mu(E_n) = 0$, tal que $x_n^* f(w) = x_n^* g(w)$ para cada $w \in \Omega \setminus E_n$. En particular, si $w \in \Omega \setminus \left(\bigcup_{n=1}^{\infty} E_n\right)$ se deduce que

$$x_n^*\big(f(w) - g(w)\big) = 0, \quad \text{para cada } n \in \mathbb{N}.$$

Al ser D^* normante, se tiene que $\big\|f(w) - g(w)\big\| = 0$ si $w \in \Omega \setminus \left(\bigcup_{n=1}^{\infty} E_n\right)$, y la prueba termina. □

La noción de μ-medibilidad para funciones $f : \Omega \longrightarrow X$ puede interpretarse en términos de medibilidad Borel.

Medibilidad Borel para funciones vectoriales La función $f : \Omega \longrightarrow X$ es *medible Borel* cuando $f^{-1}(B) \in \Sigma$ para cada subconjunto de Borel B en X. Recordemos que límites en todo punto de sucesiones de funciones medibles Borel (en general, con valores en un espacio métrico) son medibles Borel, [34, Proposition 8.1.8].

Corolario 5.1.9. *Para una función $f : \Omega \longrightarrow X$ son equivalentes:*

(i) f es μ-medible.

(ii) f es medible Borel y existe $E \in \Sigma$, con $\mu(E) = 0$, tal que $f(\Omega \setminus E)$ es separable.

Demostración. La implicación *(ii) ⇒ (i)* se sigue del teorema de medibilidad de Pettis 5.1.4. Veamos *(i) ⇒ (ii)*. Ya sabemos que f es esencialmente separable, y para demostrar que f es medible Borel, procedemos como en la prueba del teorema de medibilidad de Pettis: supongamos que $(s_n)_n$ es una sucesión de funciones simples, y sea $E \in \Sigma$ tal que $\mu(E) = 0$, con $(s_n)_n$ convergente hacia f en $\Omega \setminus E$. La sucesión de restricciones $(s_n|_{\Omega \setminus E})_n$ converge puntualmente a $f|_{\Omega \setminus E}$. Cada $s_n|_{\Omega \setminus E} : \Omega \setminus E \longrightarrow X$ es medible Borel cuando en $\Omega \setminus E$ se considera la σ-álgebra traza,

$$\Sigma_{\Omega \setminus E} = \big\{(\Omega \setminus E) \cap A : A \in \Sigma\big\}.$$

Consecuentemente, $f|_{\Omega \setminus E}$ también es medible Borel, lo cual significa que $f^{-1}(B) \cap (\Omega \setminus E) \in \Sigma$ para cada boreliano $B \subset X$. Así, se tiene que

$$f^{-1}(B) = \big[f^{-1}(B) \cap (\Omega \setminus E)\big] \cup \big[f^{-1}(B) \cap E\big] \in \Sigma,$$

dado que $\mu(E) = 0$ y (Ω, Σ, μ) es completo. □

5.2 La integral de Bochner

En esta sección nos vamos a ocupar de estudiar la extensión de la bien conocida integral de Lebesgue, al caso de funciones que toman valores en un espacio de Banach, lo que conduce a la definición de la llamada integral de Bochner.

Integración de funciones simples Si $s = \sum_{i=1}^{n} \alpha_i \chi_{A_i}$, con $A_i \in \Sigma$ y $\alpha_i \in X$, es una función simple, definimos la integral

$$\int_E s\, d\mu = \sum_{i=1}^{n} \alpha_i \mu(E \cap A_i),$$

para cada $E \in \Sigma$. Se comprueba que la definición anterior es independiente de la representación elegida para la función simple s. La integral es lineal, definida en el espacio de las funciones simples.

Integral de Bochner Una función μ-medible $f : \Omega \longrightarrow X$ se dice que es *integrable Bochner*, si existe una sucesión de funciones simples $(s_n)_n$ tal que

$$\lim_n \int_\Omega \|s_n - f\|\, d\mu = 0.$$

En este caso, para cada $E \in \Sigma$, la sucesión $\left(\int_E s_n\, d\mu\right)_n$ es de Cauchy en X, y por tanto, podemos definir

$$\int_E f\, d\mu := \lim_n \int_E s_n\, d\mu.$$

Al vector $\int_E f\, d\mu$ se le llama *integral de Bochner de f sobre E*.

- El valor $\int_E f\, d\mu$ es independiente de la sucesión de simples tomada con tal que,

$$\lim_n \int_\Omega \|s_n - f\|\, d\mu = 0.$$

- El conjunto de las funciones $f : \Omega \longrightarrow X$ integrables Bochner es un espacio vectorial.

A continuación, caracterizamos la integrabilidad de Bochner vía la integrabilidad de Lebesgue.

Teorema 5.2.1. *Sea $f : \Omega \longrightarrow X$ una función μ-medible. Son equivalentes:*

(i) f es integrable Bochner.

(ii) $\|f\|$ es integrable Lebesgue.

Demostración. Es claro que, al ser f medible, también lo es $\|f\|$, véase la página 176.

Veamos primero la implicación *(i)* $\Rightarrow$ *(ii)*. Sea $s_n : \Omega \longrightarrow X$, $n \in \mathbb{N}$, una sucesión de funciones simples con

$$\lim \int_\Omega \|s_n - f\|\, d\mu = 0.$$

Entonces, existe $n_0 \in \mathbb{N}$ tal que

$$\int_\Omega \|s_{n_0} - f\|\, d\mu < 1.$$

Por lo tanto,

$$\int_\Omega \|f\|\, d\mu = \int_\Omega \|s_{n_0} - f - s_{n_0}\|\, d\mu \leq 1 + \int_\Omega \|s_{n_0}\|\, d\mu < \infty$$

y, en consecuencia, $\|f\|$ es integrable Lebesgue.

La implicación *(ii)* $\Rightarrow$ *(i)* es como sigue. Supongamos ahora que f es μ-medible y que $\|f\|$ es integrable Lebesgue. Con ayuda del corolario 5.1.5 podemos encontrar, para cada $n \in \mathbb{N}$, una función $f_n : \Omega \longrightarrow X$ medible, que toma una cantidad numerable de valores, tal que

$$\|f_n - f\| < \frac{1}{n}, \quad \text{p.c.t.p. en } \Omega.$$

Así, se tiene que $\|f_n\| < 1/n + \|f\|$ y, como μ es finita, $\|f_n\|$ es integrable. Las funciones f_n serán de la forma

$$f_n = \sum_{m=1}^{\infty} x_{nm} \chi_{E_{nm}}, \quad \text{con } E_{nm} \in \Sigma,\ x_{nm} \in E_{nm},\ \text{y } E_{nm} \cap E_{nm'} = \emptyset \text{ si } m \neq m'.$$

Para cada n fijo, existe $p_n \in \mathbb{N}$ tal que

$$\int_{\bigcup_{m=p_n+1}^{\infty} E_{nm}} \|f_n\|\, d\mu < \frac{\mu(\Omega)}{n}.$$

Pongamos ahora $g_n = \sum_{m=1}^{p_n} x_{nm} \chi_{E_{nm}}$. Cada g_n es una función simple, y se tiene que

$$\int_\Omega \|f - g_n\|\, d\mu \leq \int_\Omega \|f - f_n\|\, d\mu + \int_\Omega \|f_n - g_n\|\, d\mu \leq 2\frac{\mu(\Omega)}{n}.$$

Esto prueba que $\lim_n \int_\Omega \|f - g_n\|\, d\mu = 0$, y así concluye la demostración. □

Proposición 5.2.2. *Si $f : \Omega \longrightarrow X$ es integrable Bochner, entonces, para todo $E \in \Sigma$, se tiene que*

$$\left\| \int_E f\, d\mu \right\| \leq \int_E \|f\|\, d\mu.$$

Demostración. Si $(s_n)_n$ es una sucesión de funciones simples con $\lim_n \int_\Omega \|s_n - f\|\, d\mu = 0$,

$$\lim_n \int_E s_n\, d\mu = \int_E f\, d\mu,$$

para cada $E \in \Sigma$. Ahora bien, $s_n = \sum_{i=1}^{m} x_{ni} \chi_{E_{ni}}$, con $E_{ni} \cap E_{nj} = \emptyset$ para $i \neq j$. Entonces,

$$\|s_n\| = \sum_{i=1}^{m} \|x_{ni}\| \chi_{E_{ni}},$$

y por tanto,

$$\left\|\int_E s_n\,d\mu\right\| = \left\|\sum_{i=1}^{m} x_{ni}\mu(E\cap E_{ni})\right\| \le \sum_{i=1}^{m}\|x_{ni}\|\,\mu(E\cap E_{ni}) = \int_E \|s_n\|\,d\mu.$$

En consecuencia,

$$\left\|\int_E f\,d\mu\right\| = \lim_n \left\|\int_E s_n\,d\mu\right\| \le \lim_n \int_E \|s_n\|\,d\mu = \int_E \|f\|\,d\mu,$$

ya que, como $\big|\|s_n\| - \|f\|\big| \le \|s_n - f\|$, se tiene que

$$\lim_n \int_E \big|\|s_n\| - \|f\|\big|\,d\mu = 0$$

para todo $E \in \Sigma$. □

$\boxed{\textit{Convergencia en medida}}$ Sean $f, f_n : \Omega \longrightarrow X$, $n \in \mathbb{N}$, funciones μ-medibles. Decimos que $(f_n)_n$ *converge hacia f en medida* si, para cada $\varepsilon > 0$,

$$\lim_{n\to\infty} \mu\Big(\big\{w \in \Omega : \big\|f_n(w) - f(w)\big\| > \varepsilon\big\}\Big) = 0.$$

Recuérdese que la convergencia en medida no implica, ni es implicada por, la convergencia en casi todo punto; sin embargo, toda sucesión convergente en medida tiene una subsucesión que converge en casi todo punto, [34, Chapter 3].

Demostramos ahora el teorema de la convergencia dominada para la integral de Bochner.

Teorema 5.2.3 (Teorema de la convergencia dominada)**.** *Sea $f_n : \Omega \longrightarrow X$, $n \in \mathbb{N}$, una sucesión de funciones integrables Bochner, que converge hacia una función μ-medible f en casi todo punto o en medida. Supongamos además que existe una función $g : \Omega \longrightarrow \mathbb{R}$, integrable Lebesgue, tal que $\|f_n\| \le g$ p.c.t.p. en Ω. Entonces, f es integrable Bochner y*

$$\lim_n \int_E f_n\,d\mu = \int_E f\,d\mu, \quad \textit{para todo } E \in \Sigma.$$

De hecho, se tiene que

$$\lim_n \int_\Omega \|f_n - f\|\,d\mu = 0.$$

Demostración. La sucesión de funciones escalares $\big(\|f_n - f\|\big)_n$ converge, en casi todo punto o en medida, a cero y, como $\|f_n - f\| \le 2g$ p.c.t.p. en Ω, el teorema de la convergencia dominada para funciones escalares, [34, Theorem 2.4.4 y Proposition 3.1.5], nos asegura que

$$\lim_n \int_\Omega \|f_n - f\|\,d\mu = 0.$$

Por lo tanto, f es integrable Bochner, y con ayuda de la desigualdad obtenida en la proposición 5.2.2 se concluye la demostración. □

$\boxed{\text{Medidas vectoriales}}$ Una *medida vectorial finitamente aditiva* es una aplicación $F : \Sigma \longrightarrow X$, donde (Ω, Σ) es un espacio medible y X es un espacio de Banach, que satisface

$$F(E_1 \cup E_2) = F(E_1) + F(E_2),$$

para cualesquiera miembros disjuntos E_1 y E_2 de Σ. F se dice que es *numerablemente aditiva* (respectivamente, *débilmente numerablemente aditiva*) si

$$F\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} F(E_n),$$

para cualquier sucesión $(E_n)_n$ de elementos de Σ disjuntos dos a dos, donde la serie es convergente en la topología de la norma (respectivamente, en la topología débil) de X.

$\boxed{\text{Variación de una medida vectorial}}$ Si $F : \Sigma \longrightarrow X$ es una medida finitamente aditiva, su *variación* $|F|$ es la función de conjunto dada, para cada $E \in \Sigma$, por la expresión

$$|F|(E) := \sup\left\{\sum_{A\in\pi} \big\|F(A)\big\| : \pi \in \Pi\right\}, \tag{5.2}$$

donde Π es la familia de todas las particiones finitas $\pi = \{A_1, \ldots, A_n\}$ de E, con $A_i \in \Sigma$. Si F es numerablemente aditiva, las particiones que intervienen en (5.2) pueden tomarse numerables. La medida F se dice de *variación acotada* si su variación total es finita, *i.e.*, si $|F|(\Omega) < \infty$. Una medida de variación acotada F es numerablemente aditiva si, y sólo si, su variación $|F|$ es numerablemente aditiva, [40, Section I.1].

Teorema 5.2.4. *Sea $f : \Omega \longrightarrow X$ una función integrable Bochner, y sea $F(E) = \int_E f\,d\mu$, $E \in \Sigma$. Entonces:*

(i) $\lim_{\mu(E)\to 0} F(E) = 0$.
(ii) *F es una medida vectorial numerablemente aditiva de variación acotada y, para cada $E \in \Sigma$, se tiene que*
$$|F|(E) = \int_E \|f\|\,d\mu.$$

Demostración. La afirmación en *(i)* se obtiene a partir de la desigualdad establecida en la proposición 5.2.2, teniendo en cuenta que, al ser $\|f\| \in \mathscr{L}^1(\mu)$, se tiene que $\lim_{\mu(E)\to 0} \int_E \|f\|\,d\mu = 0$, véase [34, Lema 4.2.1].

Establezcamos ahora la validez de *(ii)*. La linealidad de la integral de Bochner nos asegura que $F(E) = \int_E f\,d\mu$ es una medida vectorial finitamente aditiva. Se tiene así que

$$\left\|F\Big(\bigcup_{n=1}^{\infty} E_n\Big) - \sum_{n=1}^{m} f(E_n)\right\| = \left\|F\Big(\bigcup_{n=m+1}^{\infty} E_n\Big)\right\| \xrightarrow[m\to\infty]{} 0,$$

y por tanto, F es numerablemente aditiva. Veamos ahora la igualdad con la variación. Para cada $E \in \Sigma$ tenemos, por definición, que

$$|F|(E) = \sup\Big\{\sum_{A\in\pi} \big\|F(A)\big\| : \pi \in \Pi\Big\},$$

donde Π es la familia de todas las particiones finitas $\pi = \{A_1, \ldots, A_n\}$ de E, con $A_i \in \Sigma$. Como

$$\sum_{A\in\pi}\Big\|\int_A f\,d\mu\Big\| \le \sum_{A\in\pi}\int_A \|f\|\,d\mu = \int_E \|f\|\,d\mu,$$

al tomar supremos, concluimos la desigualdad $|F|(E) \le \int_E \|f\|\,d\mu$. La desigualdad contraria se establece como sigue. Sea $s_n : \Omega \longrightarrow X$, $n \in \mathbb{N}$, una sucesión de funciones simples, con

$$\lim_n \int_\Omega \|s_n - f\|\,d\mu = 0.$$

Fijado $\varepsilon > 0$, existe un $n_0 \in \mathbb{N}$, tal que $\int_\Omega \|s_n - f\|\,d\mu < \varepsilon$, para cada $n \ge n_0$. Fijados $n \ge n_0$ y una representación $s_n = \sum_{i=1}^m \alpha_i \chi_{A_i}$, donde $\{A_1, \ldots, A_m\}$ es una partición de Ω en Σ, consideramos la partición π' de E dada por $\pi' = \{A_1 \cap E, \ldots, A_m \cap E\}$. Se tiene entonces que

$$\sum_{A\in\pi'}\Big\|\int_A s_n\,d\mu\Big\| = \int_E \|s_n\|\,d\mu.$$

Tomemos ahora una partición π de E en Σ refinando π', de modo que

$$|F|(E) - \sum_{A\in\pi}\Big\|\int_A f\,d\mu\Big\| < \varepsilon.$$

Como todavía tenemos que

$$\sum_{A\in\pi}\Big\|\int_A s_n\,d\mu\Big\| = \int_E \|s_n\|\,d\mu,$$

y además se verifica que

$$\sum_{A\in\pi}\bigg|\Big\|\int_A f\,d\mu\Big\| - \Big\|\int_A s_n\,d\mu\Big\|\bigg| \le \int_E \|f - s_n\|\,d\mu < \varepsilon,$$

se tendrá que,

$$\Big||F|(E) - \int_E \|s_n\|\,d\mu\Big| < 2\varepsilon.$$

Esta desigualdad es válida para todo $n \ge n_0$ y, en consecuencia,

$$|F|(E) = \lim_n \int_E \|s_n\|\,d\mu = \int_E \|f\|\,d\mu. \qquad \square$$

$\boxed{\textit{Continuidad absoluta de medidas vectoriales}}$ La medida vectorial $F : \Sigma \longrightarrow X$ se dice que es *absolutamente continua* respecto de μ si $\mu(A) = 0$ implica $F(A) = 0$. Un conocido teorema de Pettis, [40, Section I.2], asegura que si F es numerablemente aditiva, entonces F es absolutamente continua respecto de μ si, y sólo si, $\lim_{\mu(A)\to 0} F(A) = 0$.

Corolario 5.2.5. *Si f y g son dos funciones integrables Bochner tales que*

$$\int_E f\,d\mu = \int_E g\,d\mu, \quad \textit{para todo } E \in \Sigma,$$

entonces $f = g$ p.c.t.p. en Ω.

Demostración. Consideremos la función integrable Bochner $f - g$. Entonces, la medida vectorial $F(E) = \int_E (f-g)\,d\mu$ es idénticamente nula. En consecuencia, su variación también lo es,

$$|F|(E) = \int_E \|f-g\|\,d\mu = 0,$$

para cada $E \in \Sigma$, y así, $\|f-g\| = 0$ p.c.t.p. en Ω; o lo que es lo mismo, $f = g$ p.c.t.p. en Ω. $\square$

STOP Obsérvese que el corolario anterior se puede obtener también como consecuencia de la proposición 5.1.8, teniendo en cuenta que f y g son medibles y escalarmente equivalentes ya que, para cada $x^* \in X^*$, se tiene que $\int_E x^* f\,d\mu = \int_E x^* g\,d\mu$, para cada $E \in \Sigma$.

Proposición 5.2.6. *Sea $f : \Omega \longrightarrow X$ una función integrable Bochner. Entonces:*

(i) Si Y es otro espacio de Banach y $T : X \longrightarrow Y$ es lineal y continua, entonces Tf también es integrable Bochner y

$$T\left(\int_E f\,d\mu\right) = \int_E Tf\,d\mu, \quad \textit{para cada } E \in \Sigma. \tag{5.3}$$

(ii) Para cada $x^* \in X^*$, la función $x^* f \in \mathscr{L}^1(\mu)$, y se tiene que

$$x^*\left(\int_E f\,d\mu\right) = \int_E x^* f\,d\mu, \quad \textit{para cada } E \in \Sigma. \tag{5.4}$$

Demostración. Sólo tenemos que demostrar *(i)*. Por continuidad de T, es claro que Tf es medible. Por otra parte, para una función simple $s = \sum_{i=1}^m \alpha_i \chi_{A_i}$ y $E \in \Sigma$, se tiene que

$$\begin{aligned} T\left(\int_E s\,d\mu\right) &= T\left(\sum_{i=1}^m \alpha_i \mu(A_i \cap E)\right) = \sum_{i=1}^m T(\alpha_i)\,\mu(A_i \cap E) \\ &= \int_E \left(\sum_{i=1}^m T(\alpha_i)\chi_{A_i}\right) d\mu = \int_E Ts\,d\mu. \end{aligned} \tag{5.5}$$

Así, si $s_n : \Omega \longrightarrow X$, $n \in \mathbb{N}$, son simples tales que $\lim \int_\Omega \|s_n - f\|\, d\mu = 0$, es claro que

$$\lim_n \int_\Omega \|Ts_n - Tf\|\, d\mu = 0.$$

En consecuencia, Tf es integrable Bochner y, como para simples se da la igualdad (5.5), un sencillo paso al límite permite obtener (5.3), y la prueba termina. $\square$

Utilizando la igualdad (5.4) se puede dar una demostración alternativa de la desigualdad establecida en la proposición 5.2.2:

$$\left\| \int_E f\, d\mu \right\| = \sup_{x^* \in B_{X^*}} x^*\left(\int_E f\, d\mu \right) = \sup_{x^* \in B_{X^*}} \int_E x^* f\, d\mu \leq \int_E \|f\|\, d\mu.$$

Corolario 5.2.7. *Sea $f : \Omega \longrightarrow X$ una función integrable Bochner. Entonces, para cada $E \in \Sigma$ con $\mu(E) > 0$, se tiene que*

$$\frac{1}{\mu(E)} \int_E f\, d\mu \in \overline{\operatorname{co}(f(E))}.$$

Demostración. Procedamos por reducción al absurdo, suponiendo que existe un $E \in \Sigma$ para el cual $\mu(E) > 0$, y con

$$\frac{1}{\mu(E)} \int_E f\, d\mu \notin \overline{\operatorname{co}(f(E))}.$$

Por el segundo teorema de separación, corolario 1.3.26, existen $x^* \in X^*$, $\alpha \in \mathbb{R}$ y $\varepsilon > 0$ tales que

$$x^*\left(\frac{1}{\mu(E)} \int_E f\, d\mu \right) \leq \alpha - \varepsilon < \alpha \leq x^* f(w), \quad \text{para todo } w \in E. \tag{5.6}$$

La última desigualdad proporciona

$$\alpha \mu(E) \leq \int_E x^* f(w)\, d\mu,$$

o lo que es lo mismo, teniendo en cuenta (5.4),

$$\alpha \leq x^*\left(\frac{1}{\mu(E)} \int_E f(w)\, d\mu \right),$$

lo cual contradice la primera desigualdad de (5.6), y termina la prueba. $\square$

La demostración que hicimos del teorema de Krein-Šmulian 3.2.1 se puede interpretar en términos de integración vectorial.

Corolario 5.2.8 (Krein-Šmulian, Topología débil)**.** *Sea X un espacio de Banach y sea K un subconjunto débilmente compacto de X. Entonces, la envoltura convexa y cerrada de K, $\overline{\operatorname{co}(K)}$, es débilmente compacta.*

Demostración. Como ya vimos en el teorema 3.2.1, es suficiente demostrar el resultado cuando X es separable, y a su vez, en este caso, es suficiente probar que cada probabilidad de Radon $\mu \in P(K)$ tiene un baricentro, véase la prueba del teorema 3.2.1 o, alternativamente, el lema 3.5.7. Razonamos ahora que, para cada $\mu \in P(K)$, la identidad $\mathrm{Id} : K \longrightarrow X$ es μ-integrable Bochner. En efecto, para cada $x^* \in X^*$, $x^*|_K$ es $\sigma(X,X^*)$-continua, y por tanto, μ-medible; por otra parte, como X es separable, el teorema de medibilidad de Pettis 5.1.4 nos asegura que Id es μ-medible. Además, al ser K acotado en norma, $\|\mathrm{Id}\|$ está acotada en K y, como μ es una probabilidad, se tiene que $\int_K \|\mathrm{Id}\|\, d\mu < \infty$. El teorema 5.2.1 garantiza que Id es integrable Bochner, y así, existe $x_\mu = \int_K \mathrm{Id}\, d\mu$. Utilizando ahora la igualdad (5.4), concluimos que

$$x^*(x_\mu) = \int_K x^*\, d\mu, \quad \text{para cada } x^* \in X^*,$$

lo que significa que x_μ es el baricentro de μ, y la prueba termina. $\square$

El teorema siguiente es, a la integral de Bochner, lo que el teorema de diferenciabilidad es a la integral de Lebesgue.

Teorema 5.2.9. *Sea $f : [0,1] \longrightarrow X$ una función integrable Bochner respecto a la medida de Lebesgue. Entonces, p.c.t. $s \in [0,1]$, se tiene que*

$$\lim_{h\to 0} \frac{1}{h} \int_s^{s+h} \big\|f(t) - f(s)\big\|\, dt = 0.$$

Consecuentemente, p.c.t. $s \in [0,1]$, se tiene que

$$\lim_{h\to 0} \frac{1}{h} \int_s^{s+h} f(t)\, dt = f(s).$$

Demostración. La segunda afirmación se deduce directamente de la primera, teniendo en cuenta la proposición 5.2.2. La primera es consecuencia del teorema de diferenciabilidad para la integral de Lebesgue, véase [34, Prop. 6.3.9]. Dado que f es μ-medible, no es restrictivo suponer que su rango es separable. Sea $\{x_n : n \in \mathbb{N}\}$ un subconjunto denso en $f\big([0,1]\big)$. Como cada función $\|f - x_n\|$ es integrable, se tiene, por [34, Prop. 6.3.9], que

$$\lim_{h\to 0} \frac{1}{h} \int_s^{s+h} \big\|f(t) - x_n\big\|\, dt = \big\|f(s) - x_n\big\|, \quad \text{p.c.t. } s \in [0,1].$$

Así, se verifica que

$$\begin{aligned} 0 &\le \limsup_{h\to 0} \frac{1}{h} \int_s^{s+h} \big\|f(t) - f(s)\big\|\, dt \\ &\le \limsup_{h\to 0} \frac{1}{h} \int_s^{s+h} \big\|f(t) - x_n\big\|\, dt + \big\|f(s) - x_n\big\| = 2\big\|f(s) - x_n\big\|, \end{aligned}$$

para casi todo s, independientemente de n (ya que se puede unir la cantidad numerable de conjuntos nulos). La densidad de $\{x_n : n \in \mathbb{N}\}$ en $f\big([0,1]\big)$ nos dice que $\inf_n \big\| f(s) - x_n \big\| = 0$, para todo $s \in [0,1]$, de donde concluimos que

$$\lim_{h\to 0} \frac{1}{h} \int_s^{s+h} \big\| f(t) - f(s) \big\| \, dt = 0, \quad \text{p.c.t. } s \in [0,1]. \qquad \square$$

$\boxed{\textit{La propiedad de Radon-Nikodým en espacios de Banach}}$ Un espacio de Banach X se dice que tiene la *propiedad de Radon-Nikodým* (brevemente, PRN) respecto de μ si, para cada medida vectorial $F : \Sigma \longrightarrow X$ de variación acotada que es absolutamente continua respecto de μ, existe una función integrable Bochner $f : \Omega \longrightarrow X$ tal que $F(E) = \int_E f \, d\mu$, para todo $E \in \Sigma$ (f se llama derivada de Radon-Nikodým de F respecto de μ). El clásico teorema de Radon-Nikodým para medidas escalares, [34, Theorem 4.2.2], establece que $X = \mathbb{R}$ tiene la propiedad de Radon-Nikodým. En el caso infinito-dimensional hay espacios que tienen la PRN y otros que no la tienen: entre los primeros están, por ejemplo, los espacios de Banach reflexivos, y entre los segundos están c_0, ℓ^∞, $C(K)$ para K compacto infinito, etc., véase [40].

El ejemplo que sigue muestra que c_0 no tiene la propiedad de Radon-Nikodým.

Ejemplo 5.2.10 (Una medida vectorial numerablemente aditiva en c_0, de variación acotada, absolutamente continua respecto a la medida de Lebesgue λ, y sin derivada de Radon-Nikodým respecto de λ). Sea λ la medida de Lebesgue en [0,1]. Para cada conjunto medible $E \subset [0,1]$, consideremos

$$\lambda_n(E) = \int_E \operatorname{sen}(2^n \pi t) \, dt.$$

El lema de Riemann-Lebesgue, [96, Teorema 5.15], nos asegura que la sucesión $\big(\lambda_n(E)\big)_n$ está en c_0, y así,

$$F : \Sigma \to c_0$$

$$E \to \big(\lambda_n(E)\big)_n$$

es una medida vectorial finitamente aditiva. Por otra parte,

$$\big\| F(E) \big\|_\infty \leq \sup_n \int_E \big| \operatorname{sen}(2^n \pi t) \big| \, dt \leq \lambda(E), \qquad \text{para todo } E \in \Sigma$$

y, en consecuencia, F es de variación acotada, numerablemente aditiva y absolutamente continua respecto a λ. Sin embargo, no tiene derivada de Radon-Nikodým respecto de λ pues, si

$$f : [0,1] \to c_0$$

$$t \to \big(f_n(t)\big)_n$$

fuese integrable Bochner y tal que $F(E) = \int_E f_n \, d\lambda$, para cada $E \in \Sigma$, por continuidad de las proyecciones tendríamos que

$$\lambda_n(E) = \int_E f_n \, d\lambda, \quad \text{para cada } E \in \Sigma.$$

Esto último implicaría que

$$f_n(t) = \operatorname{sen}(2^n \pi t), \quad \text{p.c.t. } t \in [0,1].$$

Consideremos

$$E_n = \left\{ t \in [0,1] : f_n(t) \geq \frac{1}{\sqrt{2}} \right\}.$$

Cada E_n se puede expresar como unión de 2^{n-1} intervalos cerrados disjuntos de longitudes $1/2^{n+1}$. La unión, E_n, tiene medida $1/4$ (véase la figura 5.1).

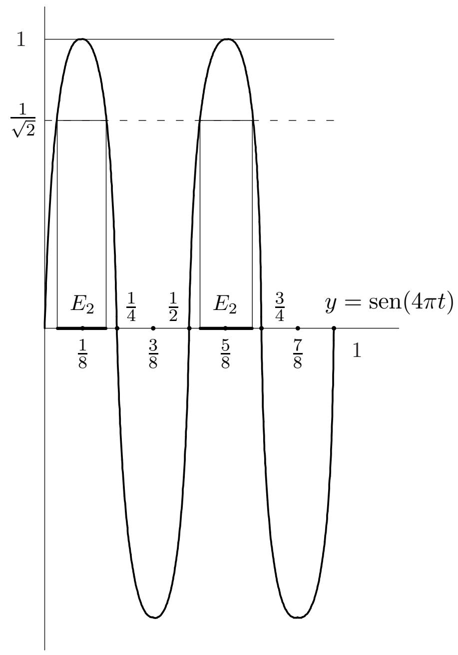


Figura 5.1: La función $y = \operatorname{sen}(4\pi t)$

Así, se tiene que $\lambda(E_n) = 1/4$. Por tanto, $\lambda\left(\limsup_j E_j\right) \geq 1/4$, donde escribimos

$$\limsup_j E_j = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n.$$

Como f toma valores en c_0, se tiene que

$$[0,1] \subset \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} E_n^c = \left(\limsup_j E_j\right)^c,$$

lo que contradice el que $\lambda\left(\limsup_j E_j\right) \geq 1/4$. $\square$

Proposición 5.2.11. *Sea $f : \Omega \longrightarrow X$ integrable Bochner. Entonces, $\left\{\int_E f\,d\mu : E \in \Sigma\right\}$ es relativamente compacto en norma.*

Demostración. Como X es completo, es suficiente ver que $\left\{\int_E f\,d\mu : E \in \Sigma\right\}$ es totalmente acotado. Para demostrar esto utilizamos la siguiente observación: *$K \subset X$ es totalmente acotado si, y sólo si, para cada $\varepsilon > 0$, existe un conjunto totalmente acotado $K_\varepsilon \subset X$ tal que $K \subset K_\varepsilon + \varepsilon B_X$*. Si $s : \Omega \longrightarrow X$ es una función simple medible, entonces, claramente, el conjunto $\left\{\int_E s\,d\mu : E \in \Sigma\right\}$ es totalmente acotado. Para f integrable, existe una sucesión de funciones simples $s_n : \Omega \longrightarrow X$ tal que $\lim_n \int_\Omega \|s_n - f\|\,d\mu = 0$. Fijado $\varepsilon > 0$, existe $N \in \mathbb{N}$ tal que $\int_\Omega \|s_N - f\|\,d\mu < \varepsilon$. Esta desigualdad y la proposición 5.2.2 implican que, para cada $E \in \Sigma$, se tiene que

$$\left\|\int_E (s_N - f)\,d\mu\right\| < \varepsilon.$$

De aquí se concluye que $\left\{\int_E f\,d\mu : E \in \Sigma\right\} \subset \left\{\int_E s_N\,d\mu : E \in \Sigma\right\} + \varepsilon B_X$, y la prueba termina. $\square$

Familias sumables Si I es un conjunto no vacío, sea $\mathscr{P}_0(I) := \{J : J \subset I,\ J \text{ es finito}\}$. El conjunto $\mathscr{P}_0(I)$ es dirigido mediante la relación de contenido conjuntista $\subset$, *i.e.*,

$$J_1 \geq J_2 \text{ en } \mathscr{P}_0(I) \text{ si, por definición, } J_2 \subset J_1.$$

Se dice que la familia $(x_i)_{i\in I}$ en el espacio normado X es *sumable*, con suma x, si la red $\left(\sum_{i\in J} x_i\right)_{J\in\mathscr{P}_0(I)}$ tiene límite x en X, y se escribe $x = \sum_{i\in I} x_i$. Cuando X es un espacio de Banach, la familia $(x_i)_{i\in I}$ es sumable si, y sólo si, la red $\left(\sum_{i\in J} x_i\right)_{J\in\mathscr{P}_0(I)}$ satisface la condición de Cauchy, que en este caso, se expresa de la siguiente forma: para cada $\varepsilon > 0$, existe $J_\varepsilon \in \mathscr{P}_0(I)$ tal que $\left\|\sum_{i\in J} x_i\right\| < \varepsilon$, para cada $J \in \mathscr{P}_0(I)$ con $J \cap J_\varepsilon = \emptyset$. Es fácil demostrar que una familia $(x_n)_{n\in\mathbb{N}}$ es sumable si, y sólo si, la serie $\sum_n x_n$ es incondicionalmente convergente, lo cual significa que, para cualquier permutación $\pi : \mathbb{N} \longrightarrow \mathbb{N}$, la serie $\sum_n x_{\pi(n)}$ es convergente al mismo valor de X. La condición de Cauchy como familia sumable permite argumentar que la serie $\sum_n x_n$ va a ser incondicionalmente convergente si, y sólo si, para cada sucesión de subconjuntos finitos no vacíos $(S_n)_n$ de $\mathbb{N}$, tales que $\max S_n < \min S_{n+1}$ para todo $n \in \mathbb{N}$, se satisface que $\liminf_n \|y_n\| = 0$, donde $y_n := \sum_{i\in S_n} x_i$ para cada $n \in \mathbb{N}$, véanse [23, p. 51-56] y [32, p. 216-251].

Teorema 5.2.12 (Orlicz-Pettis)**.** *Sea $\sum_n x_n$ una serie en el espacio de Banach X tal que,*

$$\textit{para cada } n_1 < n_2 < \ldots < n_k < \ldots \textit{ en } \mathbb{N}, \textit{ la serie } \sum_k x_{n_k} \textit{ converge débilmente.} \tag{5.7}$$

Entonces, $\sum_n x_n$ converge incondicionalmente.

Demostración. Demostramos que la familia $(x_n)_{n\in\mathbb{N}}$ es sumable. De acuerdo a nuestros comentarios anteriores, es suficiente demostrar que si $(S_n)_n$ es una sucesión de subconjuntos finitos no vacíos de $\mathbb{N}$, tales que $\max S_n < \min S_{n+1}$, para cada $n \in \mathbb{N}$, y escribimos

$$y_n := \sum_{i\in S_n} x_i,$$

para cada $n \in \mathbb{N}$, entonces $\liminf_n \|y_n\| = 0$.

Para ello, razonamos de la siguiente forma. Consideremos $\Big(\{0,1\}^{\mathbb{N}}, \mathrm{Borel}\big(\{0,1\}^{\mathbb{N}}\big), \mu\Big)$, donde μ es la probabilidad producto usual (*i.e.*, $\mu = \prod_n \mu_n$, con $\mu_n(\{0\}) = 1/2$ y $\mu_n(\{1\}) = 1/2$, para cada $n \in \mathbb{N}$). Sea $\big(\{0,1\}^{\mathbb{N}}, \Sigma, \mu\big)$ su complección. Obsérvese que, para cada $z = (z_n)_n \in \{0,1\}^{\mathbb{N}}$, la serie $\sum_n z_n y_n$ es débilmente convergente, por (5.7). Definimos entonces $\phi : \{0,1\}^{\mathbb{N}} \longrightarrow X$ mediante la expresión

$$\phi\big((z_n)_n\big) := \sigma(X,X^*)\text{-}\lim_N \sum_{n=1}^{N} z_n y_n.$$

Vamos a demostrar que ϕ es μ-integrable Bochner.

-ϕ es μ-medible: La inclusión

$$\phi\big(\{0,1\}^{\mathbb{N}}\big) \subset \overline{\mathrm{span}\{y_n : n \in \mathbb{N}\}}$$

nos dice que el rango de ϕ es separable. Demostraremos a continuación que si $\mathfrak{T}$ es la topología producto de $\{0,1\}^{\mathbb{N}}$, entonces ϕ es $\mathfrak{T}$-$\sigma(X,X^*)$-continua. Si esto es así, ϕ satisfará las hipótesis del teorema de medibilidad de Pettis 5.1.4, y por ende, concluiremos que ϕ es μ-medible.

Para demostrar que ϕ es $\mathfrak{T}$-$\sigma(X,X^*)$-continua necesitamos demostrar primero la siguiente:

Observación: *Para cada $\sigma(X,X^*)$-entorno U de 0 en X, existe $N \in \mathbb{N}$ tal que $\sum_{n\in S} z'_n y_n \in U$, para cada conjunto finito $S \subset \mathbb{N} \setminus \{1,\ldots,N\}$ y cada elección $z'_n \in \{0,1\}$, $n \in S$.*

Vemos la demostración de la observación procediendo por reducción al absurdo. Si no fuera cierta, existirían un $\sigma(X,X^*)$-entorno U de 0 en X y una sucesión de subconjuntos finitos $(T_n)_n$ no vacíos de $\mathbb{N}$, con $\max T_n < \min T_{n+1}$, junto con elecciones $z_i \in \{0,1\}$, para $i \in T_n$, $n \in \mathbb{N}$, de forma que

$$\sum_{i\in T_n} z_i y_i \notin U, \qquad \text{para cada } n \in \mathbb{N}. \tag{5.8}$$

Tomemos $z_i = 0$ si $i \in \mathbb{N} \setminus \left(\bigcup_n T_n\right)$. Entonces, la serie $\sum_i z_i y_i$ no podría ser débilmente convergente, por la propiedad (5.8). Llegamos así a una contradicción que acaba la prueba de la observación.

Veamos ahora que ϕ es $\mathfrak{T}$-$\sigma(X,X^*)$-continua. Para ello, fijamos $z = (z_n)_n \in \{0,1\}^{\mathbb{N}}$ y un $\sigma(X,X^*)$-entorno del origen U que sea absolutamente convexo y cerrado. Tomemos N como en la **observación** anterior, y sea $z' = (z'_n)_n \in \{0,1\}^{\mathbb{N}}$ tal que $z'_n = z_n$ si $1 \leq n \leq N$. Entonces,

$$\begin{aligned} \phi(z') &= \sigma(X,X^*)\text{-}\lim_p \sum_{n=1}^{p} z'_n y_n = \sum_{n=1}^{N} z_n y_n + \sigma(X,X^*)\text{-}\lim_p \sum_{n=N+1}^{p} z'_n y_n \\ &= \phi(z) - \sigma(X,X^*)\text{-}\lim_p \sum_{n=N+1}^{p} z_n y_n + \sigma(X,X^*)\text{-}\lim_p \sum_{n=N+1}^{p} z'_n y_n \\ &\in \phi(z) - U + U \subset \phi(z) + 2U, \end{aligned}$$

lo que prueba que ϕ es continua.

-ϕ es acotada: Como ϕ es $\mathfrak{T}$-$\sigma(X,X^*)$-continua, para cada $x^* \in X^*$ la función $x^*\phi$ es continua, y así, $x^*\phi\left(\{0,1\}^{\mathbb{N}}\right)$ es compacto en $\mathbb{R}$. Por lo tanto, $\phi\left(\{0,1\}^{\mathbb{N}}\right)$ es débil acotado en X, y por el teorema de la Acotación uniforme, [72, Corollary 9.4], se tiene que $\phi\left(\{0,1\}^{\mathbb{N}}\right)$ es acotado en la norma.

-ϕ es μ-integrable Bochner: La afirmación se sigue del hecho de que ϕ es μ-medible y acotada, teniendo en cuenta el teorema 5.2.1.

Definimos ahora $\psi : \{0,1\}^{\mathbb{N}} \longrightarrow X$ como $\psi := \phi \circ S$, donde $S : \{0,1\}^{\mathbb{N}} \longrightarrow \{0,1\}^{\mathbb{N}}$ es el homeomorfismo dado por

$$S\big((z_n)_n\big) := (1 - z_n)_n, \quad \text{para cada } (z_n)_n \in \{0,1\}^{\mathbb{N}}.$$

-ψ es μ-integrable Bochner: ψ es μ-medible, ya que ϕ lo es y ψ es acotada.

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

$$E_n := \left\{z \in \{0,1\}^{\mathbb{N}} : z_n = 1\right\} \in \mathrm{Borel}\big(\{0,1\}^{\mathbb{N}}\big).$$

-$\int_{E_n} (\phi - \psi)\, d\mu = (1/2)y_n$, para cada $n \in \mathbb{N}$: Podemos suponer, sin pérdida de generalidad, que $X = \mathbb{R}$, pues basta componer con cada $x^* \in X^*$, teniendo en cuenta la igualdad (5.4). Sean $f_k, g_k : \{0,1\}^{\mathbb{N}} \longrightarrow \mathbb{R}$ definidas del siguiente modo:

$$\begin{aligned} f_k(z) &= \pi_k(z) y_k = \chi_{E_k}(z) y_k, \\ g_k(z) &= \big(1 - \pi_k(z)\big) y_k = \chi_{E_k^c}(z) y_k, \end{aligned}$$

donde π_k es la k-ésima proyección canónica. Para cada $n \neq k$ se tiene que

$$\begin{aligned} \int_{E_n} (f_k - g_k)\, d\mu &= \int_{E_n} f_k\, d\mu - \int_{E_n} g_k\, d\mu \\ &= \mu(E_n \cap E_k) y_k - \mu(E_n \cap E_k^c) y_k = \frac{1}{4} y_k - \frac{1}{4} y_k = 0. \end{aligned} \tag{5.9}$$

Si escribimos $s_p := \sum_{k=1}^{p}(f_k - g_k)$, $p \in \mathbb{N}$, cada s_p es una función simple, y se verifica que $\text{lím}_p s_p(z) = \phi(z) - \psi(z)$, para cada $z \in \{0,1\}^{\mathbb{N}}$. Por otro lado se tiene que

$$s_p(z) \in \phi\big(\{0,1\}^{\mathbb{N}}\big) - \psi\big(\{0,1\}^{\mathbb{N}}\big),$$

con lo que $\sup\big\{|s_p(z)| : p \in \mathbb{N}, z \in \{0,1\}^{\mathbb{N}}\big\} < \infty$. Utilizando el teorema de la convergencia dominada 5.2.3, tenemos que

$$\text{lím}_p \int_{E_n} s_p \, d\mu = \int_{E_n} (\phi - \psi) \, d\mu.$$

Además, gracias a (5.9), para cada $p \geq n$ se verifica que

$$\int_{E_n} s_p \, d\mu = \sum_{k=1}^{p} \int_{E_n} (f_k - g_k) \, d\mu = \int_{E_n} (f_n - g_n) \, d\mu,$$

y por lo tanto,

$$\int_{E_n} (\phi - \psi) \, d\mu = \int_{E_n} (f_n - g_n) \, d\mu = \int_{E_n} (y_n \chi_{E_n} - y_n \chi_{E_n^c}) \, d\mu = \mu(E_n) y_n = \frac{1}{2} y_n,$$

como se quería demostrar.

-$\text{lím}\inf_n \|y_n\| = 0$**:** Al ser $\phi - \psi$ μ-integrable Bochner, después de la proposición 5.2.11 sabemos que el conjunto $K := \big\{\int_E (\phi - \psi) \, d\mu : E \in \Sigma\big\}$ es relativamente compacto en norma. Teniendo en cuenta lo demostrado en el apartado anterior, $\{y_n : n \in \mathbb{N}\} \subset 2K$. En consecuencia, existe una subsucesión (y_{n_k}) que converge en norma hacia un cierto $x \in X$. Como, por otro lado, existe $\sigma(X,X^*)$-$\text{lím}_p \sum_{n=1}^{p} y_n$, concluimos que existe $\sigma(X,X^*)$-$\text{lím}_n y_n$, y que vale cero. Así, $x = 0$, y finalmente, $\text{lím}\inf_n \|y_n\| = 0$. □

Como consecuencia del teorema de Orlicz-Pettis, obtenemos el siguiente resultado que utilizaremos en la proposición 5.3.4.

Corolario 5.2.13. *Sean (Ω, Σ) un espacio medible y $F : \Sigma \longrightarrow X$ una medida débilmente numerablemente aditiva. Entonces, F es numerablemente aditiva para la norma de X.*

Demostración. Sea $(E_n)_n$ una sucesión de elementos de Σ disjuntos dos a dos. Para cada sucesión creciente $n_1 < n_2 < \ldots < n_k < \ldots$, la serie $\sum_k F(E_{n_k})$ converge débilmente hacia $F\big(\bigcup_{k=1}^{\infty} E_{n_k}\big)$. Consecuentemente, la serie $\sum_n F(E_n)$ satisface las hipótesis del teorema de Orlicz-Pettis 5.2.12, y por tanto, $\sum_n F(E_n)$ es incondicionalmente convergente en norma hacia $F\big(\bigcup_{n=1}^{\infty} E_n\big)$, con lo que la prueba termina. □

Acabamos esta sección con algunos comentarios sobre los espacios $L^p(\mu, X)$.

$\boxed{\textit{Espacios de funciones p-integrables Bochner}}$ Para $1 \le p < \infty$, representamos por $\mathscr{L}^p(\mu)$ el espacio de las funciones escalares p-integrables, y si $p = \infty$, escribimos $\mathscr{L}^\infty(\mu)$ para denotar el espacio de las funciones μ-esencialmente acotadas. $\mathscr{L}^p(\mu, X)$ es el conjunto de las funciones μ-medibles de Ω en X tales que $\|f\| \in \mathscr{L}^p(\mu)$; para $f \in \mathscr{L}^p(\mu, X)$ definimos

$$\|f\|_p := \left(\int_\Omega \|f\|^p \, d\mu \right)^{1/p} \quad (= \big\| \|f\| \big\|_p \text{ en } \mathscr{L}^p(\mu)),$$

cuando $1 \le p < \infty$, y si $p = \infty$,

$$\|f\|_\infty := \big\| \|f\| \big\|_\infty \quad (\text{en } \mathscr{L}^\infty(\mu)).$$

Cada $\|\cdot\|_p$ es una seminorma, y se tiene que, para $f \in \mathscr{L}^p(\mu, X)$, se da la igualdad $\|f\|_p = 0$ si, y sólo si, $f = 0$ p.c.t.p. en Ω. Si definimos la relación de equivalencia $f \sim g$ si, y sólo si, $f = g$ p.c.t.p. en Ω, el espacio cociente $L^p(\mu, X) := \mathscr{L}^p(\mu, X)/\sim$ es un espacio de Banach cuando se dota de la norma cociente asociada a $\|\cdot\|_p$. Se satisface la inclusión

$$\mathscr{L}^q(\mu, X) \subset \mathscr{L}^p(\mu, X), \quad \text{para } q \ge p.$$

El conjunto de las funciones simples medibles es denso en $\big(L^p(\mu, X), \|\cdot\|_p\big)$, si $1 \le p < \infty$, y el conjunto de las funciones de $L^\infty(\mu, X)$ que toman una cantidad numerable de valores, es denso en este último espacio dotado de la norma $\|\cdot\|_\infty$. La descripción de los espacios duales es más problemática que en el caso de los espacios de funciones escalares, y se tiene, por ejemplo, que para $1 \le p < \infty$ se da la igualdad $\big(L^p(\mu, X), \|\cdot\|_p\big)^* = L^q(\mu, X^*)$ si, y sólo si, X^* tiene la PRN respecto de μ, véase [40, Theorem 1, p. 98].

5.3 La integral de Pettis

NO es posible desarrollar una teoría de integración similar a la de Bochner pero para funciones débilmente medibles. Hay, sin embargo, una forma simple de definir (llevados por la noción de baricentro, véase la página 130) una integral para funciones débilmente medibles, que fue introducida por Pettis y que ha tenido numerosas e importantes aplicaciones dentro del Análisis Funcional.

Un resultado importante para comprender parte de lo que se presenta en esta sección es el siguiente teorema de Dunford.

Teorema 5.3.1 (Dunford)**.** *Sea $f : \Omega \longrightarrow X$ una función débilmente medible, con $x^* f \in \mathscr{L}^1(\mu)$ para cada $x^* \in X^*$. Entonces, la aplicación*

$$\begin{aligned} X^* &\longrightarrow \mathbb{R} \\ x^* &\to \int_E x^* f \, d\mu \end{aligned}$$

es un elemento del bidual X^{**}, para cada $E \in \Sigma$.

Demostración. Observemos, en primer lugar, que la aplicación

$$T : L^1(\mu) \longrightarrow \mathbb{K}$$

$$g \to \int_\Omega g \, d\mu$$

es lineal y continua. Por lo tanto, si vemos que para cada $E \in \Sigma$, el funcional lineal

$$Q_E : X^* \longrightarrow L^1(\mu)$$

$$x^* \to x^* f \chi_E$$

es continuo, ya habremos terminado la demostración: el funcional del que queremos probar su continuidad es, precisamente, $T \circ Q_E$. Gracias al teorema de la Gráfica Cerrada, [72, Theorem 10.3], demostrar la continuidad de Q_E equivale a probar que su grafo es cerrado. Si suponemos que

$$\left.\begin{array}{ll} x_n^* \longrightarrow x^* & \text{en } X^*, \\ x_n^* f \chi_E \to g & \text{en } L^1(\mu), \end{array}\right\}$$

entonces existe una subsucesión $(x_{n_k}^*)_k$ de $(x_n^*)_n$ de forma que $\left(x_{n_k}^* f \chi_E\right)_k$ converge a g p.c.t.p. en Ω, [96, Teorema 3.11]. En consecuencia, como $\left(x_{n_k}^* f \chi_E\right)_k$ converge hacia $x^* f \chi_E$ en todos los puntos de Ω, se tiene que

$$g = x^* f \chi_E, \quad \text{p.c.t.p. en } \Omega,$$

y por tanto, serán iguales en $L^1(\mu)$, quedando así terminada la demostración. □

Integral de Dunford Se dice que una función $f : \Omega \longrightarrow X$ es *integrable Dunford*, si es débilmente medible y, para cada $x^* \in X^*$, se tiene que $x^* f \in \mathscr{L}^1(\mu)$. En este caso, para cada $E \in \Sigma$, el elemento del bidual definido por

$$x_E^{**} : X^* \longrightarrow \mathbb{K}$$

$$x^* \to \int_E x^* f \, d\mu$$

se llama *integral de Dunford* de f sobre E, y se escribe

$$x_E^{**} = (D) - \int_E f \, d\mu.$$

Integral de Pettis En las condiciones de la definición anterior, si $(D) - \int_E f \, d\mu \in X$, para cada $E \in \Sigma$, se dice que la función anterior es *integrable Pettis*, y el valor de la integral de Dunford sobre cada subconjunto $E \in \Sigma$ se denomina, en este caso, *integral de Pettis*.

Si $f : \Omega \longrightarrow X$ es una función integrable Bochner, la igualdad (5.4) nos dice que f es integrable Pettis, y para cada conjunto $E \in \Sigma$, la integral de Bochner sobre E de f es su integral de Pettis. En general, se tiene que

$$\text{Integrable Bochner} \Rightarrow \text{Integrable Pettis} \Rightarrow \text{Integrable Dunford},$$

y ninguna de estas implicaciones se puede volver hacia atrás.

Ejemplo 5.3.2 (Una función integrable Dunford que no es integrable Pettis). Consideremos la σ-álgebra de conjuntos medibles Lebesgue en $[0,1]$, y sea

$$\begin{aligned} f:[0,1] &\longrightarrow c_0 \\ t &\to \big(\chi_{(0,1]}(t), 2\chi_{(0,1/2]}(t), \ldots, n\chi_{(0,1/n]}(t), \ldots\big). \end{aligned}$$

Entonces, f es integrable Dunford y no es integrable Pettis.

Demostración. Para cada $x^* = (\alpha_n)_n \in \ell^1 = (c_0)^*$, se tiene que

$$x^* f(t) = \sum_{n=1}^{\infty} \alpha_n n \chi_{(0,1/n]}(t),$$

donde sólo un número finito de sumandos es no nulo. Además,

$$\sum_{n=1}^{\infty} \int_0^1 \left|\alpha_n n \chi_{(0,1/n]}(t)\right| dt = \sum_{n=1}^{\infty} |\alpha_n| < \infty.$$

El teorema de Beppo Levi, [34, Corollary 2.4.2], nos asegura que $x^* f \in \mathscr{L}^1(\lambda)$, y que

$$\int_0^1 x^* f \, dt = \sum_{n=1}^{\infty} \alpha_n n \int_0^1 \chi_{(0,1/n]}(t)\, dt = \sum_{n=1}^{\infty} \alpha_n = x^{**}(x^*),$$

donde $x^{**} \in \ell^\infty = (c_0)^{**}$ es el elemento definido como $x^{**} = (1,1,\ldots,1,\ldots)$. Así, tenemos que

$$(D) - \int_0^1 f \, dt = (1,1,\ldots,1,\ldots) \in \ell^\infty \setminus c_0,$$

y por tanto, f es integrable Dunford pero no integrable Pettis. $\square$

Ejemplo 5.3.3 (Una función integrable Pettis que no es integrable Bochner). La función

$$\begin{aligned} f:\ [0,1] &\longrightarrow \ell^2\big([0,1]\big) \\ t &\to e_t, \end{aligned}$$

es integrable Pettis respecto a la medida de Lebesgue λ, pero no es integrable Bochner.

Demostración. Según hemos establecido en el ejemplo 5.1.7, para cada $x^* \in \big(\ell^2([0,1])\big)^*$ se tiene que $x^* f = 0$ p.c.t.p. en $[0,1]$. Consecuentemente, f es débilmente medible e integrable Pettis, con integral nula sobre cada conjunto medible. Tal y como vimos en el mencionado ejemplo, f no es λ-medible, y así, en particular, no puede ser integrable Bochner. □

Proposición 5.3.4. *Sea $f : \Omega \longrightarrow X$ una función integrable Pettis. Entonces, su integral indefinida $F : \Sigma \longrightarrow X$ dada por*

$$F(E) = (P) - \int_E f\, d\mu, \quad E \in \Sigma,$$

es una medida numerablemente aditiva en Σ que satisface

$$\lim_{\mu(E)\to 0} F(E) = 0. \tag{5.10}$$

Demostración. Dados $(E_n)_n$ una sucesión de elementos de Σ disjuntos dos a dos y $E := \bigcup_n E_n$,

$$\begin{aligned} x^*\big(F(E)\big) &= x^*\left((P) - \int_E f\, d\mu\right) = \int_{\bigcup_n E_n} x^* f\, d\mu \\ &= \sum_n \int_{E_n} x^* f\, d\mu = \sum_n x^*\left((P) - \int_{E_n} f\, d\mu\right) = \sum_n x^*\big(F(E_n)\big). \end{aligned}$$

Esto demuestra que F es una medida débilmente numerablemente aditiva y, después del corolario 5.2.13, concluimos que F es numerablemente aditiva en norma. Por ser F y μ numerablemente aditivas en una σ-álgebra, para demostrar (5.10) es suficiente probar, por el teorema de Pettis, que $\mu(E) = 0$ implica $F(E) = 0$, lo que resulta evidente a partir de la definición de la integral de Pettis, ya que si $\mu(E) = 0$, entonces, para cada $x^* \in X^*$,

$$x^*\big(F(E)\big) = \int_E x^* f\, d\mu = 0.$$ □

La proposición anterior no es cierta para funciones integrables Dunford.

Ejemplo 5.3.5 (Una función integrable Dunford cuya integral indefinida no es numerablemente aditiva). Sea f la función definida en el ejemplo 5.3.2, y consideremos la sucesión decreciente de conjuntos medibles $(0,1/n] \searrow \emptyset$. Para cada $m \in \mathbb{N}$ fijo y cada $x^* = (\alpha_n)_n \in \ell^1 = (c_0)^*$, se tiene que

$$\int_0^{1/m} x^* f\, dt = \sum_{n=1}^{m-1} \alpha_n \frac{n}{m} + \sum_{n=m}^{\infty} \alpha_n = x_m^{**}(x^*),$$

donde $x_m^{**} = \left(\frac{1}{m}, \frac{2}{m}, \dots, \frac{m-1}{m}, 1, 1, \dots\right)$. Así, se verifica que

$$(D) - \int_{(0,1/m]} f\, dt = \left(\frac{1}{m}, \dots, \frac{m-1}{m}, 1, 1, \dots\right),$$

de donde se deduce que

$$\left\| (D) - \int_{(0,1/m]} f\, dt \right\|_{\ell^\infty} = 1,$$

para todo $m \in \mathbb{N}$, no pudiendo ser F, por lo tanto, una medida numerablemente aditiva. □

El siguiente teorema aclara la relación existente entre integrabilidad Dunford y Pettis de una función medible, y el hecho de que la integral indefinida sea una medida numerablemente aditiva.

Teorema 5.3.6. *Sea $f : \Omega \longrightarrow X$ una función medible e integrable Dunford. Las siguientes condiciones son equivalentes:*

(i) *f es integrable Pettis.*

(ii) *La integral indefinida de Dunford, $F(E) = (D) - \int_E f\, d\mu$, $E \in \Sigma$, es una medida numerablemente aditiva.*

Demostración. La implicación *(i)* $\Rightarrow$ *(ii)* es la proposición 5.3.4. Veamos ahora cómo *(ii)* $\Rightarrow$ *(i)*. Dado que f es μ-medible, existen $E \in \Sigma$ nulo y $E_n \nearrow \Omega \setminus E$, con $E_n \in \Sigma$, tales que $f\chi_{E_n}$ está acotada. Por lo tanto, $f\chi_{E_n}$ es integrable Bochner, y así, para cada $E \in \Sigma$,

$$(D) - \int_{E\cap E_n} f\, d\mu = (D) - \int_E f\chi_{E_n}\, d\mu = (B) - \int_E f\chi_{E_n}\, d\mu \in X.$$

Por otra parte, al ser F numerablemente aditiva, existe el límite

$$\lim_n (D) - \int_{E\cap E_n} f\, d\mu \in X.$$

Ahora bien, como para cada $x^* \in X^*$ se tiene que $x^* f \in \mathscr{L}^1(\mu)$,

$$\begin{aligned} x^*\left(\lim_n (D) - \int_{E\cap E_n} f\, d\mu\right) &= \lim_n x^*\left((D) - \int_{E\cap E_n} f\, d\mu\right) \\ &= \lim_n \int_{E\cap E_n} x^* f\, d\mu = \int_E x^* f\, d\mu = \left((D) - \int_E f\, d\mu\right)(x^*), \end{aligned}$$

y por tanto, $\lim_n (D) - \int_{E\cap E_n} f\, d\mu = (D) - \int_E f\, d\mu \in X$, como queríamos demostrar. $\square$

Integración en espacios de Banach.

«En el desarrollo de las teorías de integración y diferenciación de funciones definidas en un espacio euclídeo con valores en un espacio de Banach B, las figuras pioneras, después de los artículos de Bochner de 1933, fueron Dunford y Gel'fand. Fue el estudio de la diferenciación de funciones lo que hizo aparecer el espacio de Banach B. A pesar de que algunas funciones, tales como las dadas por integrales de Bochner, eran diferenciables en casi todo punto independientemente de B, muchas otras no lo eran, y su diferenciabilidad dependía de las características de los espacios rango; más precisamente, dependía de propiedades asociadas al conjunto rango de la

función como subconjunto de B (Clarkson inventó los espacios uniformemente convexos con el propósito de que tuvieran propiedades de diferenciabilidad universal; los espacios de Banach reflexivos reaparecieron con el mismo propósito). Además, la teoría de diferenciación, al margen de su interés intrínseco, fue fundamental en los esfuerzos para representar operadores lineales por medio de integrales: teoremas tipo Radon-Nikodým.

B. J. Pettis (sacado del preámbulo de [40]).»

Las propiedades básicas de la integral de Pettis fueron establecidas por B. J. Pettis en 1938, aunque en 1936 algunas de estas propiedades ya habían sido estudiadas por N. Dunford: correctamente, la integral de Pettis debería llamarse *segunda integral de Dunford*.

5.4 Conexión con el proyecto de investigación: las propiedades de Radon-Nikodým y débil de Radon-Nikodým

UN espacio de Banach X se dice que tiene la propiedad de Radon-Nikodým –brevemente PRN (respectivamente, la propiedad débil de Radon-Nikodým –brevemente PDRN–) si, para cada espacio de medida finito completo (Ω,Σ,μ), y cada medida vectorial $F:\Sigma\longrightarrow X$ de variación acotada y absolutamente continua respecto de μ (*i.e.*, si $E\in\Sigma$ y $\mu(E)=0$, entonces $|F|(E)=0$; en este caso se escribe $|F|\ll\mu$), existe una función $f:\Omega\longrightarrow X$ integrable Bochner (respectivamente, integrable Pettis) tal que

$$F(E)=\int_E f\,d\mu, \quad \text{para cada } E\in\Sigma.$$

La PDRN es estrictamente más general que la PRN, [104, Chapter VIII].

La propiedad de Radon-Nikodým se puede caracterizar de dos formas: vía el concepto de *dentabilidad* y vía el concepto de *martingala acotada*, véase [40, Chapter V]. De cualquiera de estas dos caracterizaciones se sigue que un espacio de Banach con la PRN no puede contener un ε-árbol acotado infinito.

ε-árbol Un *ε-árbol* infinito en X, véase la página 57, es una sucesión $(x_n)_n$ infinita en X tal que $x_n=(x_{2n}+x_{2n+1})/2$, para todo $n\in\mathbb{N}$.

En general, no es cierto que un espacio de Banach que no contiene ε-árboles infinitos acotados tenga la PRN: Bourgain y Rosenthal construyeron en $L^1\big([0,1]\big)$ un subespacio cerrado sin la PRN que no contenía ε-árboles infinitos, [13]. Sorprendentemente, para espacios de Banach duales, la situación es mucho más placentera: la combinación de resultados debidos a Fitzpatrick, Huff y Morris, Stegall, Namioka y Namioka-Phelps, permiten establecer el siguiente resultado:

Teorema 5.4.1 (Theorem 4.2.13, [15])**.** *Para un espacio de Banach X, las siguientes afirmaciones son equivalentes:*

(i) X^* *tiene la PRN.*
(ii) X^* *no contiene* ε*-árboles infinitos para ningún* $\varepsilon > 0$.
(iii) *Para cada subespacio separable* $Y \subset X$*, el dual* Y^* *es separable.*
(iv) *Cada subconjunto* $\sigma(X^*,X)$*-compacto de* B_{X^*} *tiene* $\sigma(X^*,X)$*-abiertos relativos de diámetro, en norma, arbitrariamente pequeño.*
(v) *Para cada* $D \subset B_{X^*}$ $\sigma(X^*,X)$*-compacto, la identidad* $\mathrm{Id} : \big(D, \sigma(X^*,X)\big) \longrightarrow \big(D, \|\cdot\|\big)$ *tiene un punto de continuidad sobre* D*.*
(vi) X *es un espacio de Asplund,* i.e.*, cada función real, continua y convexa, definida en un subconjunto abierto y convexo de* X*, es diferenciable Fréchet en un* G_δ *denso de su dominio.*

La caracterización de los espacios duales con la PDRN es también espectacular, y en ella han contribuido numerosos matemáticos: Musial, Odell y Rosenthal, Jacnica y Haydon, entre otros:

Teorema 5.4.2 (Chapter VI en [104])**.** *Para un espacio de Banach* X*, las siguientes afirmaciones son equivalentes:*

(i) X^* *tiene la PDRN.*
(ii) $\ell^1 \not\subset X$.
(iii) *La inyección canónica* $\mathrm{i} : B_{X^*} \longrightarrow X^*$ *es integrable Pettis para cada medida de Radon en* $\big(B_{X^*}, \sigma(X^*,X)\big)$.
(iv) *Para cada conjunto* $\sigma(X^*,X)$*-compacto* $K \subset X^*$*, se tiene que* $\overline{\mathrm{co}(K)}^{\sigma(X^*,X)} = \overline{\mathrm{co}(K)}^{\|\cdot\|}$.
(v) *Para cada conjunto* $\sigma(X^*,X)$*-compacto y convexo* $K \subset X^*$*, se tiene que* $K = \overline{\mathrm{co}(\mathrm{Ext}(K))}^{\|\cdot\|}$.

El teorema 5.4.1 llama especialmente la atención porque conecta propiedades de naturaleza extremadamente distinta: mientras *(i)* tiene que ver con teoría de la medida, *(iv)* es puramente topológica y *(vi)* puramente analítica; *(iii)* establece una condición práctica que se puede comprobar en distintos casos concretos, y es la propiedad que liga entre sí a todas las demás.

Fragmentabilidad Sean (Y,τ) un espacio topológico, d una métrica en Y y $\varepsilon > 0$; el espacio Y se dice que está ε*-fragmentado* por d si, para cada subconjunto no vacío C de Y, existe un subconjunto τ-abierto V de Y con $C \cap V \neq \emptyset$ y d-diam$(C \cap V) < \varepsilon$, [66]. Cuando Y está ε-fragmentado por d para cada $\varepsilon > 0$, decimos que Y está *fragmentado* por d.

σ-fragmentabilidad Un espacio topológico (X,τ) se dice que es σ*-fragmentable* por una métrica d si, para cada $\varepsilon > 0$, $X = \bigcup_{n=1}^{\infty} X_{n,\varepsilon}$, donde $X_{n,\varepsilon}$ está ε-fragmentado por d.

Los subconjuntos acotados de un espacio de Banach con la PRN están fragmentados por la norma; el recíproco no es cierto. En el teorema 5.4.1, la condición *(iv)* se lee: $\big(B_{X^*}, \sigma(X^*,X)\big)$ *está fragmentada por la métrica asociada a la norma dual.* La clase de los espacios de Banach que, dotados con su topología débil, son σ-fragmentados por la norma, ha sido extensamente estudiada por Jayne, Namioka y Rogers, véanse [64, 65]. Namioka, [82], estudió intensivamente los compactos que son fragmentados por métricas inferiormente semicontinuas, demostrando que éstos son homeomorfos a subconjuntos débil*-compactos de duales de espacios de Asplund. Hay una gran variedad de resultados conectando fragmentabilidad y σ-fragmentabilidad con la existencia de buenos renormamientos. Por ejemplo, si X admite un renormamiento diferenciable Gâteaux, entonces $\big(B_{X^*}, \sigma(X^*, X)\big)$ está fragmentado. Desde otro punto de vista, la noción de σ-fragmentabilidad ha desempeñado un papel importante en cuestiones de selectores y diferenciabilidad en espacios de Banach. Los artículos [24] y [25] ofrecen nuevas caracterizaciones topológicas de los conceptos de fragmentabilidad y σ-fragmentabilidad, y contienen numerosas aplicaciones topológicas y analíticas. La noción de fragmentabilidad se puede utilizar, por ejemplo, como herramienta auxiliar para estudiar el problema de la frontera presentado en la página 131, tal y como se expone en [29].

La PDRN es, incluso para espacios de Banach duales, estrictamente más general que la PRN: la integral de Pettis es estrictamente más general que la integral de Bochner. Entre la integral de Bochner y la integral de Pettis existe una noción de integral que es la versión vectorial de la reinterpretación que Fréchet hizo, en 1915, de la integral de Lebesgue. Fréchet consideró en [52] funciones $f : \Omega \longrightarrow \mathbb{R}$ y, para cada partición $\Gamma = (A_n)_n$ de Ω en una cantidad contable (finita o numerable) de elementos de Σ, asignó una integral *superior* y una integral *inferior* relativas mediante las expresiones

$$J^*(f,\Gamma) = \sum_n \mu(A_n) \sup f(A_n) \quad \text{y} \quad J_*(f,\Gamma) = \sum_n \mu(A_n) \operatorname{ínf} f(A_n),$$

respectivamente, suponiendo que ambas series están bien definidas y son absolutamente convergentes. Se tiene la desigualdad $J_*(f,\Gamma) \le J^*(f,\Gamma')$ siempre que $J_*(f,\Gamma)$ y $J^*(f,\Gamma')$ estén definidas. Por tanto, la intersección de los *rangos integrales relativos*

$$J_*(f,\Gamma) \le x \le J^*(f,\Gamma),$$

para Γ variable, es no vacío. Esta intersección es un único punto x si, y sólo si, f es integrable Lebesgue y $x = \int_\Omega f \, d\mu$.

Las ideas de Fréchet inspiraron a Birkhoff para dar la siguiente definición:

Integral de Birkhoff Sea $f : \Omega \longrightarrow X$ una función. Si $\Gamma = (A_n)_n$ es una partición contable de Ω en Σ, la función f se dice *sumable* respecto de Γ si la restricción $f|_{A_n}$ es acotada cuando $\mu(A_n) > 0$, y el conjunto de sumas

$$J(f,\Gamma) = \left\{ \sum_n \mu(A_n) f(t_n) : t_n \in A_n \right\}$$

está formado por series incondicionalmente convergentes. La función f se dice *integrable Birkhoff* si, para cada $\varepsilon > 0$, existe una partición contable $\Gamma = (A_n)_n$ de Ω en Σ para la que f es sumable y $\|\cdot\|$-diam$\big(J(f,\Gamma)\big) < \varepsilon$. Cuando f es integrable según la definición previa, la *integral de Birkhoff* de f se define como el único punto en la intersección

$$\bigcap\Big\{\overline{\mathrm{co}\big(J(f,\Gamma)\big)} : f \text{ es sumable respecto de } \Gamma\Big\},$$

véase [8, Theorem 12].

Como se ha indicado antes, la noción de integrabilidad Birkhoff es intermedia entre las de Bochner y Pettis, en el sentido de que, para una función $f:\Omega\longrightarrow X$, se tiene que:

$$f \text{ integrable Bochner} \implies f \text{ integrable Birkhoff} \implies f \text{ integrable Pettis},$$

véase [8]. Ninguno de los recíprocos es válido en general, aunque integrabilidad Birkhoff y Pettis coinciden cuando X es separable. Si f es integrable Birkhoff, entonces sus integrales de Birkhoff y de Pettis coinciden.

En [27] se ha estudiado la integral de Birkhoff, poniendo de manifiesto que, lo que durante muchos años se había considerado como condición suficiente para caracterizar integrabilidad Pettis, realmente proporciona una caracterización de la integrabilidad Birkhoff. Como consecuencia, se demuestra el siguiente resultado, que supone una mejora cualitativa de la PDRN:

Teorema 5.4.3 ([27]). *Para un espacio de Banach X, son equivalentes:*

(i) *X^* tiene la PDRN.*

(ii) *Para cada espacio de medida finito completo (Ω,Σ,μ), y cada medida vectorial $F:\Sigma\longrightarrow X$ de variación acotada con $|F|\ll\mu$, existe una función $f:\Omega\longrightarrow X$ integrable Birkhoff de forma que*

$$F(E)=\int_E f\,d\mu, \quad \text{para cada } E\in\Sigma.$$

En estos últimos años (2000–actualidad), hay una corriente creciente de trabajos dedicados al estudio de la integral de Pettis en el ámbito de las multifunciones, véanse [45, 62]. La integral de Bochner fue transportada a multifunciones por Debreu en 1967, quien la utilizó para reemplazar *sumas* en ciertos modelos matemáticos en Economía. El artículo [26] avanza algunas cuestiones nuevas sobre la integral de Pettis de multifunciones, y propone la integral de Birkhoff de multifunciones como noción intermedia entre la de Debreu y la de Pettis.

PARA SABER MÁS

- Los libros [40], [42] y [103] son tres buenas referencias para el estudio de cuestiones tratadas en este capítulo. El primero de ellos presenta, de forma organizada, numerosos aspectos concernientes a integración vectorial, medidas vectoriales, propiedad de Radon-Nikodým y diversas aplicaciones de tipo analítico y geométrico en espacios de Banach. El segundo libro estudia, de forma exhaustiva, medidas vectoriales, considerando especialmente las que se definen en espacios topológicos. También presta atención a las diversas nociones de medibilidad asociadas a funciones vectoriales y a los espacios L^p vectoriales, *lifting*, etc. El libro [103] pasa revista, de forma autocontenida, a diversos aspectos de la teoría de la medida e integración en espacios de Banach, proporcionando criterios generales para asegurar integrabilidad Pettis en espacios de Banach, y analizando con detalle la propiedad débil de Radon-Nikodým. El estudio que se lleva a cabo se basa en herramientas topológicas y de teoría de la medida abstracta. Los resultados más espectaculares se obtienen a partir de profundos resultados sobre conjuntos compactos de funciones medibles.

► Los libros [15] y [104] son dos referencias básicas para el estudio de las propiedades de Radon-Nikodým y débil de Radon-Nikodým en espacios de Banach. En el primero de ellos, se estudia de forma detallada la propiedad de Radon-Nikodým y propiedades relacionadas, tanto a nivel global de los espacios involucrados, como a nivel localizado de subconjuntos de los mismos. Se presentan los resultados centrales mediante los cuales se caracterizan los espacios de Asplund X como aquéllos para los que X^* tiene la propiedad de Radon-Nikodým. El libro [104] está dedicado al estudio sistemático de los espacios X que no contienen ℓ^1, a la postre, condición equivalente a que X^* tenga la propiedad débil de Radon-Nikodým. Resultados sobre existencia de *lifting* de funciones, funciones de la primera clase de Baire, espacios de medidas de Radon, criterios de compacidad puntual en espacios de funciones medibles, etc., se presentan con gran detalle.

Índice terminológico

Bibliografía

[1] D. Amir and J. Lindenstrauss, *The structure of weakly compact sets in Banach spaces*, Ann. Math. **88** (1968), no. 2, 35–46. MR 37 #4562

[2] T. M. Apostol, *Análisis matemático*, segunda ed., Editorial Reverté, s.a., 1976.

[3] N. Aronszajn and K. T. Smith, *Invariant subspaces of completely continuous operators*, Ann. of Math. (2) **60** (1954), 345–350. MR 16,488b

[4] E. Asplund, *Fréchet differentiability of convex functions*, Acta Math. **121** (1968), 31–47. MR 37 #6754

[5] S. Banach, *Théorie des óperations linéaires*, Chelsea Publishing Co., New York, 1955. MR 17,175h

[6] B. Beauzamy, *Un opérateur sans sous-espace invariant: simplification de l'exemple de P. Enflo*, Integral Equations Operator Theory **8** (1985), no. 3, 314–384. MR 88b:47011

[7] A. R. Bernstein and A. Robinson, *Solution of an invariant subspace problem of K. T. Smith and P. R. Halmos*, Pacific J. Math. **16** (1966), 421–431. MR 33 #1724

[8] G. Birkhoff, *Integration of functions with values in a Banach space*, Trans. Amer. Math. Soc. **38** (1935), no. 2, 357–378. MR 1 501 815

[9] S. Bochner, *Integration von Funktionen, deren Werte die Elemente eines Vektorraumes sind*, Fund. Math. **20** (1933), 262–270.

[10] W. M. Boothby, *An introduction to differentiable manifolds and Riemannian geometry*, second ed., Pure and Applied Mathematics, vol. 120, Academic Press Inc., Orlando, FL, 1986. MR 87k:58001

[11] N. Bourbaki, *Elementos de historia de las matemáticas*, Alianza Universidad, Madrid, 1976.

[12] J. Bourgain, D. H. Fremlin, and M. Talagrand, *Pointwise compact sets of Baire-measurable functions*, Amer. J. Math. **100** (1978), no. 4, 845–886. MR 80b:54017

[13] J. Bourgain and H. P. Rosenthal, *Martingales valued in certain subspaces of L^1*, Israel J. Math. **37** (1980), no. 1-2, 54–75. MR 82g:46044

[14] J. Bourgain and M. Talagrand, *Compacité extrémale*, Proc. Amer. Math. Soc. **80** (1980), no. 1, 68–70. MR 81h:46011

[15] R. D. Bourgin, *Geometric aspects of convex sets with the Radon-Nikodým property*, Lecture Notes in Mathematics, vol. 993, Springer-Verlag, Berlin, 1983. MR 85d:46023

[16] H. Brézis and F. E. Browder, *A general principle on ordered sets in nonlinear functional analysis*, Advances in Math. **21** (1976), no. 3, 355–364. MR 54 #13641

[17] A. L. Brown, *Some problems in linear analysis*, Ph.D. thesis, Cambridge University, 1961.

[18] J. Caristi, *Fixed point theorems for mappings satisfying inwardness conditions*, Trans. Amer. Math. Soc. **215** (1976), 241–251. MR 52 #15132

[19] H. Cartan, *Teoría elemental de las funciones analíticas de una y varias variables complejas*, Selecciones científicas. Madrid, 1968.

[20] ______, *Cálculo diferencial*, Ediciones Omega S. A. Barcelona, 1978.

[21] B. Cascales and G. Godefroy, *Angelicity and the boundary problem*, Mathematika **45** (1998), no. 1, 105–112. MR 99f:46019

[22] B. Cascales, G. Manjabacas, and G. Vera, *A Krein-Šmulian type result in Banach spaces*, Quart. J. Math. Oxford Ser. (2) **48** (1997), no. 190, 161–167. MR 99c:46009

[23] B. Cascales and J. M. Mira, *Análisis funcional*, ICE- Universidad de Murcia - DM, 2002.

[24] B. Cascales and I. Namioka, *The Lindelöf property and σ-fragmentability*, Fund. Math. **180** (2003), no. 2, 161–183. MR MR2064322 (2005b:54036)

[25] B. Cascales, I. Namioka, and J. Orihuela, *The Lindelöf property in Banach spaces*, Studia Math. **154** (2003), no. 2, 165–192. MR MR1949928 (2003m:54028)

[26] B. Cascales and J. Rodríguez, *Birkhoff integral for multi-valued functions*, J. Math. Anal. Appl. **297** (2004), no. 2, 540–560, Special issue dedicated to John Horváth. MR 2088679 (2005f:26021)

[27] ______, *The Birkhoff integral and the property of Bourgain*, Math. Ann. **331** (2005), no. 2, 259–279. MR 2115456

[28] B. Cascales and R. Shvydkoy, *On the Krein-Šmulian theorem for weaker topologies*, Illinois J. Math. **47** (2003), no. 4, 957–976. MR MR2036985 (2004m:46044)

[29] B. Cascales and G. Vera, *Topologies weaker than the weak topology of a Banach space*, J. Math. Anal. Appl. **182** (1994), no. 1, 41–68. MR 95c:46017

[30] ______, *Norming sets and compactness*, Rocky Mountain J. Math. **25** (1995), no. 3, 919–925. MR 96i:46011

[31] M. Castellet and I. Llerena, *álgebra lineal y geometría*, Reverté, 1991.

[32] G. Choquet, *Topology*, Translated from the French by Amiel Feinstein. Pure and Applied Mathematics, Vol. XIX, Academic Press, New York, 1966. MR 33 #1823

[33] ______, *Lectures on analysis. Vol. II: Representation theory*, Edited by J. Marsden, T. Lance and S. Gelbart, W. A. Benjamin, Inc., New York-Amsterdam, 1969. MR 40 #3253

[34] D. L. Cohn, *Measure theory*, Birkhäuser, Boston, Mass., 1980. MR 81k:28001

[35] J. B. Conway, *Functions of one complex variable*, second ed., Graduate Texts in Mathematics, vol. 11, Springer-Verlag, New York, 1978. MR 80c:30003

[36] ______, *A course in functional analysis*, Graduate Texts in Mathematics, vol. 96, Springer-Verlag, New York, 1985. MR 86h:46001

[37] F. K. Dashiell and J. Lindenstrauss, *Some examples concerning strictly convex norms on $C(K)$ spaces*, Israel J. Math. **16** (1973), 329–342. MR 50 #964

[38] R. Deville, G. Godefroy, and V. Zizler, *Smoothness and renormings in Banach spaces*, Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 64, Longman Scientific & Technical, Harlow, 1993. MR 94d:46012

[39] J. Diestel, *Sequences and series in Banach spaces*, Graduate Texts in Mathematics, vol. 92, Springer-Verlag, New York, 1984. MR 85i:46020

[40] J. Diestel and J. J. Uhl Jr, *Vector measures*, Mathematical Surveys, vol. 15, American Mathematical Society, Providence, R.I., 1977, With a foreword by B. J. Pettis. MR 56 #12216

[41] J. Dieudonné, *History of functional analysis*, North-Holland Mathematics Studies, vol. 49, North-Holland Publishing Co., Amsterdam, 1981, Notas de Matemática [Mathematical Notes], 77. MR 83d:46001

[42] N. Dinculeanu, *Vector measures*, International Series of Monographs in Pure and Applied Mathematics, Vol. 95, Pergamon Press, Oxford, 1967. MR 34 #6011b

[43] N. Dunford and J. T. Schwartz, *Linear Operators. I. General Theory*, With the assistance of W. G. Bade and R. G. Bartle. Pure and Applied Mathematics, Vol. 7, Interscience Publishers, Inc., New York, 1958. MR 22 #8302

[44] I. Ekeland, *On the variational principle*, J. Math. Anal. Appl. **47** (1974), 324–353. MR 49 #11344

[45] K. El Amri and C. Hess, *On the Pettis integral of closed valued multifunctions*, Set-Valued Anal. **8** (2000), no. 4, 329–360. MR 1802239 (2002e:26025)

[46] P. Enflo, *Banach spaces which can be given an equivalent uniformly convex norm*, Proceedings of the International Symposium on Partial Differential Equations and the Geometry of Normed Linear Spaces (Jerusalem, 1972), vol. 13, 1972, pp. 281–288 (1973). MR 49 #1073

[47] ———, *On the invariant subspace problem for Banach spaces*, Acta Math. **158** (1987), no. 3-4, 213–313. MR 88j:47006

[48] R. Engelking, *General topology*, PWN—Polish Scientific Publishers, Warsaw, 1977, Translated from the Polish by the author, Monografie Matematyczne, Tom 60. [Mathematical Monographs, Vol. 60]. MR 58 #18316b

[49] M. Fabian, *Gâteaux differentiability of convex functions and topology*, Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley & Sons Inc., New York, 1997, Weak Asplund spaces, A Wiley-Interscience Publication. MR 98h:46009

[50] M. Fabian, P. Habala, P. Hájek, V. Montesinos Santalucía, J. Pelant, and V. Zizler, *Functional analysis and infinite-dimensional geometry*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, vol. 8, Springer-Verlag, New York, 2001. MR 2002f:46001

[51] K. Floret, *Weakly compact sets*, Lecture Notes in Mathematics, vol. 801, Springer, Berlin, 1980, Lectures held at S.U.N.Y., Buffalo, in Spring 1978. MR 82b:46001

[52] M. Frechet, *Sur l'intégrale d'une fonctionnelle étendue à un ensemble abstrait*, Bull. Soc. Math. France **43** (1915), 248–265.

[53] G. Godefroy, *Boundaries of a convex set and interpolation sets*, Math. Ann. **277** (1987), no. 2, 173–184. MR 88f:46037

[54] ———, *Five lectures in geometry of Banach spaces*, Seminar on Functional Analysis, 1987 (Murcia, 1987), Notas Mat., vol. 1, Univ. Murcia, Murcia, 1988, pp. 9–67. MR 93a:46015

[55] K. Goebel and W. A. Kirk, *Topics in metric fixed point theory*, Cambridge Studies in Advanced Mathematics, vol. 28, Cambridge University Press, Cambridge, 1990. MR 92c:47070

[56] A. Granas and J. Dugundji, *Fixed point theory*, Springer Monographs in Mathematics, Springer-Verlag, New York, 2003. MR 2004d:58012

[57] L. M. Graves, *Riemann integration and Taylor's theorem in general analysis*, Trans. Amer. Math. Soc. **29** (1927), no. 1, 163–177. MR 1501382

[58] A. Grothendieck, *Critères de compacité dans les espaces fonctionnels généraux*, Amer. J. Math. **74** (1952), 168–186. MR 13,857e

[59] P. R. Halmos, *Invariant subspaces of polynomially compact operators*, Pacific J. Math. **16** (1966), 433–437. MR 33 #1725

[60] G. Hämmerlin and K-H. Hoffmann, *Numerical mathematics*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1991, Translated from the German by Larry Schumaker, Readings in Mathematics. MR 92d:65001

[61] R. Haydon, *Trees in renorming theory*, Proc. London Math. Soc. **78** (1999), no. 3, 541–584. MR 2000d:46011

[62] C. Hess, *Set-valued integration and set-valued probability theory: an overview*, Handbook of measure theory, Vol. I, II, North-Holland, Amsterdam, 2002, pp. 617–673. MR 1954624 (2004g:60020)

[63] T. H. Hildebrandt, *Integration in abstract spaces*, Bull. Amer. Math. Soc. **59** (1953), 111–139. MR 0053191 (14,735c)

[64] J. E. Jayne, I. Namioka, and C. A. Rogers, *σ-fragmentable Banach spaces*, Mathematika **39** (1992), no. 2, 197–215. MR 94c:46028

[65] ______, *Topological properties of Banach spaces*, Proc. London Math. Soc. (3) **66** (1993), no. 3, 651–672. MR 94c:46041

[66] J. E. Jayne and C. A. Rogers, *Borel selectors for upper semicontinuous set-valued maps*, Acta Math. **155** (1985), no. 1-2, 41–79. MR 87a:28011

[67] W. F. Donoghue Jr., *The lattice of invariant subspaces of a completely continuous quasi-nilpotent transformation*, Pacific J. Math. **7** (1957), 1031–1035. MR 19,1066f

[68] J. L. Kelley, *The Tychonoff product theorem implies the axiom of choice*, Fund. Math. **37** (1950), 75–76. MR 12,626d

[69] ______, *Topología general*, Editorial Universitaria de Buenos Aires, Buenos Aires, 1962. MR 33 #6566

[70] W. A. Kirk, *Caristi's fixed point theorem and metric convexity*, Colloq. Math. **36** (1976), no. 1, 81–86. MR 55 #9061

[71] G. Köthe, *Topological vector spaces. I*, Translated from the German by D. J. H. Garling. Die Grundlehren der mathematischen Wissenschaften, Band 159, Springer-Verlag New York Inc., New York, 1969. MR 40 #1750

[72] B. V. Limaye, *Functional analysis*, Wiley Eastern Ltd., New Delhi, 1981. MR 83b:46001

[73] J. Lindenstrauss, *On nonseparable reflexive Banach spaces*, Bull. Amer. Math. Soc. **72** (1966), 967–970. MR 34 #4875

[74] J. Lindenstrauss and L. Tzafriri, *On the complemented subspaces problem*, Israel J. Math. **9** (1971), 263–269. MR 43 #2474

[75] V. I. Lomonosov, *Invariant subspaces of the family of operators that commute with a completely continuous operator*, Funkcional. Anal. i Priložen. **7** (1973), no. 3, 55–56. MR 54 #8319

[76] ______, *A counterexample to the Bishop-Phelps theorem in complex spaces*, Israel J. Math. **115** (2000), 25–28. MR 2000k:46016

[77] A. J. Michaels, *Hilden's simple proof of Lomonosov's invariant subspace theorem*, Adv. Math. **25** (1977), no. 1, 56–58. MR 58 #17893

[78] A. Moltó, J. Orihuela, and S. Troyanski, *Locally uniformly rotund renorming and fragmentability*, Proc. London Math. Soc. (3) **75** (1997), no. 3, 619–640. MR 98e:46011

[79] A. Moltó, J. Orihuela, S. Troyanski, and M. Valdivia, *A non linear transfer technique for renorming*, To appear in Lectures Notes in Mathematics, 2006.

[80] K. Musiał, *Topics in the theory of Pettis integration*, Rend. Istit. Mat. Univ. Trieste **23** (1991), no. 1, 177–262 (1993), School on Measure Theory and Real Analysis (Grado, 1991). MR 1248654 (94k:46084)

[81] K. Musial, *Pettis integral*, Handbook of measure theory, Vol. I, II, North-Holland, Amsterdam, 2002, pp. 531–586. MR 1954622 (2004d:28026)

[82] I. Namioka, *Radon-Nikodým compact spaces and fragmentability*, Mathematika **34** (1987), no. 2, 258–281. MR 89i:46021

[83] L. Pasicki, *A short proof of the Caristi theorem*, Comment. Math. Prace Mat. **20** (1977/78), no. 2, 427–428. MR 80a:54086

[84] B. J. Pettis, *On integration in vector spaces*, Trans. Amer. Math. Soc. **44** (1938), no. 2, 277–304. MR 1 501 970

[85] R. R. Phelps, *Convex functions, monotone operators and differentiability*, second ed., Lecture Notes in Mathematics, vol. 1364, Springer-Verlag, Berlin, 1993. MR 94f:46055

[86] ______, *Lectures on Choquet's theorem*, second, 1975 ed., Lecture Notes in Mathematics, vol. 1757, Springer-Verlag, Berlin, 2001. MR 2002k:46001

[87] F. Puerta, *álgebra lineal*, Marcombo, 1975.

[88] H. Radjavi and P. Rosenthal, *Invariant subspaces*, Springer-Verlag, New York, 1973, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 77. MR 51 #3924

[89] M. Raja, *Locally uniformly rotund norms*, Mathematika **46** (1999), no. 2, 343–358. MR 2002e:46019

[90] C. J. Read, *A solution to the invariant subspace problem*, Bull. London Math. Soc. **16** (1984), no. 4, 337–401. MR 86f:47005

[91] ______, *A short proof concerning the invariant subspace problem*, J. London Math. Soc. (2) **34** (1986), no. 2, 335–348. MR 87m:47020

[92] J. Rodríguez, *On the existence of Pettis integrable functions which are not Birkhoff integrable*, Proc. Amer. Math. Soc. **133** (2005), no. 4, 1157–1163. MR 2117218 (2005k:28021)

[93] C. A. Rogers, *A less strange version of Milnor's proof of Brouwer's fixed-point theorem*, Amer. Math. Monthly **87** (1980), no. 7, 525–527. MR 82b:55004

[94] N. M. Roy, *Extreme points of convex sets in infinite-dimensional spaces*, Amer. Math. Monthly **94** (1987), no. 5, 409–422. MR 89a:52001

[95] W. Rudin, *Análisis funcional*, Reverté, 1979.

[96] ______, *Análisis real y complejo*, tercera ed., McGraw-Hill, 1988.

[97] B. R. Salinas and F. Bombal, *The Tychonoff product theorem for compact Hausdorff spaces does not imply the axiom of choice: a new proof. Equivalent propositions*, Collect. Math. **24** (1973), 219–230. MR 49 #3812

[98] R. Schneider, *Convex bodies: the Brunn-Minkowski theory*, Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1993. MR 94d:52007

[99] Z. Semadeni, *Banach spaces of continuous functions. Vol. I*, PWN—Polish Scientific Publishers, Warsaw, 1971, Monografie Matematyczne, Tom 55. MR 45 #5730

[100] J. Siegel, *A new proof of Caristi's fixed point theorem*, Proc. Amer. Math. Soc. **66** (1977), no. 1, 54–56. MR 56 #16606

[101] S. Simons, *A convergence theorem with boundary*, Pacific J. Math. **40** (1972), 703–708. MR 47 #755

[102] I. Singer, *Best approximation in normed linear spaces by elements of linear subspaces*, Translated from the Romanian by Radu Georgescu. Die Grundlehren der mathematischen Wissenschaften, Band

171, Publishing House of the Academy of the Socialist Republic of Romania, Bucharest, 1970. MR 42 #4937

[103] M. Talagrand, *Pettis integral and measure theory*, Mem. Amer. Math. Soc. **51** (1984), no. 307, ix+224. MR 86j:46042

[104] D. van Dulst, *Characterizations of Banach spaces not containing l^1*, CWI Tract, vol. 59, Centrum voor Wiskunde en Informatica, Amsterdam, 1989. MR 90h:46037

[105] L. Vašák, *On one generalization of weakly compactly generated Banach spaces*, Studia Math. **70** (1981), no. 1, 11–19. MR 83h:46028

[106] G. Vera, *Introducción al análisis complejo*, Publicaciones del Departamento Matemáticas, Universidad de Murcia. 169 págs., 2000.

[107] E. Zeidler, *Nonlinear functional analysis and its applications. I*, Springer-Verlag, New York, 1986, Fixed-point theorems, Translated from the German by Peter R. Wadsack. MR 87f:47083

[108] ______, *Applied functional analysis. Applications to mathematical physics*, Applied Mathematical Sciences, vol. 108, Springer-Verlag, New York, 1995. MR 96i:00005

[109] ______, *Applied functional analysis. Main principles and their applications*, Applied Mathematical Sciences, vol. 109, Springer-Verlag, New York, 1995. MR 96i:00006

[110] M. Zisman, *Topología algebraica elemental*, Paraninfo, Madrid, 1979, Translated from the French by Emilio Romero Ros. MR 84e:55001

[111] V. Zizler, *Nonseparable Banach spaces*, Handbook of the Geometry of Banach spaces vol. 2., Elsevier Science B. V., 2003, pp. 1745–1816. MR 1 229 122