Haïm Brézis

Análisis funcional
Teoría y aplicaciones

Versión española de
Juan Ramón Esteban

Alianza
Editorial

Título original:
Analyse fonctionnelle. La edición original de esta obra ha sido publicada en francés por Masson Editeur, de París.

© Masson, París, 1983
© Ed. cast.: Alianza Editorial, S. A., Madrid, 1984
Calle Milán, 38; ☎ 200 00 45
ISBN: 84-206-8088-5
Depósito legal: M. 34.699-1984
Fotocomposición: EFCA. Avda. de Pablo Iglesias, 17. 28003 Madrid
Impreso en Artes Gráficas Ibarra, S. A., Matilde Hernández, 31. 28019 Madrid
Printed in Spain

INDICE

NOTACIONES

Notaciones generales

E'	espacio dual de E
$<\,,\,>$	producto escalar en la dualidad E′, E.
$[f = \alpha] = \{x\,;f(x) = \alpha\}$	
$B(x_0, r) = \{x\,;\,\|x - x_0\| < r\}$ bola abierta, centrada en x_0, de radio r	
$B_E = \{x \in E;\ \|x\| \leqslant 1\}$	
epi $\varphi = \{(x, \lambda);\ \varphi(x) \leqslant \lambda\}$ epigrafo de φ	
φ^*	función conjugada de φ
$\mathscr{L}(E, F)$	espacio de los operadores continuos de E en F
$M^\perp$	ortogonal de M
D(A)	dominio del operador A
G(A)	grafo del operador A
N(A)	núcleo del operador A
R(A)	imagen del operador A
$\sigma(E, E')$	topología débil definida en E
$\sigma(E', E)$	topología débil * definida en E′
$\rightharpoonup$	convergencia débil
J	inyección canónica de E en E″
p'	exponente conjugado de p, es decir, $\frac{1}{p} + \frac{1}{p'} = 1$
c.t.p.	para casi todo punto
$\|A\|$	medida (de Lebesgue) del conjunto A
Supp f	soporte de la función f
$f * g$	producto de convolución
ρ_n	sucesión regularizante
$(\tau_h f)(x) = f(x + h)$	trasladada de la función f
$\omega \subset\subset \Omega$	abierto ω fuertemente incluido en Ω, es decir $\bar{\omega}$ es compacto y $\bar{\omega} \subset \Omega$.
P_K	proyección sobre el convexo cerrado K
$\|\ \|$	norma Hilbertiana

$\rho(T)$ conjunto resolvente del operador T
$\sigma(T)$ espectro del operador T
$VP(T)$ valores propios del operador T
$J_\lambda = (I + \lambda A)^{-1}$ resolvente del operador A
$A_\lambda = AJ_\lambda$ regularización Yosida del operador A

$$\nabla u = \left(\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \ldots \frac{\partial u}{\partial x_N}\right) = \text{grad } u$$

$$D^\alpha u = \frac{\partial^{\alpha_1 + \alpha_2 + \ldots \alpha_N}}{\partial x_1^{\alpha_1} . x_2^{\alpha_2} \ldots \partial x_N^{\alpha_N}} u, \quad |\alpha| = \sum_{i=1}^{N} \alpha_i$$

$$\Delta u = \sum_{i=1}^{N} \frac{\partial^2 u}{\partial x_i^2} = \text{Laplaciano de } u$$

$\mathbb{R}^N_+ = \{x = (x', x_N) \in \mathbb{R}^{N-1} \times \mathbb{R},\ x_N > 0\}$
$Q = \{x = (x', x_N) \in \mathbb{R}^{N-1} \times \mathbb{R};\ |x'| < 1 \text{ y } |x_N| < 1\}$
$Q_+ = Q \cap \mathbb{R}^N_+$
$Q_0 = \{x \in Q;\ x_N = 0\}$

$$(D_h u)(x) = \frac{1}{|h|}(u(x+h) - u(x))$$

$\dfrac{\partial u}{\partial n}$ derivada normal exterior

Espacios funcionales

$\Omega \subset \mathbb{R}^N$ abierto
$\partial\Omega = \Gamma =$ frontera de Ω
$L^p(\Omega) = \{u$ medible en Ω y $\int_\Omega |u|^p\, dx < \infty\}$, $1 \le p < \infty$,
$L^\infty(\Omega) = \{u$ medible en Ω y existe C tal que $|u(x)| \le C$ c.t.p. en $\Omega\}$
$C_c(\Omega)$ funciones continuas con soporte compacto en Ω
$C^k(\Omega)$ funciones k veces continuamente diferenciables en Ω ($k \ge 0$)
$C^\infty(\Omega) = \bigcap_{k \geq 0} C^k(\Omega)$
$C^k_c(\Omega) = C^k(\Omega) \cap C_c(\Omega)$
$C^\infty_c(\Omega) = C^\infty(\Omega) \cap C_c(\Omega) = \mathscr{D}(\Omega)$
$C(\bar\Omega)$ funciones continuas en $\bar\Omega$
$C^k(\bar\Omega)$ funciones u de $C^k(\Omega)$ tales que para cada multi-índice α, $|\alpha| \le k$, la aplicación $x \in \Omega \mapsto D^\alpha u(x)$ se extiende con continuidad a $\bar\Omega$
$C^\infty(\bar\Omega) = \bigcap_k C^k(\bar\Omega)$

$$C^{0,\alpha}(\bar\Omega) = \left\{u \in C(\Omega);\ \sup_{x,y\in\Omega} \frac{|u(x) - u(y)|}{|x-y|^\alpha} < \infty\right\} \text{ con } 0 < \alpha < 1$$

$C^{k,\alpha}(\bar\Omega) = \{u \in C^k(\Omega);\ D^j u \in C^{0,\alpha}(\bar\Omega)\ \ \forall j,\ |j| \leq k\}$
$W^{1,p}$, $W^{1,p}_0$, $W^{m,p}$, H^1, H^1_0, H^m espacios de Sobolev.

INTRODUCCION

Este texto recoge en una forma sensiblemente más elaborada un curso de Maîtrise impartido en la Universidad Pierre y Marie Curie (París VI). Supone conocidos los elementos básicos de la Topología General, de la Integración y del Cálculo Diferencial.

La primera parte del curso (Capítulos I a VII) desarrolla los resultados «abstractos» del Análisis Funcional. La segunda parte (Capítulos VIII a X) se dedica al estudio de espacios funcionales «concretos» que intervienen en la teoría de ecuaciones en derivadas parciales; en ella se muestra cómo los teoremas de existencia «abstractos» permiten resolver ecuaciones en derivadas parciales. Estas dos ramas del Análisis están estrechamente ligadas. Históricamente, el Análisis Funcional «abstracto» se desarrolló en primer lugar para responder a cuestiones planteadas por la resolución de ecuaciones en derivadas parciales. Inversamente, los progresos del Análisis Funcional «abstracto» han estimulado considerablemente la teoría de ecuaciones en derivadas parciales. Este curso no contiene ninguna referencia histórica; recomendamos al lector la consulta del texto de J. Dieudonné [3]. Deseamos que este libro pueda ser útil tanto a los estudiantes interesados en las «Matemáticas Puras» como a quienes deseen orientarse hacia las «Matemáticas Aplicadas».

Mi agradecimiento a:

—M. G. Tronel, que me sugirió numerosas mejoras.

—Mm. Ph. Ciarlet y P. Rabinowitz por sus valiosos consejos y estímulos.

—Mm. Berestycki, Gallouet, Kavian, Mc Intosh por sus útiles comentarios.

—El Mathematics Reseach Center, University of Wisconsin, y al Department of Mathematics, University of Chicago, donde fueron redactadas algunas partes de este libro.

Dedico este libro a la memoria de Guido Stampacchia, en homenaje a un Maestro del Análisis Funcional, prematuramente desaparecido.

H. BREZIS

Observaciones

1) La notación [BT] hace referencia a la obra de H. Brézis-G. Tronel, *Analyse Fonctionnelle, Recueil de Problèmes et Exercices* Masson.

Algunos resultados *enunciados* en este volumen *se demuestran* en ejercicios de [BT].

2) Ciertos enunciados o párrafos aparecen precedidos del símbolo •; se trata de pasajes **muy importantes.** El símbolo * precede a ciertos enunciados que se pueden omitir en primera lectura.

3) Hemos adoptado una numeración continua para las proposiciones, teoremas y corolarios; únicamente los lemas están numerados de forma separada.

4) En todo el texto consideramos sólo *espacios vectoriales sobre* $\mathbb{R}$ (lo cual es reprobable, pero simplifica la presentación). La mayoría de los enunciados son válidos para los espacios vectoriales sobre $\mathbb{C}$; a veces son necesarias algunas modificaciones. En [BT] se establece la lista de cambios a introducir cuando se trabaja en espacios vectoriales sobre $\mathbb{C}$.

Capítulo I
LOS TEOREMAS DE HAHN-BANACH. INTRODUCCION A LA TEORIA DE LAS FUNCIONES CONVEXAS CONJUGADAS

I.1. Forma analítica del teorema de Hahn-Banach: extensión de formas lineales

Sea E un espacio vectorial sobre $\mathbb{R}$. Recordemos que una **forma lineal** es una aplicación lineal definida sobre E, o sobre un subespacio vectorial de E, **con valores en** $\mathbb{R}$. El resultado esencial de §I.1 hace referencia a la extensión de una forma definida sobre un subespacio vectorial de E a una forma lineal definida sobre todo E.

Teorema I.1 (Hahn-Banach, forma analítica).—*Sea p: $E \to \mathbb{R}$ una aplicación que verifica*

$$p(\lambda x) = \lambda p(x) \qquad \forall x \in E \qquad y \qquad \forall \lambda > 0, \tag{1}$$

$$p(x + y) \leqslant p(x) + p(y) \qquad \forall x, y \in E. \tag{2}$$

Sean también $G \subset E$ *un subespacio vectorial y* g: $G \to \mathbb{R}$ *una aplicación lineal tal que*

$$g(x) \leqslant p(x) \qquad \forall x \in G. \tag{3}$$

Entonces existe una forma lineal f *definida sobre* E *que extiende a* g, *i.e.*

$$g(x) = f(x) \qquad \forall x \in G$$

y tal que

$$f(x) \leqslant p(x) \qquad \forall x \in E. \tag{4}$$

La demostración del teorema I.1 requiere la utilización del lema de Zorn, cuyo enunciado recordamos. Comencemos precisando algunas nociones de la teoría de los conjuntos ordenados.

Sea P un conjunto dotado de un relación de orden (parcial) notada $\leqslant$. Se dice que un subconjunto $Q \subset P$ está **totalmente ordenado** si para todo a, b de Q se tiene (al menos) una de las relaciones $a \leqslant b$ o $b \leqslant a$.

Sea Q $\subset$ P un subconjunto de P; se dice que $c \in$ P es una **cota superior de Q** si para todo $a \in$ Q se tiene $a \leqslant c$.

Se dice que $m \in$ P es un elemento **maximal** de P si para todo $x \in$ P tal que $m \leqslant x$ se tiene necesariamente $x = m$.

Finalmente, se dice que P es **inductivo** si todo subconjunto totalmente ordenado de P admite una cota superior.

Lema I.1 (Zorn).—*Todo conjunto ordenado, inductivo y no vacío admite un elemento maximal.*

Se encontrará una demostración del lema de Zorn (a partir del Axioma de Elección) en N. Dunford-J. Schwartz [1] (Vol. 1, Teorema 1.2.7.) o bien en P. Dubreil-M. L. Dubreil Jacotin [1] (Cap. 6).

NOTA 1.—No es indispensable para un analista conocer la demostración del lema de Zorn, sin embargo es **esencial** entender bien su enunciado y saberlo utilizar. El lema de Zorn tiene numerosas y muy importantes aplicaciones en Análisis; es una herramienta indispensable para establecer ciertos resultados de **existencia.**

DEMOSTRACIÓN DEL TEOREMA I.1. - Se considera el conjunto

$$\mathrm{P} = \left\{ h \;\middle|\; \begin{array}{l} h\colon \mathrm{D}(h) \subset \mathrm{E} \to \mathbb{R} \text{ con } \mathrm{D}(h) \text{ subespacio vectorial de E, } h \text{ lineal,} \\ G \subset \mathrm{D}(h),\ h \text{ extiende a } g \text{ y } h(x) \leq p(x) \quad \forall x \in \mathrm{D}(h) \end{array} \right\}.$$

P está dotado de la relación de orden

$$(h_1 \leq h_2) \Leftrightarrow (\mathrm{D}(h_1) \subset \mathrm{D}(h_2) \text{ y } h_2 \text{ extiende a } h_1).$$

Es claro que P no es vacío, ya que $g \in$ P. Por otra parte, P es inductivo. En efecto sea Q $\subset$ P un subconjunto totalmente ordenado; denotado por Q $= (h_i)_{i\in \mathrm{I}}$. Se define

$$\mathrm{D}(h) = \bigcup_{i\in \mathrm{I}} \mathrm{D}(h_i) \qquad \text{y} \qquad h(x) = h_i(x) \quad \text{si} \quad x \in \mathrm{D}(h_i).$$

Se comprueba que esta definición tiene sentido, que $h \in$ P y que h es una cota superior de Q. Resulta del lema de Zorn que P admite un elemento maximal, notado por f. Probemos que D(f) = E — lo que terminará la demostración del teorema I.1. Razonemos por reducción al absurdo y supongamos que D(f) $\neq$ E. Sea $x_0 \in$ D(f); pongamos D(h) = D(f) + $\mathbb{R}x_0$ y para $x \in$ D(f), $h(x + tx_0) = f(x) + t\alpha$ ($t \in \mathbb{R}$) donde α es una constante que se fijará posteriormente de forma que $h \in$ P. Nos debemos asegurar de que

$$f(x) + t\alpha \leqslant p(x + tx_0) \qquad \forall x \in \mathrm{D}(f), \quad \forall t \in \mathbb{R}.$$

Gracias a (1) basta comprobar que

$$\begin{cases} f(x) + \alpha \leqslant p(x + x_0) & \forall x \in \mathrm{D}(f) \\ f(x) - \alpha \leqslant p(x - x_0) & \forall x \in \mathrm{D}(f). \end{cases}$$

O lo que es igual, hay que elegir α verificando

$$\operatorname*{Sup}_{y\in \mathrm{D}(f)} \{f(y) - p(y - x_0)\} \leqslant \alpha \leqslant \operatorname*{Inf}_{x\in \mathrm{D}(f)} \{p(x + x_0) - f(x)\}.$$

Tal elección es posible ya que

$$f(y) - p(y - x_0) \leqslant p(x + x_0) - f(x) \qquad \forall x \in D(f), \quad \forall y \in D(f);$$

en efecto, obsérvese que

$$f(x) + f(y) \leqslant p(x + y) \leqslant p(x + x_0) + p(y - x_0)$$

gracias a (2).
Se concluye que f está mayorada por h y que $f \neq h$; esto contradice la maximalidad de f.

Indiquemos ahora algunas aplicaciones sencillas del teorema I.1 cuando E es un **espacio vectorial normado** (e.v.n.) de norma $\| \ \|$.
Notación: Se designa por E′ el dual (topológico) ([1]) de E, i.e. el espacio de las **formas lineales y continuas sobre E;** E′ está dotado de la **norma dual** ([2])

(5)
$$\|f\|_{E'} = \sup_{\substack{x \in E \\ \|x\| \leqslant 1}} |f(x)| = \sup_{\substack{x \in E \\ \|x\| \leqslant 1}} f(x).$$

Cuando $f \in E'$ y $x \in E$ se notará generalmente $\langle f, x \rangle$ en lugar de $f(x)$; se dice que $\langle \, , \rangle$ es el **producto escalar en la dualidad** E′, E.

• **Corolario 1.2.**—*Sea G un subespacio vectorial de* E, *y sea* g: G → ℝ *una aplicación lineal y continua de norma*

$$\|g\|_{G'} = \sup_{\substack{x \in G \\ \|x\| \leqslant 1}} g(x).$$

Entonces existe $f \in E'$ *que extiende a* g *y tal que*

$$\|f\|_{E'} = \|g\|_{G'}.$$

DEMOSTRACION.—Aplicar el teorema I.1 con $p(x) = \|g\|_{G'} \|x\|$.

• **Corolario 1.3**—*Para todo* $x_0 \in E$ *existe* $f_0 \in E'$ *tal que*

$$\|f_0\| = \|x_0\| \qquad y \qquad \langle f_0, x_0 \rangle = \|x_0\|^2.$$

DEMOSTRACIÓN.—Aplicar el corolario I.2 con $G = \mathbb{R}x_0$ y $g(tx_0) = t\|x_0\|^2$ de forma que $\|g\|_{G'} = \|x_0\|$.

NOTA 2.—El elemento f_0 definido en el corolario I.3 no es único en general (intentar construir un ejemplo o ver [BT]). Sin embargo, si E′ es **estrictamente convexo** ([3]) — lo cual es cierto por ejemplo si E es un espacio de Hilbert (ver capítulo V) o si $E = L^p(\Omega)$ con $1 < p < \infty$ (ver capítulo IV) — entonces f_0 es único. De forma general, se designa, para cada $x_0 \in E$,

$$F(x_0) = \{f_0 \in E'; \ \|f_0\| = \|x_0\| \ \text{y} \ \langle f_0, x_0 \rangle = \|x_0\|^2\}.$$

([1]) En la literatura americana el dual topológico de E se designa por E*. ¡Atención a las confusiones!
([2]) En general se escribirá simplemente $\|f\|$ en lugar de $\|f\|_{E'}$, excepto si hay ambigüedad.
([3]) Se dice que un espacio vectorial normado E es **estrictamente convexo** si $\forall x, y \in E$ con $\|x\| = \|y\| = 1$ y $x \neq y$ se tiene $\|tx + (1 - t)y\| < 1$ $\forall t \in]0, 1[$.

La aplicación (multívoca) $x_0 \longmapsto F(x_0)$ es **la aplicación de dualidad** de E en E'; se encontrarán algunas de sus propiedades en [BT].

• **Corolario 1.4.**—*Para todo $x \in E$ se tiene*

$$\|x\| = \sup_{\substack{f \in E' \\ \|f\| \leq 1}} |\langle f, x\rangle| = \max_{\substack{f \in E' \\ \|f\| \leq 1}} |\langle f, x\rangle|. \tag{6}$$

DEMOSTRACIÓN.—Supongamos que $x \neq 0$. Claramente se tiene

$$\sup_{\substack{f \in E' \\ \|f\| \leq 1}} |\langle f, x\rangle| \leq \|x\|.$$

Por otra parte (corolario I.3) se sabe que existe $f_0 \in E'$ tal que $\|f_0\| = \|x\|$ y $\langle f_0, x\rangle = \|x\|^2$. Se pone $f_1 = \|x\|^{-1} f_0$ de forma que $\|f_1\| = 1$ y $\langle f_1, x\rangle = \|x\|$.

NOTA 3.—Conviene distinguir la fórmula (5), que es una **definición**, de la fórmula (6), que es un **resultado.** En general el «Sup» que aparece en (5) no es un «Max» i.e. no se alcanza (ver un ejemplo en [BT]). Sin embargo, este «Sup» se alcanza si E es un espacio de Banach reflexivo (ver capítulo III); un teorema difícil debido a R. C. James afirma el recíproco: si E en un espacio de Banach tal que para todo $f \in E'$ el «Sup» en (5) se alcanza, entonces E es reflexivo (ver por ejemplo Diestel [1], capítulo I, u Holmes [1]).

I.2. Formas geométricas del teorema de Hahn-Banach: separación de conjuntos convexos

Comencemos con algunos preliminares sobre hiperplanos. En todo lo que sigue E designa un e.v.n.

Definición.—*Un* **hiperplano** *(afín) es un conjunto de la forma*

$$H = \{x \in E;\ f(x) = \alpha\}$$

donde f es una forma lineal (¹) *sobre E,* **no idénticamente nula** *y $\alpha \in \mathbb{R}$. Se dice que* H *es el* **hiperplano de ecuación** $[f = \alpha]$.

Proposición 1.5.—***El hiperplano de ecuación $[f = \alpha]$ es cerrado si y solamente si f es continua.***

DEMOSTRACIÓN.—Es claro que si f es continua entonces H es cerrado. Recíprocamente, supongamos que H es cerrado. El complementario $\complement H$ de H es abierto y no vacío (ya que $f \not\equiv 0$).

(¹) No necesariamente continua (cuando E es de dimensión infinita, siempre existen formas lineales no continuas; ver [BT]).

Sea $x_0 \in \complement H$ y supongamos (para fijar ideas) que $f(x_0) < \alpha$. Sea $r > 0$ tal que $B(x_0, r) \subset \complement H$ donde

$$B(x_0, r) = \{x \in E;\ \|x - x_0\| < r\}.$$

Se tiene

$$(7) \qquad f(x) < \alpha \qquad \forall x \in B(x_0, r).$$

En efecto, supongamos que $f(x_1) > \alpha$ para algún $x_1 \in B(x_0, r)$. El segmento

$$\{x_t = (1 - t)x_0 + tx_1\ ;\ t \in [0, 1]\}$$

está contenido en $B(x_0, r)$ y así $f(x_t) \neq \alpha$ $\forall t \in [0, 1]$; sin embargo, $f(x_t) = \alpha$ para $t = \dfrac{f(x_1) - \alpha}{f(x_1) - f(x_0)}$ lo cual es absurdo, y así (7) queda demostrado. Resulta de (7) que

$$f(x_0 + rz) < \alpha \qquad \forall z \in B(0, 1).$$

Por consiguiente f es continua y $\|f\| < \dfrac{1}{r}(\alpha - f(x_0))$.

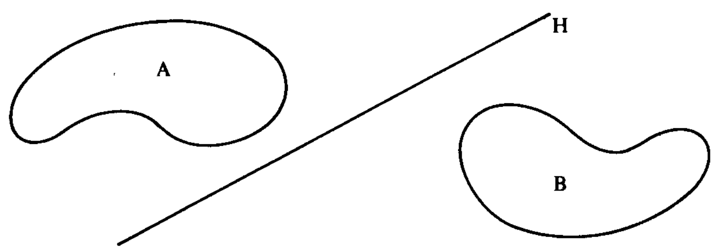
Definición.—Sean $A \subset E$ y $B \subset E$. Se dice que el hiperplano H de ecuación $[f = \alpha]$ **separa A y B en sentido amplio** si se verifica

$$f(x) \leqslant \alpha \quad \forall x \in A \qquad \text{y} \qquad f(x) \geqslant \alpha \quad \forall x \in B.$$

Se dice que H **separa A y B en sentido estricto** si existe $\varepsilon > 0$ tal que

$$f(x) \leqslant \alpha - \varepsilon \quad \forall x \in A \qquad \text{y} \qquad f(x) \geqslant \alpha + \varepsilon, \quad \forall x \in B.$$

Geométricamente la separación significa que A y B se sitúan **«de un lado y de otro de H».**



Recordemos finalmente que un conjunto $A \subset E$ es **convexo** si

$$tx + (1 - t)y \in A \qquad \forall x, y \in A, \qquad \forall t \in [0, 1].$$

• **Teorema I.6 (Hahn-Banach, primera forma geométrica).**—*Sean* **A** $\subset$ **E** *y* **B** $\subset$ **E** *dos conjuntos convexos, no vacíos y disjuntos. Supongamos que* **A** *es* **abierto.** *Entonces existe un hiperplano cerrado que separa* **A** *y* **B** *en sentido amplio.*

La demostración del teorema I.6 se basa en los dos lemas siguientes

Lema 1.2 (Funcional de Minkowski de un convexo).—*Sea* **C** $\subset$ **E** *un convexo abierto con* $0 \in C$. *Para todo* $x \in E$ *se define:*

$$(8) \qquad p(x) = \operatorname{Inf}\{\alpha > 0;\ \alpha^{-1}x \in C\}$$

(se dice que p es el funcional de Minkowski de **C***).*

Entonces p verifica **(1)**, **(2)** *y*

(9) $\qquad$ *existe* **M** *tal que* $0 \leqslant p(x) \leqslant M\|x\| \quad \forall x \in E,$

(10) $\qquad C = \{x \in E;\ p(x) < 1\}.$

Demostración del lema I.2.—Sea $r > 0$ tal que $B(0, r) \subset C$; es claro que

$$p(x) \leqslant \frac{1}{r}\|x\| \qquad \forall x \in E.$$

De donde (9).
La propiedad (1) es evidente.

Demostremos (10). Supongamos primero que $x \in C$; como C es abierto, $(1 + \varepsilon)x \in C$ para $\varepsilon > 0$ suficientemente pequeño. Así $p(x) \leqslant \dfrac{1}{1 + \varepsilon} < 1$. Inversamente, si $p(x) < 1$, existe $0 < \alpha < 1$ tal que $\alpha^{-1}x \in C$ y así $x = \alpha(\alpha^{-1}x) + (1 - \alpha)0 \in C$.

Demostremos (2). Sean $x, y \in E$ y sea $\varepsilon > 0$. De (1) y de (10) se sabe que $\dfrac{x}{p(x) + \varepsilon} \in C$ y $\dfrac{y}{p(y) + \varepsilon} \in C$. Así $\dfrac{tx}{p(x) + \varepsilon} + \dfrac{(1 - t)y}{p(y) + \varepsilon} \in C$ para todo $t \in [0, 1]$. En particular, para $t = \dfrac{p(x) + \varepsilon}{p(x) + p(y) + 2\varepsilon}$ se obtiene $\dfrac{x + y}{p(x) + p(y) + 2\varepsilon} \in C$. Se deduce de ello, gracias a (1) y (10), que $p(x + y) < p(x) + p(y) + 2\varepsilon \quad \forall \varepsilon > 0$. De donde se sigue (2).

Lema I.3.—***Sea*** $C \subset E$ ***un convexo abierto no vacío y sea*** $x_0 \in E$ ***con*** $x_0 \notin C$. ***Entonces existe*** $f \in E'$ ***tal que*** $f(x) < f(x_0)\ \forall x \in C$. ***En particular el hiperplano de ecuación*** $[f = f(x_0)]$ ***separa*** $\{x_0\}$ ***de C en sentido amplio.***

Demostración del lema I.3.—Por traslación se puede siempre suponer que $0 \in C$ e introducir el funcional de Minkowski de C (lema I.2) denotado por p. Se consideran $G = \mathbb{R}x_0$ y la forma lineal g definida en G por

$$g(tx_0) = t, \qquad t \in \mathbb{R}.$$

Es claro que

$$g(x) \leqslant p(x) \qquad \forall x \in G$$

(tomar $x = tx_0$ y distinguir los casos $t > 0$ y $t \leqslant 0$). Gracias al teorema I.1, existe una forma lineal f sobre E, que extiende a g, y tal que

$$f(x) \leqslant p(x) \qquad \forall x \in E.$$

En particular se tiene $f(x_0) = 1$ y f es continua en virtud de (9). Por otra parte se deduce de (10) que $f(x) < 1$ para todo $x \in C$.

Demostración del teorema I.6.—Se pone $C = A - B$ de forma que C es convexo (esto es fácil de comprobar), C es abierto (obsérvese que $C = \bigcup_{y \in B} (A - y)$) y $0 \notin C$ (ya que $A \cap B = \varnothing$). Según el lema I.3 existe $f \in E'$ tal que

$$f(z) < 0 \qquad \forall z \in C$$

es decir

$$f(x) < f(y) \qquad \forall x \in \mathrm{A}, \qquad \forall y \in \mathrm{B}.$$

Se fija $\alpha \in \mathbb{R}$ con

$$\operatorname*{Sup}_{x \in \mathrm{A}} f(x) \leqslant \alpha \leqslant \operatorname*{Inf}_{y \in \mathrm{B}} f(y)$$

y entonces el hiperplano de ecuación $[f = \alpha]$ separa en sentido amplio A y B.

• **Teorema I.7 (Hahn-Banach, segunda forma geométrica).**—*Sean* **A** $\subset$ **E** *y* **B** $\subset$ **E** *dos conjuntos convexos, no vacíos y disjuntos. Supongamos que* **A** *es* **cerrado** *y que* **B** *es* **compacto.** *Entonces existe un hiperplano cerrado que separa* **A** *y* **B** *en sentido estricto.*

DEMOSTRACIÓN.—Para $\varepsilon > 0$ se pone $A_\varepsilon = A + B(0, \varepsilon)$ y $B_\varepsilon = B = B(0, \varepsilon)$ de forma que A_ε y B_ε son convexos, abiertos y no vacíos. Además, para $\varepsilon > 0$ suficientemente pequeño, A_ε y B_ε son disjuntos (en caso contrario, se podrían encontrar sucesiones $\varepsilon_n \to 0$, $x_n \in A$, e $y_n \in B$ tales que $\|x_n - y_n\| < 2\varepsilon_n$; y se podría extraer una subsucesión $y_{ni} \to y \in A \cap B$). Por el teorema 1.6 existe un hiperplano cerrado de ecuación $[f = \alpha]$ que separa A_ε y B_ε en el sentido amplio. Se tiene entonces

$$f(x + \varepsilon z) \leqslant \alpha \leqslant f(y + \varepsilon z) \qquad \forall x \in \mathrm{A}, \quad \forall y \in \mathrm{B}, \quad \forall z \in \mathrm{B}(0, 1).$$

de donde resulta que

$$f(x) + \varepsilon\|f\| \leqslant \alpha \leqslant f(y) - \varepsilon\|f\|, \qquad \forall x \in \mathrm{A}, \qquad \forall y \in \mathrm{B}.$$

Y se concluye que A y B están separados en sentido estricto por el hiperplano $[f = \alpha]$ ya que $\|f\| \neq 0$.

NOTA 4.—Sean $A \subset E$ y $B \subset E$ dos conjuntos convexos, no vacíos y disjuntos. **Sin una hipótesis suplementaria** no se pueden separar A y B en sentido amplio con un hiperplano cerrado. Incluso se puede construir un ejemplo donde A y B son dos convexos cerrados, no vacíos y disjuntos tales que no existe ningún hiperplano cerrado que separe A y B en sentido amplio; ver [BT]. Sin embargo, si E es un espacio de **dimensión finita,** siempre se pueden separar en sentido amplio dos convexos A y B, no vacíos, disjuntos (¡sin hipótesis suplementaria!); ver [BT].

Indiquemos por último un corolario **muy útil** cuando se trata de probar que un subespacio vectorial es **denso.**

• **Corolario I.8.**—*Sea* **F** $\subset$ **E** *un subespacio vectorial tal que* $\overline{\mathbf{F}} \neq \mathbf{E}$. *Entonces existe* $f \in \mathbf{E}'$, $f \neq 0$ *tal que*

$$\langle f, x\rangle = 0 \qquad \forall x \in \mathrm{F}.$$

DEMOSTRACIÓN.—Sea $x_0 \in E$, $x_0 \notin \overline{F}$. Se aplica el teorema I.7 con $A = \overline{F}$ y $B = \{x_0\}$. Existe entonces $f \in E'$, $f \neq 0$ tal que el hiperplano de ecuación $[f = \alpha]$ separa en sentido estricto $\overline{F}$ y $\{x_0\}$. Se tiene

$$\langle f, x\rangle < \alpha < \langle f, x_0\rangle \qquad \forall x \in \mathrm{F}.$$

De donde resulta que $\langle f, x\rangle = 0$ $\forall x \in F$, ya que $\lambda\langle f, x\rangle < \alpha$ para todo $\lambda \in \mathbb{R}$.

• NOTA 5.—Frecuentemente se aplica el corolario I.8 para demostrar que un subespacio vectorial $F \subset E$ es **denso.** Se considera una forma lineal y continua f sobre E tal que $f = 0$ sobre F y **se prueba** que f es idénticamente nula sobre E.

I.3. Introducción a la teoría de las funciones convexas conjugadas

Comencemos con algunos preliminares sobre las funciones semicontinuas inferiormente y sobre las funciones convexas.
En esta sección se consideran funciones φ definidas sobre un conjunto E y con valores en $]-\infty, +\infty]$; así pues φ puede tomar el valor $+\infty$, (pero el valor $-\infty$ queda excluido). Se designa con $D(\varphi)$ el **dominio de** φ es decir, el conjunto

$$D(\varphi) = \{x \in E; \quad \varphi(x) < +\infty\}$$

Notación. El **epigrafo** de φ es el conjunto

$$\text{epi}\, \varphi = \{[x, \lambda] \in E \times \mathbb{R}; \quad \varphi(x) \leqslant \lambda\}\ (^1)$$

Se supone ahora que E es un **espacio topológico.** Recordemos la

Definición.—Una función $\varphi : E \to]-\infty, +\infty]$ se dice **semicontinua inferiormente (s.c.i.)** si para todo $x \in E$ se tiene

$$\liminf_{y \to x} \varphi(y) \geqslant \varphi(x).$$

Utilizaremos algunas propiedades elementales de las funciones s.c.i. (ver Choquet [1] o Dixmier [1]):

(*a*) Si φ es s.c.i., entonces epi φ es cerrado en $E \times \mathbb{R}$; y recíprocamente.

(*b*) Si φ es s.c.i., entonces para todo $\lambda \in \mathbb{R}$ el conjunto $[\varphi \leqslant \lambda] = \{x \in E; \varphi(x) \leqslant \lambda\}$ es cerrado; y recíprocamente.

(*c*) Si φ_1 y φ_2 son s.c.i., entonces $\varphi_1 + \varphi_2$ es s.c.i.

(*d*) Si $(\varphi_i)_{i \in I}$ es una familia de funciones s.c.i. entonces **la envolvente superior** de las (φ_i) es s.c.i., es decir, la función φ definida por

$$\varphi(x) = \operatorname*{Sup}_{i \in I} \varphi_i(x)$$

es s.c.i.

(*e*) Si E es **compacto** y si φ es s.c.i., entonces φ alcanza su cota inferior sobre E.

Se supone ahora que E es un **espacio vectorial.** Recordemos la

Definición.—Una función $\varphi : E \to]-\infty, +\infty]$ se dice **convexa** si

$$\varphi(tx + (1-t)y) \leqslant t\varphi(x) + (1-t)\varphi(y) \qquad \forall x, y \in E, \quad \forall t \in]0, 1[.$$

Utilizaremos algunas propiedades elementales de las funciones convexas:

(*a*) Si φ es una función convexa, entonces epi φ es un conjunto convexo en $E \times \mathbb{R}$; y recíprocamente.

(*b*) Si φ es una función convexa, entonces, para todo $\lambda \in \mathbb{R}$ el conjunto $[\varphi \leqslant \lambda]$ es convexo; pero el recíproco no es cierto.

---

(1) Insistamos en el hecho de que $\mathbb{R} =]-\infty, +\infty[$ y así λ no toma el valor $+\infty$.

(*c*) Si φ_1 y φ_2 son funciones convexas, entonces $\varphi_1 + \varphi_2$ es convexa.

(*d*) Si $(\varphi_i)_{i\in I}$ es una familia de funciones convexas, entonces la envolvente superior de las (φ_i) es convexa.

En todo lo que sigue se supone que E es un **e.v.n.**

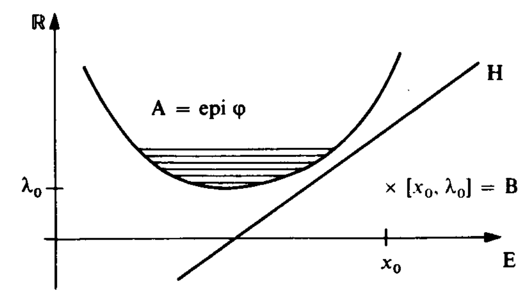
Definición.—Dada una función $\varphi : E \to]-\infty, +\infty]$ tal que $\varphi \not\equiv +\infty$ (i.e. $D(\varphi) \neq \emptyset$) se define la función $\varphi^* : E' \to]-\infty, +\infty]$, **conjugada** de φ por

$$\varphi^*(f) = \operatorname*{Sup}_{x\in E} \{\langle f, x\rangle - \varphi(x)\} \quad (f \in E').$$

Notemos que φ^* es una función convexa s.c.i. sobre E'. En efecto, para cada $x \in E$ **fijo** la aplicación $f \longmapsto \langle f, x\rangle - \varphi(x)$ es convexa y continua, y por tanto s.c.i. Por consiguiente, la envolvente superior de estas funciones (cuando x recorre el conjunto de índices E) es convexa y s.c.i.

Proposición I.9.—*Supongamos que* φ *es convexa, s.c.i. y* $\varphi \not\equiv +\infty$. *Entonces* $\varphi^* \not\equiv +\infty$.

DEMOSTRACIÓN.—Sea $x_0 \in D(\varphi)$ y sea $\lambda_0 < \varphi(x_0)$. Se aplica el teorema I.7 (Hahn-Banach, segunda forma geométrica) en el espacio $E \times \mathbb{R}$ con $A = \text{epi}\ \varphi$ y $B = \{[x_0, \lambda_0]\}$.



Existe entonces un hiperplano cerrado H en $E \times \mathbb{R}$ de ecuación $[\Phi = \alpha]$ que separa estrictamente A y B. Obsérvese que la aplicación $x \in E \mapsto \Phi([x, 0])$ es una forma lineal y continua sobre E y así $\Phi([x, 0]) = \langle f, x\rangle$ para algún $f \in E'$. Tomando $k = \Phi([0, 1])$ se tiene entonces que

$$\Phi([x, \lambda]) = \langle f, x\rangle + k\lambda \quad \text{para todo} \quad [x, \lambda] \in E \times \mathbb{R}.$$

Escribiendo $\Phi > \alpha$ sobre A y $\Phi < \alpha$ sobre B se obtiene:

$$\langle f, x\rangle + k\lambda > \alpha, \qquad \forall [x, \lambda] \in \text{epi}\ \varphi$$

y

$$\langle f, x_0\rangle + k\lambda_0 < \alpha.$$

En particular se tiene

(11) $$\langle f, x\rangle + k\varphi(x) > \alpha \qquad \forall x \in D(\varphi)$$

y entonces

$$\langle f, x_0\rangle + k\varphi(x_0) > \alpha > \langle f, x_0\rangle + k\lambda_0.$$

De donde $k > 0$. Se deduce de (11) que

$$\langle -\frac{1}{k}f, x\rangle - \varphi(x) < -\frac{\alpha}{k} \qquad \forall x \in D(\varphi)$$

y por tanto $\varphi^*\left(-\frac{1}{k}f\right) < +\infty$.

Se define ahora, cuando $\varphi^* \not\equiv +\infty$, la función $\varphi^{**} : E \to]-\infty, +\infty]$ por

$$\varphi^{**}(x) = \operatorname*{Sup}_{f\in E'} \{\langle f, x\rangle - \varphi^*(f)\}.$$

• **Teorema I.10 (Fenchel-Moreau).**—*Supongamos que* φ *es convexa, s.c.i. y* $\varphi \not\equiv +\infty$. *Entonces* $\varphi^{**} = \varphi$.

DEMOSTRACIÓN.—Se procede en dos etapas:

1.ª etapa. Se supone además que $\varphi \geqslant 0$. Ciertamente, $\varphi^{**} \leqslant \varphi$; en efecto, de la definición de φ^* se tiene

$$\langle f, x\rangle \leqslant \varphi(x) + \varphi^*(f) \qquad \forall x \in E, \quad \forall f \in E'.$$

Para probar que $\varphi^{**} = \varphi$ razonemos por reducción al absurdo suponiendo que existe un $x_0 \in E$ tal que

$$\varphi^{**}(x_0) < \varphi(x_0).$$

Posiblemente se tiene $\varphi(x_0) = +\infty$, pero siempre es $\varphi^{**}(x_0) < +\infty$. Se aplica el teorema 1.7 (Hahn-Banach, segunda forma geométrica) en el espacio $E \times \mathbb{R}$ con $A = \text{epi}\,\varphi$ y $B = [x_0, \varphi^{**}(x_0)]$. Existen entonces —como en la demostración de la proposición I.9— $f \in E'$; $k \in \mathbb{R}$ y $\alpha \in \mathbb{R}$ tales que

(12) $$\langle f, x\rangle + k\lambda > \alpha, \quad \forall [x, \lambda] \in \text{epi}\,\varphi$$

(13) $$\langle f, x_0\rangle + k\varphi^{**}(x_0) < \alpha.$$

Resulta de ello que $k \geqslant 0$ (elegir en (12), $x \in D(\varphi)$ y $\lambda = n \to +\infty$). [Aquí no se puede concluir, como en la demostración de la proposición I.9, que $k > 0$; se podría tener posiblemente $k = 0$, lo que correspondería a un hiperplano H «vertical» en $E \times \mathbb{R}$]. Sea $\varepsilon > 0$; como $\varphi \geqslant 0$, en virtud de (12) se tiene:

$$\langle f, x\rangle + (k+\varepsilon)\varphi(x) \geqslant \alpha \qquad \forall x \in D(\varphi).$$

De donde $\varphi^*\left(-\frac{f}{k+\varepsilon}\right) \leqslant -\frac{\alpha}{k+\varepsilon}$; por la definición de $\varphi^{**}(x_0)$, resulta

$$\varphi^{**}(x_0) \geqslant \langle -\frac{f}{k+\varepsilon}, x_0\rangle - \varphi^*\left(-\frac{f}{k+\varepsilon}\right) \geqslant \langle -\frac{f}{k+\varepsilon}, x_0\rangle + \frac{\alpha}{k+\varepsilon}.$$

Por consiguiente

$$\langle f, x_0\rangle + (k+\varepsilon)\varphi^{**}(x_0) \geqslant \alpha \quad \forall \varepsilon > 0,$$

lo cual contradice (13).

2.ª etapa: Caso general. Sea $f_0 \in D(\varphi^*)$($D(\varphi^*) \neq \varnothing$ por la definición I.9). Para situarnos en el caso precedente se introduce la función

$$\bar{\varphi}(x) = \varphi(x) - \langle f_0, x\rangle + \varphi^*(f_0)$$