



UNIVERSITAT DE
BARCELONA

Facultat de Matemàtiques
i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

TEOREMA DE HAHN-BANACH

Autor: Elies Espadaler Eljarrat

Director: Dr. Xavier Massaneda
Realitzat a: Universitat de Barcelona

Barcelona, 14 de gener de 2025

Introducció

El Teorema de Hahn-Banach és un dels pilars fonamentals de l'anàlisi funcional, una branca de les matemàtiques que estudia els espais vectorials amb estructura topològica. Aquest teorema, formulat inicialment per Hans Hahn el 1927 i ampliat posteriorment per Stefan Banach, proporciona una eina essencial per a l'extensió de funcionals lineals definits en subespais a tot l'espai vectorial sense augmentar-ne la norma. Aquesta propietat, aparentment senzilla, té implicacions profundes en l'anàlisi moderna, tant en matemàtiques pures com en àrees aplicades.

El Teorema de Hahn-Banach sorgeix en un moment clau en el desenvolupament de les matemàtiques al segle XX. La seva formulació està profundament arrelada en la necessitat de construir fonaments sòlids per a l'anàlisi funcional, un camp emergent en aquell temps. Hans Hahn, conegut pel seu treball en topologia i anàlisi, va introduir aquest resultat en un article centrat en funcional lineals. Stefan Banach, un dels matemàtics més influents del segle XX, va generalitzar el resultat al context dels espais vectorials normats complexos.

Aquest treball té com a objectiu explorar el Teorema de Hahn-Banach des de diverses perspectives. Primerament, es presenta el context teòric necessari per entendre el resultat, incloent conceptes bàsics sobre espais vectorials normats i funcionals lineals, així com altres resultats rellevants de l'anàlisi funcional com el Teorema de la Categoria de Baire, el Teorema de l'aplicació oberta i el Teorema de Banach-Steinhaus. Posteriorment, es detalla la demostració del teorema, tant en el cas real com en el complex, fent ús del Lema de Zorn com a eina principal.

Finalment, es discuteixen algunes de les seves aplicacions. Les aplicacions del Teorema de Hahn-Banach són àmplies i abasten múltiples àrees de les matemàtiques i disciplines relacionades. En primer lloc, el Teorema de Separació permet dividir conjunts convexos en espais vectorials normats mitjançant hiperplans, una eina clau en optimització i teoria de jocs. En segon lloc, en teoria de conjunts, aquest teorema s'utilitza per demostrar resultats com l'existència de conjunts no mesurables de Lebesgue o per entendre la sorprenent Paradoxa de Banach-Tarski, que desafia la intuïció geomètrica. Finalment, en teoria de jocs, constitueix la base matemàtica del Teorema Minimax de von Neumann, que assegura l'existència d'estratègies òptimes en jocs de suma zero. Aquestes aplicacions il·lustren la versatilitat i la potència d'aquest teorema, unificant conceptes diversos sota un marc teòric comú.

Així doncs, amb aquest treball es vol donar una visió global del Teorema de Hahn-Banach, des de la seva formulació i demostració fins a les seves aplicacions més destacades. L'objectiu és destacar la seva importància dins l'anàlisi funcional i mostrar com connecta diferents àrees de les matemàtiques, convertint-se en una eina essencial tant en la teoria com en les aplicacions pràctiques.

Agraïments

Vull donar les gràcies al meu tutor, el Dr. Xavier Massaneda, per tot el suport, les correccions i l'ajut que m'ha brindat en tot moment a l'hora de realitzar el treball.

Índex

1	Teoria d'operadors	1
1.1	Continuïtat, acotació i norma d'un operador	1
1.2	Espai dels operadors lineals i acotats	4
2	Teoremes clàssics de l'Anàlisi Funcional	7
2.1	Teorema de la Categoria de Baire	7
2.2	Teorema de l'aplicació oberta	8
2.3	Teorema de Banach-Steinhaus	10
3	Teorema de Hahn-Banach	13
3.1	Lema de Zorn	13
3.2	Demostració del Teorema de Hahn-Banach	14
4	Aplicacions del Teorema de Hahn-Banach	19
4.1	Teorema de Separació	19
4.2	Teorema Minimax	23
4.3	Teoria de conjunts i Teorema de Hahn-Banach	32
4.3.1	Reformulació del Teorema de Hahn-Banach en la teoria de conjunts	32
4.3.2	Conjunts no mesurables de Lebesgue	33
4.3.3	Paradoxa de Banach-Tarsky	38

Capítol 1

Teoria d'operadors

En aquest capítol explicarem les propietats fonamentals dels operadors lineals. Veurem les propietats de continuïtat i d'acotació d'un operador lineal i la relació entre les dues. A partir d'aquestes propietats, definirem la norma d'un operador lineal i acotat. Per últim, considerarem el conjunt de tots els operadors lineals acotats entre espais normats, i veurem que formen un espai vectorial normat. Totes les referències d'aquest capítol són del llibre [9, Capítols 5.1, 5.2].

Passem, primer de tot, a definir què és un operador lineal:

Definició 1.0.1. *Siguin X i Y dos espais vectorials reals (o complexos). Una aplicació $T : X \rightarrow Y$ és un operador lineal si el domini de T , D_T , és un subespai vectorial de X , i a més*

$$T(\alpha x + \beta y) = \alpha Tx + \beta Ty$$

per tot $x, y \in D_T$ i per tots escalars α i β .

S'acostuma a escriure Tx en comptes de $T(x)$. Considerarem, a menys que s'especifiqui, que el domini D_T és tot X . En el cas en que X sigui un espai vectorial sobre $\mathbb{R}$ (respectivament $\mathbb{C}$), un operador lineal de X a $\mathbb{R}$ (respectivament $\mathbb{C}$) s'anomena *funcional lineal*.

1.1 Continuïtat, acotació i norma d'un operador

Començem amb la continuïtat d'un operador lineal. Recordem que, donat X un espai mètric i $T : X \rightarrow Y$ un operador, T és contínu si per tota successió $\{x_n\}_{n=1}^\infty \subseteq X$ tal que $\lim_n \|x_n - x_0\| = 0$ se satisfà que $\lim_n \|Tx_n - Tx_0\| = 0$. Veiem ara que, pel fet de ser lineal, si l'operador és contínu en un únic punt aleshores ho és en tots els punts, és a dir, per veure que un operador lineal sigui contínu, només ens caldrà veure que ho és en un punt.

Teorema 1.1.1. *Siguin X i Y espais lineals normats. Aleshores un operador lineal $T : X \rightarrow Y$ és continu en tots els punts si ho és en un únic punt.*

Demostració. Suposem que T és contínua en x_0 . Sigui x un punt qualsevol de X i sigui $\{x_n\}_{n=1}^{\infty}$ una successió en X que convergeix a x . Considerem ara la successió desplaçada $\{x_n - x + x_0\}_{n=1}^{\infty}$. Aquesta successió convergeix a x_0 i, a més, com T és contínua en x_0 , tenim que $\{T(x_n - x + x_0)\}_{n=1}^{\infty}$ convergeix a Tx_0 . Com T és lineal, se satisfà que

$$T(x_n - x + x_0) = Tx_n - Tx + Tx_0.$$

Per tant, $\{T(x_n)\}_{n=1}^{\infty}$ convergeix a Tx . Per l'arbitrarietat de x , T és contínua a tot $x \in X$. $\square$

Seguim ara definint l'acotació d'un operador lineal.

Definició 1.1.1. *Siguin X i Y espais lineals normats i $T : X \rightarrow Y$ un operador lineal. Diem que T està acotat si existeix $M > 0$ tal que*

$$\|Tx\|_Y \leq M\|x\|_X$$

per tot $x \in X$.

Notem ara que, donat un operador lineal acotat tenim que, per tot $x \in X$,

$$\frac{\|Tx\|_Y}{\|x\|_X} \leq M.$$

Això porta a la següent definició:

Definició 1.1.2. *Siguin X i Y espais lineals normats i $T : X \rightarrow Y$ un operador lineal acotat. Es defineix la norma de T com*

$$\|T\| = \inf\{M > 0 : \|Tx\|_Y \leq M\|x\|_X, \forall x \in X\} = \sup_{x \neq 0} \frac{\|Tx\|_Y}{\|x\|_X} \quad (1.1.1)$$

Per a operadors lineals, la propietat de continuïtat i la propietat d'acotació són equivalents:

Teorema 1.1.2. *Siguin X i Y espais lineals normats. Un operador lineal $T : X \rightarrow Y$ és continu en X si, i només si, T està acotat en X .*

Demostració. Demostrem les dues implicacions:

T continu $\Rightarrow T$ acotat: Suposem que T és continu en X , en particular ho és en el 0; per tant, existeix $\delta > 0$ tal que $\|x_0\|_X \leq \delta$ implica que

$$\|Tx_0\|_Y = \|Tx_0 - T0\|_Y \leq 1.$$

Sigui $x \in X$ arbitrari i $x_0 = \delta \frac{x}{\|x\|_X}$. Aleshores tenim que $\|x_0\| = \delta$ i, per la linealitat de T ,

$$1 \geq \|Tx_0\|_Y = \frac{\delta}{\|x\|_X} \|Tx\|_Y,$$

Per tant,

$$\|Tx\|_Y \leq \frac{1}{\delta} \|x\|_X$$

per tot $x \in X$. En altres paraules, T està acotat, i la seva norma és, com a molt $1/\delta$.

T acotat $\Rightarrow T$ continu: Suposem ara que T està acotat en X . Pel Teorema 1.1.1, és suficient veure que T és contínua en el 0.

Segui $\{x_n\}_{n=1}^\infty$ una successió en X que convergeix a 0. Com T està acotat, tenim que

$$\|Tx_n\|_Y \leq \|T\| \cdot \|x_n\|_X$$

per tot n . Per tant,

$$\|Tx_n\|_Y \rightarrow 0 \text{ quan } n \rightarrow \infty,$$

és dir, que T és continu en el 0. □

Fem ara un exemple del càlcul de la norma d'un operador lineal acotat:

Exemple: Considerem $X = \mathcal{C}[0, 1] = Y$. Aquest és un espai normat amb la norma

$$\|f\|_X = \|f\|_Y = \max_{t \in [0, 1]} |f(t)|, \text{ per tot } f \in \mathcal{C}[0, 1].$$

Considerem l'operador "multiplicar per t ":

$$\begin{aligned} T : X &\rightarrow Y \\ f(t) &\mapsto tf(t) \end{aligned}$$

La norma de l'operador T és, en aquest cas,

$$\|T\| = \sup_{f \neq 0} \frac{\|Tf\|_Y}{\|f\|_X} = \sup_{f \neq 0} \frac{\max_{t \in [0, 1]} |tf(t)|}{\max_{t \in [0, 1]} |f(t)|}$$

Notem que, com $t \in [0, 1]$,

$$|tf(t)| \leq |f(t)| \leq \max_{t \in [0, 1]} |f(t)| = \|f\|_X$$

de manera que

$$\|Tf\|_X = \max_{t \in [0, 1]} |tf(t)| \leq \|f\|_X$$

i per tant

$$\frac{\|Tf\|_X}{\|f\|_X} \leq 1$$

és a dir,

$$\|T\| = \sup_{f \neq 0} \frac{\|Tf\|_Y}{\|f\|_X} \leq 1.$$

Veiem ara que $\|T\| = 1$. Prenem per exemple $f(t) = t$ (podríem considerar qualsevol funció f amb $\max_{t \in [0, 1]} |f(t)| = |f(1)|$). Aleshores tenim $\|Tf\|_X = \|f\|_X$, i per tant

$$\|T\| = \sup_{f \neq 0} \frac{\|Tf\|_Y}{\|f\|_X} \geq 1.$$

Tot plegat; $\|T\| = 1$.

De manera anàloga, es pot veure que tota $g \in \mathcal{C}[0, 1]$ defineix un operador $T_g : \mathcal{C}[0, 1] \rightarrow \mathcal{C}[0, 1]$ donat per $T_g(f) = f \cdot g$, i que té $\|T_g\| = \|g\|_{\mathcal{C}[0, 1]}$

1.2 Espai dels operadors lineals i acotats

Siguin X i Y espais normats. Considerem $\mathcal{B}(X, Y)$ el conjunt de tots els operadors lineals acotats $T : X \rightarrow Y$, és a dir,

$$\mathcal{B}(X, Y) = \{T : X \rightarrow Y \mid T \text{ operador lineal i acotat}\}$$

En cas que $X = Y$, escrivim $\mathcal{B}(X) = \mathcal{B}(X, X)$. La norma de l'operador T , vist com a operador de X a Y , es denota per $\|T\|_{\mathcal{B}(X, Y)}$. Cal fer atenció en que aquesta norma depèn dels espais X i Y .

Teorema 1.2.1. *Siguin X i Y espais vectorials normats. El conjunt $\mathcal{B}(X, Y)$ amb la norma $\|\cdot\|_{\mathcal{B}(X, Y)}$ forma un espai vectorial normat.*

Demostració. Comprovem que satisfà les propietats d'espai vectorial normat:

- Donat $T \in \mathcal{B}(X, Y)$, tenim que $\|T\| = \sup_{x \neq 0} \frac{\|Tx\|_Y}{\|x\|_X} \geq 0$ per tot $x \in X$, ja que $\|Tx\|_Y \geq 0$ i $\|x\|_X \geq 0$. A més, $\|T\| = \sup_{x \neq 0} \frac{\|Tx\|_Y}{\|x\|_X} = 0$ per tot $x \in X$ si, i només si, $T = 0$.
- Donat $T \in \mathcal{B}(X, Y)$ i donat $\alpha \in \mathbb{R}$, tenim que

$$\|\alpha T\| = \sup_{x \neq 0} \frac{\|(\alpha T)x\|_Y}{\|x\|_X} = \sup_{x \neq 0} \frac{|\alpha| \|Tx\|_Y}{\|x\|_X} = |\alpha| \|T\|,$$

ja que $\|\alpha Tx\|_Y = |\alpha| \|Tx\|_Y$

- Donats $T, F \in \mathcal{B}(X, Y)$ volem veure la desigualtat triangular. Tenim que

$$\|T + F\| = \sup_{x \neq 0} \frac{\|(T + F)x\|_Y}{\|x\|_X} \leq \sup_{x \neq 0} \frac{\|Tx\|_Y + \|Fx\|_Y}{\|x\|_X} = \|T\| + \|F\|.$$

□

Recordem que un espai de Banach és un espai vectorial normat i complet. Un espai és complet si tota successió de Cauchy és convergent.

Teorema 1.2.2. *Siguin X i Y espais lineals normats. Si Y és un espai de Banach aleshores el conjunt $\mathcal{B}(X, Y)$ és un espai de Banach.*

Demostració. Pel Teorema 1.2.1, només cal veure que $\mathcal{B}(X, Y)$ és complet. Sigui $\{T_n\}_{n=1}^{\infty}$ una successió de Cauchy de $\mathcal{B}(X, Y)$. Aleshores, per tot $\epsilon > 0$, existeix n_0 tal que per tot $n, m \geq n_0$,

$$\|T_n - T_m\|_{\mathcal{B}(X, Y)} < \epsilon.$$

A més, com tota successió de Cauchy està acotada, existeix $M > 0$ tal que

$$\|T_n\|_{\mathcal{B}(X, Y)} \leq M$$

per tot n .

Cal veure que aquesta successió és convergent. Considerem la successió $\{T_n x\}_{n=1}^{\infty}$, que és de Cauchy per cada $x \in X$. Com Y és complet, tenim que $\{T_n x\}$ convergeix per cada $x \in X$. Denotem per $Tx = \lim_{n \rightarrow \infty} T_n x$ el límit puntual d'aquesta successió. Observem que l'operador T és lineal i satisfà que

$$\|Tx\|_Y \leq M\|x\|_X$$

per tot $x \in X$, és a dir, T és acotat. Per tant, $T \in \mathcal{B}(X, Y)$. Veiem ara que $\{T_n\}_n \rightarrow T$:

Per $n > m \geq n_0$, tenim

$$\|T_n - T_m\|_{\mathcal{B}(X, Y)} < \epsilon$$

per tant,

$$\|T_n x - T_m x\|_{\mathcal{B}(X, Y)} < \epsilon \|x\|_X$$

per tot $x \in X$. Fixant m , tenim que

$$\lim_{n \rightarrow \infty} \|T_n x - T_m x\|_Y = \|Tx - T_m x\|_Y,$$

i, per la desigualtat anterior,

$$\|Tx - T_m x\|_Y \leq \epsilon \|x\|_X$$

per tot $x \in X$. És a dir,

$$\|T - T_m\|_Y \leq \epsilon$$

per tot $m > n_0$, que demostra que $\lim_{n \rightarrow \infty} T_n = T$.

□

Capítol 2

Teoremes clàssics de l'Anàlisi Funcional

En aquest capítol presentarem dos dels teoremes més clàssics de l'anàlisi funcional: El teorema de l'aplicació oberta i el teorema de Banach-Steinhaus (o també anomenat el principi de l'acotació uniforme). El teorema de Hahn-Banach, que explicarem més endavant al Capítol 3, és també un dels resultats més importants en aquest àmbit.

Abans de presentar aquests teoremes, ens cal parlar del Teorema de la Categoria de Baire, un resultat necessari per la demostració d'aquests resultats.

2.1 Teorema de la Categoria de Baire

Totes les referències d'aquesta secció són del llibre [8, Secció 2.2].

Recordem que, donat X un espai mètric complet, un conjunt $A \subseteq X$ és *dens* en X si per tot obert U de X tenim que $A \cap U \neq \emptyset$. Equivalentment, $A \subseteq X$ és dens en X si, i només si, $\overline{A} = X$.

Definició 2.1.1. *Sigui X un espai mètric i sigui $A \subseteq X$ un subconjunt. Diem que el conjunt A és enlloc dens si $X \setminus \overline{A}$ és dens en X . Equivalentment, A és un conjunt enlloc dens en X si, i només si, $\text{int}(\overline{A}) = \emptyset$.*

Teorema 2.1.1. (Teorema de la Categoria de Baire): *Si X és un espai mètric complet, aleshores la intersecció de tota col·lecció numerable de subconjunts densos oberts de X és densa en X .*

Equivalentment, si X és un espai mètric complet, tota unió numerable de conjunts tancats amb interiors buits té interior buit.

Demostració. Sigui X un espai mètric complet no buit. Suposem que $V_1, V_2, V_3, \dots$ són subconjunts oberts densos de X . Sigui B_0 un conjunt obert arbitrari no buit a X . Volem veure que

$$B_0 \cap \bigcap_{n \in \mathbb{N}} V_n \neq \emptyset,$$

i, per tant, que $\bigcap_{n \in \mathbb{N}} V_n$ és dens.

Per això, farem un procés iteratiu per B_0 . Com cada V_j és dens en X , tenim que $V_j \cap B_0 \neq \emptyset$ per tot j . En particular, tenim que $V_1 \cap B_0 \neq \emptyset$ i és obert, ja que és intersecció numerable de dos oberts. Per tant, podem trobar un obert $B_1 \neq \emptyset$ tal que $\overline{B_1} \subseteq V_1 \cap B_0$. Notem que ara podem repetir el procés de la següent manera: com V_2 és dens en X , tenim que $V_2 \cap B_1 \neq \emptyset$. Per tant, existeix un obert $B_2 \neq \emptyset$ tal que $\overline{B_2} \subseteq V_2 \cap B_1$.

Per tant, iterant aquest procés, veiem que podem trobar una successió d'oberts $B_n \neq \emptyset$ tal que

$$\overline{B_n} \subseteq V_n \cap B_{n-1}.$$

Podem suposar que B_n és una bola oberta de la forma

$$B_{\epsilon_n} := \{x \in X : d(z_n, x) < \epsilon_n\}, \text{ amb } \lim_{n \rightarrow \infty} \epsilon_n = 0,$$

ja que tot obert conté una bola oberta.

Considerem ara

$$K := \bigcap_{j=1}^{\infty} \overline{B_j},$$

i demostrem, per contradicció, que $K \neq \emptyset$ i que

$$K \subseteq B_0 \cap \bigcap_{n \in \mathbb{N}} V_n.$$

El fet que $K \neq \emptyset$ surt de comprovar que, com $\lim_{n \rightarrow \infty} \epsilon_n = 0$, la successió $\{z_n\}_{n \in \mathbb{N}}$ és de Cauchy, i per tant, convergent a un cert $z_0 \in K$, ja que X és complet. Comprovem doncs que és de Cauchy: sigui $n > m$, aleshores

$$z_n \in B_n \subseteq B_{n-1} \subseteq \dots \subseteq B_m,$$

i per tant, $|z_n - z_m| < \epsilon_m$.

Finalment, per construcció tenim que $K \subseteq B_0$ i $K \subseteq V_j$ per cada j . La inclusió doncs és directa.

□

2.2 Teorema de l'aplicació oberta

En aquesta secció seguirem el llibre [9, Capítol 6.3].

Siguin X un espai vectorial, $\alpha > 0$ un escalar i $A \subseteq X$ un subespai. Denotem $\alpha A = \{\alpha x | x \in A\}$. Notem que $\alpha B_\beta(x) = B_{\alpha\beta}(x)$ per $\beta > 0$, on $B_r(x)$ denota la bola oberta de radi r i centre x .

Teorema 2.2.1. (Teorema de l'aplicació oberta): *Siguin X i Y espais de Banach i sigui $T \in \mathcal{B}(X, Y)$. Si T és exhaustiva, aleshores $T(U)$ és un obert de Y per tot U obert de X .*

Demostració. Separem la demostració en 3 passos:

- (i) $\exists \epsilon > 0$ tal que $B_\epsilon(0) \subseteq \overline{T(B_{\frac{1}{2}}(0))}$,
- (ii) Per $\epsilon > 0$ trobat a l'apartat (i), $B_\epsilon(0) \subseteq T(B_1(0))$,
- (iii) $T(U)$ és un obert de Y per tot U obert de X .

Demostrem cada apartat per separat:

- (i) Com T és exhaustiva se satisfà que

$$Y = T(X) = T\left(\bigcup_{n=1}^{\infty} B_{\frac{n}{2}}(0)\right) = \bigcup_{n=1}^{\infty} T(B_{\frac{n}{2}}(0))$$

El teorema de la categoria de Baire implica que existeix $N \in \mathbb{N}$ tal que $T(B_{\frac{N}{2}}(0))$ no és un conjunt enllloc dens en Y . Usant la definició, tenim que

$$\text{int}\left(\overline{T\left(B_{\frac{N}{2}}(0)\right)}\right) \neq \emptyset,$$

i, per tant, existeixen $y_0 \in Y$ i $r > 0$ tals que

$$B_r(y_0) \subseteq \overline{T\left(B_{\frac{N}{2}}(0)\right)}.$$

Prenem $\epsilon = \frac{r}{2N}$. Donat $y \in B_r(y_0)$, tenim que $y = Tx$ amb $\|x\| \leq N/2$. Aleshores $\frac{y}{N} = T\left(\frac{x}{N}\right)$ per $\|x\| \leq 1/2$, que implica que $\frac{y}{N} \in B_{\frac{r}{N}}\left(\frac{y_0}{N}\right) \subseteq \overline{T\left(B_{\frac{1}{2}}(0)\right)}$ per tant,

$$\frac{y_0}{N} + y \in \overline{T\left(B_{\frac{1}{2}}(0)\right)}, \text{ per tot } y \in B_{2\epsilon}(0).$$

Considerem ara qualsevol $y \in B_\epsilon(0)$. Com tant $\frac{y_0}{N}$ com $-\frac{y_0}{N}$ estan en $\overline{T\left(B_{\frac{1}{2}}(0)\right)}$ tenim que:

$$y = \frac{-y_0}{2N} + \left(\frac{y_0}{2N} + y\right) = \frac{-y_0}{2N} + \frac{1}{2}\left(\frac{y_0}{N} + 2y\right) \in \overline{T\left(B_{\frac{1}{4}}(0)\right)} + \overline{T\left(B_{\frac{1}{4}}(0)\right)} \subseteq \overline{T\left(B_{\frac{1}{2}}(0)\right)}$$

- (ii) Sigui $y \in B_\epsilon(0)$ qualsevol. Per l'apartat (i), podem escollir $x_1 \in B_{\frac{1}{2}}(0)$ tal que Tx_1 i y siguin suficientment propers, és a dir, que satisfacin

$$\|y - Tx_1\|_Y \leq \frac{\epsilon}{2}.$$

Això implica que

$$y - Tx_1 \in B_{\frac{\epsilon}{2}}(0) = \frac{1}{2}B_\epsilon(0) \subseteq \frac{1}{2}\overline{T\left(B_{\frac{1}{2}}(0)\right)} = \overline{T\left(B_{\frac{1}{4}}(0)\right)}.$$

Ara podem escollir $x_2 \in B_{\frac{1}{4}}(0)$ tal que Tx_2 i $y - Tx_1$ siguin suficientment propers, és a dir, que satisfacin

$$\|(y - Tx_1) - Tx_2\|_Y \leq \frac{\epsilon}{4}.$$

Iterant aquest procés, podem crear una successió $x_n \in B_{2^{-n}}(0)$ tal que

$$\|y - \sum_{i=1}^n Tx_i\|_Y \leq \frac{\epsilon}{2^n}. \quad (2.2.1)$$

Notem que $\|x_n\| \leq 1/2^n$, i per tant, posant $x = \sum_{i=1}^{\infty} x_i$ tenim que

$$\|x\|_X \leq \sum_{i=1}^{\infty} \|x_i\|_X < \sum_{i=1}^{\infty} 2^{-i} = 1,$$

és a dir, $x \in B_1(0)$. Finalment, tenim que, per (2.2.1), $y = \sum_{i=1}^{\infty} Tx_i$, i per continuïtat de T , tenim que $y = T(\sum_{i=1}^{\infty} x_i) = Tx$.

- (iii) Sigui U un obert de X . Considerem un element $Tx \in T(U)$. Com $x \in U$ i U és obert, existeix $\delta > 0$ tal que $B_{\delta}(x) \subseteq U$.

És suficient veure que $B_{\delta\epsilon}(Tx) \subseteq T(U)$. Sigui $y \in B_{\delta\epsilon}(Tx)$. Podem escriure $y = Tx + y_1$ per algun $y_1 \in B_{\delta\epsilon}(0)$. Considerem ara $y_2 = \frac{y_1}{\delta}$. Tenim que

$$\|y_2\| = \left\| \frac{y_1}{\delta} \right\| = \frac{1}{\delta} \|y_1\| \leq \frac{1}{\delta} \delta\epsilon = \epsilon.$$

Aleshores $y_2 = \frac{y_1}{\delta} \in B_{\epsilon}(0) \subseteq T(B_1(0))$. Escrivim $y_2 = Tx_2$ per algun $x_2 \in B_1(0)$. Aleshores

$$y = Tx + y_1 = Tx + \delta Tx_2 = T(x + \delta x_2).$$

Veiem doncs que $\|x + \delta x_2 - x\| = \delta \|x_2\| \leq \delta$. Per tant, $x + \delta x_2 \in B_{\delta}(x) \subseteq U$. Per tant, $y \in T(U)$ com volíem.

□

2.3 Teorema de Banach-Steinhaus

Passem ara al Teorema de Banach-Steinhaus, també conegut com Principi d'acotació uniforme. Aquest resultat afirma que, per a una família d'operadors lineals acotats amb domini un espai de Banach, l'acotació puntual és equivalent a l'acotació uniforme segons una norma d'un operador. Totes les referències d'aquesta secció són del llibre [9, Capítol 6.3].

Teorema 2.3.1. (Teorema de Banach-Steinhaus): *Siguin X un espai de Banach i Y un espai vectorial normat. Si $\mathcal{A} \subseteq \mathcal{B}(X, Y)$ satisfà*

$$\sup\{\|Tx\|_Y \mid T \in \mathcal{A}\} < \infty$$

per cada $x \in X$, aleshores

$$\sup\{\|T\|_{\mathcal{B}(X, Y)} \mid T \in \mathcal{A}\} < \infty$$

Demostració. Considerem la família de conjunts

$$E_n = \{x \in X : \|Tx\|_Y \leq n \text{ per tot } T \in \mathcal{A}\} = \bigcap_{T \in \mathcal{A}} \{x \in X : \|Tx\|_Y \leq n\}$$

Tots els conjunts E_n són conjunts tancats ja que $\{x \in X : \|Tx\|_Y \leq n\}$ és tancat per la completitud de X i Y , i la intersecció de tancats és també tancat. Per hipòtesi, $\bigcup_{n=1}^{\infty} E_n = X$. Usant el teorema de categoria de Baire tenim que existeix $N \in \mathbb{N}$ tal que $\text{int}(E_N)$ és no buit. Per definició d'interior, existeixen x_0 i $r > 0$ tal que $B_r(x_0) \subseteq E_N$, és a dir, per tot $x \in B_r(x_0)$ i $T \in \mathcal{A}$ se satisfà que $\|Tx\|_Y \leq N$.

Volem veure que existeix un nombre K tal que $\|Tx\|_Y \leq K$ per tot $x \in X$ de norma 1 i per tot $T \in \mathcal{A}$. Prenem arbitràriament un element $x \in X$ amb $\|x\|_X = 1$ i considerem l'element $y = \frac{r}{2}x + x_0$. Tenim doncs que $y \in B_r(x_0)$ i per tant $\|Ty\|_Y \leq N$. A més,

$$\frac{r}{2}\|Tx\|_Y - \|Tx_0\|_Y \leq \left\| \frac{r}{2}Tx + Tx_0 \right\|_Y = \|Ty\|_Y \leq N,$$

o equivalentment,

$$\|Tx\|_Y \leq \frac{2}{r}(N + \|Tx_0\|_Y),$$

que és independent de x . Definint $K = \frac{2}{r}(N + \|Tx_0\|_Y)$, ja hem acabat. $\square$

Capítol 3

Teorema de Hahn-Banach

En aquest capítol presentarem el Teorema de Hahn-Banach, un dels resultats fonamentals de l'anàlisi funcional i l'objecte d'estudi d'aquest treball. Aquest resultat, publicat l'any 1927 per Hans Hahn i ampliat posteriorment per Stefan Banach, afirma que podem estendre un funcional lineal acotat definit en un subespai vectorial a tot l'espai vectorial, de tal manera que la norma de l'extensió no augmenti. La importància d'aquest teorema radica en el fet que la norma de l'extensió no augmenti. Totes les referències d'aquest capítol són del llibre [9, Capítol 6.3].

Definició 3.0.1. *Sigui X un espai vectorial normat. La col·lecció de tots els funcionals lineals continus en X s'anomena espai dual de X , i és denota X^**

Notem que $X^* = \mathcal{B}(X, \mathbb{R})$ (respectivament $\mathcal{B}(X, \mathbb{C})$) si X és un espai vectorial sobre $\mathbb{C}$).

Sigui X un espai vectorial normat i sigui M un subespai propi de X . Sigui λ un funcional lineal en M que pren valors a $\mathbb{R}$ o $\mathbb{C}$. Aleshores un funcional lineal Λ en X s'anomena *extensió* de λ si $\lambda(x) = \Lambda(x)$ per tot $x \in M$.

Teorema 3.0.1. (Teorema de Hahn-Banach): *Siguin X un espai vectorial normat real o complex, i M un subespai. Si $\lambda \in M^*$, aleshores existeix $\Lambda \in X^*$ tal que $\Lambda = \lambda$ en M i, a més,*

$$\|\Lambda\| = \|\lambda\|.$$

Observem que les dues normes que apareixen en el teorema són diferents. La igualtat, escrita amb més presició és:

$$\|\Lambda\|_{\mathcal{B}(X, \mathbb{R})} = \|\lambda\|_{\mathcal{B}(M, \mathbb{R})}.$$

Per demostrar aquest teorema utilitzarem el Lema de Zorn, que ara tot seguit recordem.

3.1 Lema de Zorn

Un conjunt *parcialment ordenat* és un conjunt no buit S amb una relació " $\preceq$ " que satisfà:

- Propietat reflexiva: $x \preceq x$ per cada $x \in S$.
- Propietat transitiva: Si $x \preceq y$ i $y \preceq z$, aleshores $x \preceq z$ per tot $x, y, z \in S$.

Si per tot x i y en un conjunt parcialment ordenat S , o bé $x \preceq y$ o bé $y \preceq x$ aleshores deim que S es un conjunt *totalment ordenat*.

Considerem T un subconjunt d'un conjunt parcialment ordenat S . Diem que un element $x \in S$ és una *cota superior* de T si $y \preceq x$ per tot $y \in T$. Diem que $x \in S$ és un *element maximal* si $x \preceq y$ implica $y \preceq x$.

Lema 3.1.1. (Lema de Zorn): *Tot conjunt parcialment ordenat S en el qual tots els seus subconjunts totalment ordenats tenen una cota superior, ha de contenir un element maximal.*

3.2 Demostració del Teorema de Hahn-Banach

Ara ja tenim totes les eines necessàries per demostrar el Teorema de Hahn-Banach. Separarem la demostració del cas real i del cas complex. Començem pel cas real:

Demostració del Teorema de Hahn-Banach (Cas Real): Considerem la funció real $p(x) = \|\lambda\| \cdot \|x\|$ definida a tot $x \in X$. Observem que, per tot $x, y \in X$,

$$p(x + y) \leq p(x) + p(y)$$

ja que $p(x + y) = \|\lambda\| \cdot \|x + y\| \leq \|\lambda\| \cdot (\|x\| + \|y\|) = p(x) + p(y)$, per la desigualtat triangular. Observem també que per $\alpha \in \mathbb{R}$ i $x \in X$

$$p(\alpha x) \leq \alpha p(x)$$

ja que $p(\alpha x) = \|\lambda\| \cdot \|\alpha x\| = |\alpha| \cdot \|\lambda\| \cdot \|x\|$.

Observem també que, per a tot $x \in X$,

$$|\lambda(x)| \leq p(x)$$

ja que, $\|\lambda\| = \sup_{x \neq 0} \frac{|\lambda(x)|}{\|x\|}$, que implica que, $\|\lambda\| \geq \frac{|\lambda(x)|}{\|x\|}$, i per tant, $|\lambda(x)| \leq \|\lambda\| \cdot \|x\| = p(x)$.

Considerem $z \in X \setminus M$ fix. El primer pas de la prova és veure com λ es pot estendre al subespai generat per M i z . Aleshores per tot $x, y \in M$ tenim que

$$\begin{aligned} \lambda(x) - \lambda(y) &= \lambda(x - y) \\ &\leq p(x - y) \\ &= p((x + z) + (-z - y)) \\ &\leq p(x + z) + p(-z - y). \end{aligned}$$

Per tant,

$$-p(-z - y) - \lambda(y) \leq p(x + z) - \lambda(x)$$

per tot $x, y \in M$. A més, com $y \in M$,

$$s = \sup_{y \in M} \{-p(-z - y) - \lambda(y)\} \leq p(x + z) - \lambda(x),$$

i per tant,

$$s \leq \inf_{x \in M} \{p(x + z) - \lambda(x)\}.$$

Definim ara el subespai generat per M i z , M_z de X com

$$M_z := \{x + \alpha z : x \in M, \alpha \in \mathbb{R}\}.$$

Observem que la representació $w = x + \alpha z$ és única per $w \in M_z$. Definim

$$\tilde{\lambda}(w) = \lambda(x) + \alpha s$$

en M_z . Comprovem ara que $\tilde{\lambda}$ és una extensió de λ a M_z . Tenim que $\tilde{\lambda}$ és lineal i $\tilde{\lambda}(x) = \lambda(x)$ per cada $x \in M$. Per tant, acabem d'extendre λ des de M a un subespai més gran M_z de X .

La idea a partir d'aquí és anar iterant aquest procés fins trobar una extensió de λ a tot X . Separem dos casos:

- Si $M_z = X$: Com $\frac{x}{\alpha} \in M$ per tot $\alpha \neq 0$, tenim que

$$-p(-z - \frac{x}{\alpha}) - \lambda(\frac{x}{\alpha}) \leq s \leq p(\frac{x}{\alpha} + z) - \lambda(\frac{x}{\alpha}), \quad \alpha \neq 0.$$

D'aquesta desigualtat podem deduir que $|\tilde{\lambda}(w)| \leq p(w)$ per cada $w \in M_z$ (usant la primera desigualtat si $\alpha < 0$ i la segona si $\alpha > 0$).

- Si $M_z \neq X$: Recordem que volem estendre λ de M a X . Podríem repetir el procés anterior, extenent λ de M_z a un subespai de X més gran. Tot i així, no tenim la certesa de poder estendre-ho a tot X amb aquest mètode. Per comprovar que efectivament és així, utilitzarem el Lema de Zorn.

Considerem el conjunt

$$S = \left\{ (M', \lambda') : \begin{array}{l} M' \text{ és un subespai de } X \text{ tal que } M \subseteq M' \subseteq X \\ \lambda' \text{ és una extensió de } \lambda \text{ de } M \text{ en } M' \text{ tal que } \lambda' \leq p \text{ en } M' \end{array} \right\}$$

Definim ara la relació " $\preceq$ " en S com:

$$(M', \lambda') \preceq (M'', \lambda'') \text{ si } M' \text{ és un subespai propi de } M'' \text{ i } \lambda' = \lambda'' \text{ en } M'.$$

Això defineix un ordre parcial en S .

Sigui ara $P = \{(M_a, \lambda_a)\}_{a \in A}$ un subconjunt de S totalment ordenat. Tenim doncs que el parell $(\bigcup_{a \in A} M_a, \tilde{\lambda})$, on $\tilde{\lambda}(x) = \lambda_a(x)$ per tot $x \in M_a$, és un element de S , i de fet, és una cota superior per P . Com P és un subespai totalment ordenat arbitrari de S , pel Lema de Zorn tenim que S té un element maximal, que podem anomenar $(M_\infty, \lambda_\infty)$. Notem que λ_∞ és una extensió

de λ de M a M_∞ , que satisfà que $\lambda_\infty \leq p$ en M_∞ . Per tant, si veiem que $M_\infty = X$, haurem provat el que volíem.

Suposem doncs que $M_\infty \neq X$ i arribem a contradicció :

Suposem que $M_\infty \neq X$. Podem doncs, aplicant el mètode anterior per passar de M a M_z , crear un element $(M_{\infty z}, \lambda_{\infty z})$ de S amb $z \in X \setminus M_\infty$. Aquest element satisfà que

$$(M_\infty, \lambda_\infty) \preceq (M_{\infty z}, \lambda_{\infty z})$$

Però, com $(M_\infty, \lambda_\infty)$ és maximal en S , aleshores s'ha de satisfer que $M_\infty = M_{\infty z}$, contradient la definició de " $\preceq$ ". Per tant, $M_\infty = X$.

Posem ara $\Lambda = \lambda_\infty$, que és una extensió de λ a tot X que satisfà

$$|\Lambda(x)| \leq p(x) = \|\lambda\| \cdot \|x\|$$

per tot $x \in X$, per tant, $\|\Lambda\| \leq \|\lambda\|$.

Només cal veure que aquesta desigualtat és realment una igualtat. Per cada $\epsilon > 0$, escollim $x \in M$ tal que $\|x\| \leq 1$ i $|\lambda(x)| > \|\lambda\| - \epsilon$. Aleshores, $|\Lambda(x)| > \|\lambda\| - \epsilon$, i, consegüentment, com que això val per tot $\epsilon > 0$, $\|\Lambda\| \geq \|\lambda\|$. Com hem vist les dues desigualtats, tenim que $\|\Lambda\| = \|\lambda\|$, finalitzant la demostració.

□

Veiem ara el cas complex del teorema. En aquest cas considerem un espai vectorial complex X , un subespai M de X i un funcional lineal complex acotat λ en M . Volem estendre el funcional λ a un funcional Λ definit a tot X de manera que $\|\Lambda\| = \|\lambda\|$.

Demostració del Teorema de Hahn-Banach (Cas Complex): Definim els funcionals reals lineals λ_1 i λ_2 com

$$\lambda(z) = \lambda_1(z) + i\lambda_2(z) = \operatorname{Re}\lambda(z) + i\operatorname{Im}\lambda(z)$$

per $z \in M$. Notem que λ_i , per $i=1,2$, satisfà que

$$|\lambda_i(z)| \leq |\lambda(z)| \leq \|\lambda\| \cdot \|z\|$$

per cada $z \in M$.

Si veiem X com un espai vectorial sobre $\mathbb{R}$, la versió real del Teorema de Hahn-Banach ens diu que existeix un funcional lineal real acotat Λ_1 en X tal que $\Lambda_1(z) = \lambda_1(z)$ per cada $z \in M$ i $\|\Lambda_1\| = \|\lambda_1\|$.

Recordem que, donat $z \in \mathbb{C}$, se satisfà que $-\operatorname{Re}(iz) = \operatorname{Im}(z)$. Per tant, tenim que $-\operatorname{Re}(i\Lambda z) = \operatorname{Im}(\Lambda z)$. Tenint en compte això, definim ara

$$\Lambda(z) = \Lambda_1(z) - i\Lambda_1(iz).$$

Aquesta és l'extensió que busquem. Provem que satisfà les propietats demanades:

- Λ és $\mathbb{C}$ -lineal en X : Clarament, per $z_1, z_2 \in X$,

$$\begin{aligned}\Lambda(z_1 + z_2) &= \Lambda_1(z_1 + z_2) - i\Lambda_1(iz_1 + iz_2) \\ &= \Lambda_1(z_1) - i\Lambda_1(iz_1) + \Lambda_1(z_2) - i\Lambda_1(iz_2) = \Lambda(z_1) + \Lambda(z_2).\end{aligned}$$

Per altra banda, prenent $z = iy$, tenim que:

$$\Lambda(iy) = \Lambda_1(iy) - i\Lambda_1(-y) = \Lambda_1(iy) + i\Lambda_1(y) = i(\Lambda_1(y) - i\Lambda_1(iy)) = i\Lambda(y).$$

Amb aquesta igualtat és fàcil veure que,

$$\begin{aligned}\Lambda((a + ib)z) &= \Lambda(az) + \Lambda(ibz) \\ &= a\Lambda(z) + ib\Lambda(z) \\ &= (a + ib)\Lambda(z).\end{aligned}$$

- Λ és una extensió de λ : Per $z \in M$,

$$\lambda_1(iz) + i\lambda_2(iz) = \lambda(iz) = i\lambda(z) = -\lambda_2(z) + i\lambda_1(z).$$

Com λ_1 i λ_2 prenen valors reals, tenim, igualant les parts reals,

$$-\lambda_2(z) = \lambda_1(iz).$$

Per tant,

$$\lambda(z) = \lambda_1(z) + i\lambda_2(iz) = \lambda_1(z) - i\lambda_1(iz) = \Lambda_1(z) - i\Lambda_1(iz) = \Lambda(z)$$

per cada $z \in M$, el que prova que Λ és una extensió de λ .

- $\|\Lambda\| = \|\lambda\|$: Tenim que $\|\lambda\| \leq \|\Lambda\|$. Veiem l'altra desigualtat: per $z \in X$, escrivim $\Lambda(z)$ en forma polar $|\Lambda(z)|e^{i\theta}$, amb $\theta \in \mathbb{R}$. Tenim doncs que

$$|\Lambda(z)| = e^{-i\theta} \Lambda(z) = \Lambda(e^{-i\theta} z),$$

cosa que prova que $\Lambda(e^{-i\theta} z)$ és real, i per tant, igual a la part real de Λ , $\Lambda_1(e^{-i\theta} z)$. Així,

$$|\Lambda(z)| = \Lambda_1(e^{i\theta} z) \leq \|\Lambda_1\| \cdot \|e^{i\theta} z\| \leq \|\lambda\| \cdot \|z\|,$$

que implica $\|\Lambda\| \leq \|\lambda\|$.

Com hem vist les dues desigualtats, hem provat la igualtat.

□

El Teorema de Hahn-Banach té tot de conseqüències que generalitzen, a espais normats, fets ben coneguts a $\mathbb{R}^n$. Donem ara alguns corol·laris directes del Teorema de Hahn-Banach:

Corol·lari 3.2.1. ([2, Corollary 3.1.3]) Per tot $x_0 \in \mathcal{S}(X) = \{x \in X : \|x\| = 1\}$ (l'esfera unitat de X), existeix $\lambda \in X^*$ tal que $\|\lambda\|_{X^*} = 1$, i $\lambda(x_0) = 1$. Per tant, el funcional lineal λ assoleix el seu suprem en la bola unitat en el vector x_0 .

Demostració. Considerem el subespai 1-dimensional $M_0 = \{\mu x_0 : \mu \in \mathbb{R}\}$ i el funcional lineal $\lambda_0(\mu x_0) = \mu$. Aleshores, $\|\lambda_0\|_{M_0^*} = 1$. Pel Teorema de Hahn-Banach, existeix una extensió λ de λ_0 amb les propietats desitjades. $\square$

Corol·lari 3.2.2. ([2, Corollary 3.1.4]) Per tot $x_0 \in X$, existeix $\lambda \in X^* \setminus \{0\}$ tal que $\lambda(x_0) = \|\lambda\| \cdot \|x_0\|$.

Demostració. Apliquem el Corol·lari 3.2.1 al vector $x_0/\|x_0\| \in \mathcal{S}(X)$. $\square$

Corol·lari 3.2.3. ([2, Corollary 3.1.5]) Per tot $x_1, x_2 \in X$ tal que $x_1 \neq x_2$, existeix $\lambda \in X^*$ tal que $\lambda(x_1) \neq \lambda(x_2)$.

Demostració. Apliquem el Corol·lari 3.2.2 amb $x_0 = x_1 - x_2$. $\square$

Corol·lari 3.2.4. ([2, Corollary 3.1.6]) X^* és un conjunt total, és a dir, si $\lambda(x) = 0$ per tot $\lambda \in X^*$, aleshores $x = 0$.

Demostració. Surt de forma immediata del Corol·lari 3.2.3. $\square$

Capítol 4

Aplicacions del Teorema de Hahn-Banach

En aquest capítol veurem diferents aplicacions del Teorema de Hahn-Banach, tant en altres branques de les matemàtiques (com la geometria o la teoria de conjunts), com en altres àmbits més aplicats com l'economia o la teoria de jocs.

4.1 Teorema de Separació

En aquesta Secció presentarem el Teorema de Separació de Hahn-Banach. El teorema de separació de Hahn-Banach és un dels pilars fonamentals de l'anàlisi funcional. Aquest teorema formalitza el concepte intuïtiu de "separar" dos conjunts convexos disjunts en un espai vectorial mitjançant un hiperplà (o una funció lineal).

La importància del teorema radica en la seva capacitat de proporcionar eines per analitzar propietats geomètriques i funcionals dels espais vectorials, permetent establir separacions clares entre conjunts que comparteixen certes característiques com la convexitat i la topologia.

El resultat s'utilitza àmpliament en camps com l'optimització, la teoria de conjunts convexos, la programació lineal i la teoria de jocs, així com en l'estudi de problemes de dualitat i extensions de funcionals lineals.

En el seu nucli, el teorema demostra que és possible trobar un funcional lineal continu que actui com a "frontera" entre dos conjunts convexos disjunts, preservant la seva separació. Aquesta idea té aplicacions profundes en la resolució de problemes variats, des de l'anàlisi de sistemes lineals fins a la física matemàtica.

El Teorema de separació de Hahn-Banach és una generalització del Teorema de Separació per Hiperplans en espais vectorials normats.

Teorema de Separació per Hiperplans a $\mathbb{R}^n$

Començem definint que és un hiperplà a $\mathbb{R}^n$:

Definició 4.1.1. *Siguin $p \in \mathbb{R}^n$ diferent de zero i $c \in \mathbb{R}$ una constant. El conjunt $\mathcal{H}(p, c)$ format per tots els vectors $x \in \mathbb{R}^n$ tal que*

$$p \cdot x = p_1x_1 + p_2x_2 + \dots + p_nx_n = c$$

és un subconjunt de $\mathbb{R}^n$ anomenat hiperplà. Més concretament,

$$\mathcal{H}(p, c) = \{x \in \mathbb{R}^n \mid p \cdot x = c\} \subseteq \mathbb{R}^n.$$

Notem també que p pot ser pensat com un vector ortogonal a cada punt de l'hiperplà.

Com podem veure a la definició, els hiperplans tenen una dimensió menys que l'espai on estàn definits. Per exemple, en $\mathbb{R}^2$, els hiperplans són rectes, mentre que en $\mathbb{R}^3$ són plans. A més, els hiperplans divideixen $\mathbb{R}^n$ en dues regions, que denotarem:

$$\mathcal{O}(p, c) = \{x \in \mathbb{R}^n \mid p \cdot x > c\},$$

$$\mathcal{P}(p, c) = \{x \in \mathbb{R}^n \mid p \cdot x < c\}.$$

També podem considerar les clausures d'aquests dues regions, que són:

$$\overline{\mathcal{O}(p, c)} = \{x \in \mathbb{R}^n \mid p \cdot x \geq c\},$$

$$\overline{\mathcal{P}(p, c)} = \{x \in \mathbb{R}^n \mid p \cdot x \leq c\}.$$

Donem ara la definició de separació en aquest context:

Definició 4.1.2.

- *Direm que dos conjunts no buits A i B en $\mathbb{R}^n$ estan separats per un hiperplà si existeix un vector $p \in \mathbb{R}^n$ no nul i una constant $c \in \mathbb{R}$ tal que:*

$$p \cdot a \geq c \geq p \cdot b \quad \text{per a tot } a \in A \text{ i tot } b \in B.$$

- *Direm que dos conjunts no buits A i B en $\mathbb{R}^n$ estan estrictament separats per un hiperplà si existeix un vector $p \in \mathbb{R}^n$ no nul i una constant $c \in \mathbb{R}$ tal que:*

$$p \cdot a > c > p \cdot b \quad \text{per a tot } a \in A \text{ i tot } b \in B.$$

Les definicions anteriors estableixen que A i B es troben en diferents regions $\overline{\mathcal{O}(p, c)}$ i $\overline{\mathcal{P}(p, c)}$ (en el cas de separació), o bé en diferents $\mathcal{O}(p, c)$ i $\mathcal{P}(p, c)$ (en el cas de separació estricta).

Podem ja presentar els dos teoremes de separació per hiperplans:

Teorema 4.1.1 (Teorema de separació dèbil per hiperplans). (*[7, Theorem 3.6]*) *Siguin A i B dos subconjunts no buits i convexos de $\mathbb{R}^n$. Si A i B són disjunts, aleshores A i B poden ser separats per un hiperplà.*

Teorema 4.1.2 (Teorema de separació forta per hiperplans). (*[7, Theorem 3.7]*) *Siguin A i B dos subconjunts no buits i convexos de $\mathbb{R}^n$. Suposem que A és tancat i B és compacte. Si A i B són disjunts, aleshores A i B poden ser estrictament separats per un hiperplà.*

Teorema de Separació per Hiperplans a espais normats

Definim primer que és un hiperplà a un espai normat:

Definició 4.1.3. *Sigui V un espai vectorial normat real, F un funcional lineal continu, i $c \in \mathbb{R}$ una constant. Un hiperplà en V és el conjunt de la forma*

$$\mathcal{H}(F, c) = \{x \in V \mid F(x) = c\}.$$

Observem que els hiperplans en $\mathbb{R}^n$ són només un cas particular d'aquesta definició, on el funcional lineal continu és el producte escalar.

Definició 4.1.4. *Es diu que dos subconjunts no buits A i B d'un espai vectorial normat real V estan separats per un hiperplà si existeix un funcional lineal continu F en V i una constant $c \in \mathbb{R}$ tals que*

$$F(a) \geq c \geq F(b) \quad \text{per a tot } a \in A \text{ i } b \in B.$$

Considerem ara les dues regions tancades d'un espai vectorial normat real V

$$\overline{\mathcal{O}(F, c)} = \{x \in V \mid F(x) \geq c\},$$

$$\overline{\mathcal{P}(F, c)} = \{x \in V \mid F(x) \leq c\},$$

llavors aquesta definició de separació estableix que A i B es troben en regions diferents $\overline{\mathcal{O}(F, c)}$ i $\overline{\mathcal{P}(F, c)}$.

Donem ara la definició de conjunt convex en un espai normat:

Definició 4.1.5. *Un conjunt A en un espai vectorial V és convex si, per tot $x, y \in A$ i per tot $t \in [0, 1]$, se satisfà que $tx + (1 - t)y \in A$.*

Finalment doncs presentem el Teorema de Separació de Hahn-Banach, que és una generalització del Teorema de separació dèbil per hiperplans.

Teorema 4.1.3 (Teorema de separació de Hahn-Banach). *([7, Theorem 4.24]) Si guin A i B dos subconjunts no buits i convexos d'un espai vectorial normat real V . Supposem que A i B són disjunts i que A té un punt interior. Aleshores A i B poden ser separats per un hiperplà.*

Per poder demostrar el Teorema de Separació de Hahn-Banach necessitem els següents lemes:

Lema 4.1.1. *([7, Lemma 1.5]) Si A i B són subconjunts convexos d'un espai vectorial V , llavors $A + B$ és convex.*

Demostració. Sigui x i y dos elements de $A + B$.

Aleshores $x = a_1 + b_1$ per $a_1 \in A$, $b_1 \in B$, i $y = a_2 + b_2$ per $a_2 \in A$, $b_2 \in B$.

Així doncs, per a tot $t \in [0, 1]$,

$$\begin{aligned} tx + (1 - t)y &= t(a_1 + b_1) + (1 - t)(a_2 + b_2) \\ &= ta_1 + tb_1 + (1 - t)a_2 + (1 - t)b_2 \\ &= \underbrace{[ta_1 + (1 - t)a_2]}_{\in A} + \underbrace{[tb_1 + (1 - t)b_2]}_{\in B}. \end{aligned}$$

Per tant, per a tot $x, y \in A + B$ i tot $t \in [0, 1]$, tenim que $tx + (1 - t)y \in A + B$, la qual cosa significa que el conjunt és convex. $\square$

Lema 4.1.2. ([7, Lemma 4.23]) *Sigui C un subconjunt convex no buit d'un espai vectorial normat real V tal que C conté un punt interior, i sigui $x_0 \notin C$. Aleshores existeix un funcional lineal continu F a V tal que $F(x) \leq F(x_0)$ per a tot $x \in C$.*

Tenim ja totes les eines per demostrar el Teorema de Separació de Hahn-Banach:

Demostració del Teorema de Separació de Hahn-Banach. Considerem el conjunt $Z = B - A$, que és convex pel Lema 4.1.1. Com que A i B són disjunts, $0 \notin Z$. A més, és cert que Z té necessàriament un punt interior. Per tant, el conjunt Z amb el punt 0 compleix les condicions del Lema 4.3.10, de manera que existeix un funcional lineal continu F tal que $F(z) \leq F(0) = 0$ per a tot $z \in Z$ (recordem que per a qualsevol funcional lineal f , tenim que $f(0) = 0$).

Ara bé, observem que tots els $z \in Z$ es poden escriure com $b - a$ per a $a \in A$ i $b \in B$. Així doncs, tenim que

$$F(z) = F(b - a) = F(b) - F(a) \leq 0,$$

o equivalentment, $F(a) \geq F(b)$ per a tot $a \in A$ i $b \in B$. Per tant,

$$\inf_{a \in A} F(a) \geq \sup_{b \in B} F(b),$$

i sabem que $\inf_{a \in A} F(a) > -\infty$, ja que està acotat inferiorment per $F(b_0)$, on b_0 pot ser qualsevol punt de B (i $F(b_0) \neq -\infty$ per la continuïtat de F). De manera similar, sabem que $\sup_{b \in B} F(b) < \infty$.

Així, podem triar qualsevol nombre real c entre $\sup_{b \in B} F(b)$ i $\inf_{a \in A} F(a)$ per obtenir que

$$F(a) \geq c \geq F(b)$$

per a tot $a \in A$ i $b \in B$. $\square$

Hi ha molts altres resultats de separació que es poden establir en un espai vectorial normat amb el teorema de Hahn-Banach, incloent una generalització del Teorema Fort de Separació per Hiperplans.

4.2 Teorema Minimax

El Teorema Minimax, formulat per John von Neumann el 1928, és un dels resultats fonamentals en la Teoria de Jocs, especialment en jocs de suma zero de dos jugadors. Aquest resultat estableix que, en aquest tipus de jocs, cada jugador pot adoptar una estratègia òptima que minimitzi la seva pèrdua màxima possible, assegurant un resultat equilibrat entre els dos jugadors, anomenat equilibri Minimax.

Conceptes bàsics de la Teoria de Jocs

Per entendre aquest resultat és essencial entendre alguns conceptes bàsics de la Teoria de jocs.

1. Jocs de suma zero

Un joc de suma zero és un tipus de joc en què el guany d'un jugador és exactament igual a la pèrdua de l'altre jugador. Siguin u_1 i u_2 els resultats dels jugadors, llavors és complex que

$$u_1 + u_2 = 0.$$

Això implica que qualsevol avantatge obtingut per un jugador es tradueix en una pèrdua equivalent per l'altre.

Exemple: Suposem que dues parts, A el comprador i B el venedor, acorden un contracte de futurs per a la compra de petroli. Les condicions són les següents:

- El preu pactat per barril de petroli és de 70 USD.
- La transacció s'executarà d'aquí a un mes.
- El preu real del petroli al mercat al final del mes determinarà els resultats del contracte.

Hi ha dos escenaris possibles:

1. Si el preu del petroli al final del mes és superior a 70 USD, per exemple 80 USD:
 - El comprador A (que ha comprat el futur) obté un guany de $80 - 70 = 10$ USD per barril.
 - El venedor B (que ha venut el futur) perd exactament 10 USD per barril.
 - La suma neta dels pagaments és: $+10 - 10 = 0$.
2. Si el preu del petroli al final del mes és inferior a 70 USD, per exemple 60 USD:
 - El comprador A perd $70 - 60 = 10$ USD per barril.
 - El venedor B guanya 10 USD per barril.

- La suma neta dels pagaments continua sent $-10 + 10 = 0$.

Observem que els guanys d'un són exactament iguals a les pèrdues de l'altre, per tant, és un joc de suma zero. Aquest tipus de transacció té lloc diàriament en els mercats financers, especialment en derivats com futurs, opcions i swaps.

2. Estratègies

Una estratègia és el pla d'acció que segueix un jugador per maximitzar els seus beneficis o minimitzar les seves pèrdues. N'hi ha dos tipus principals:

- **Estratègia pura:** El jugador escull una acció específica i la segueix sempre.
- **Estratègia mixta:** El jugador assigna probabilitats a diverses accions possibles i escull de manera aleatòria d'acord amb aquestes probabilitats.

Exemple: El joc pedra-paper-tisora és també un joc de suma zero. Una estratègia pura seria "jugar sempre pedra", mentre que una estratègia mixta podria ser "jugar pedra, paper i tisora amb probabilitats iguals del 33,3%".

3. Matriu de pagaments

Un joc es pot representar mitjançant una matriu de pagaments A , on cada entrada $A(i, j)$ indica el guany del jugador 1 quan:

- El jugador 1 escull l'estratègia i .
- El jugador 2 escull l'estratègia j .

El guany del jugador 2 serà simplement $-A(i, j)$.

Exemple: Dues empreses, A i B , competeixen per un mercat limitat. Cada empresa pot triar entre dues estratègies:

- (i) Reduir preus,
- (ii) Mantenir preus.

Els beneficis de cada empresa depenen de les decisions de l'altra, i el joc és de suma zero (els guanys d'una són exactament iguals a les pèrdues de l'altra).

La matriu de pagaments per a l'empresa A és:

	Reduir preus (B)	Mantenir preus (B)
Reduir preus (A)	-5	+10
Mantenir preus (A)	+5	0

- Si ambdues empreses decideixen reduir preus, A perd 5 i B guanya 5.

- Si A redueix preus i B manté preus, A guanya 10 i B perd 10.
- Si A manté preus i B redueix preus, A perd 5 i B guanya 5.
- Si ambdues empreses mantenen preus, ningú guanya ni perd (0).

4. Valor del joc

El valor del joc, denotat per v , és el resultat que cada jugador pot garantir-se en condicions òptimes. Aquest valor depèn de les estratègies adoptades:

$$v = \begin{cases} > 0 & \text{si el joc afavoreix el jugador 1,} \\ < 0 & \text{si el joc afavoreix el jugador 2,} \\ = 0 & \text{si el joc és equilibrat.} \end{cases}$$

5. Estratègies òptimes

Les estratègies òptimes són aquelles que garanteixen el millor resultat possible per a un jugador, independentment del que faci l'altre.

Exemple: Tornem al joc de les dues empreses, on les estratègies es descriuen per la matriu següent:

	Reduir preus (B)	Mantenir preus (B)
Reduir preus (A)	-5	+10
Mantenir preus (A)	+5	0

Sigui:

- p : probabilitat que l'empresa A redueixi els preus.
- $1 - p$: probabilitat que l'empresa A mantingui els preus.
- q : probabilitat que l'empresa B redueixi els preus.
- $1 - q$: probabilitat que l'empresa B mantingui els preus.

L'objectiu és trobar les probabilitats òptimes que equilibrin el joc per ambdues empreses.

El valor esperat, és a dir, el valor mitjà del resultat per l'empresa A quan implementa una estratègia mixta es calcula amb la següent fórmula:

$$E_A = p \cdot [q \cdot (-5) + (1 - q) \cdot 10] + (1 - p) \cdot [q \cdot 5 + (1 - q) \cdot 0]$$

Desenvolupem cada component de la fórmula:

$$E_A = p \cdot (-5q + 10 - 10q) + (1 - p) \cdot (5q)$$

Distribuïm els termes:

$$E_A = p \cdot (-15q + 10) + 5q(1 - p)$$

L'objectiu és trobar p i q tal que ambdós jugadors estiguin en equilibri. Aquests càlculs es resolden analitzant l'estratègia òptima en equilibris de Nash. Les probabilitats òptimes obtingudes són:

$$p = \frac{1}{3}, \quad q = \frac{2}{3}$$

L'empresa A ha de reduir els preus amb probabilitat $\frac{1}{3}$ i mantenir els preus amb probabilitat $\frac{2}{3}$.

L'empresa B ha de reduir els preus amb probabilitat $\frac{2}{3}$ i mantenir els preus amb probabilitat $\frac{1}{3}$.

Aquest equilibri garanteix que cap empresa pugui ser explotada per l'altra, mantenint un joc estable entre ambdues.

6. Teorema Minimax

Presentem ara el Teorema Minimax generalitzat. Recordem que aquest teorema garanteix l'existència d'estratègies òptimes per als dos jugadors.

Teorema 4.2.1 (Teorema Minimax Generalitzat). *Siguin X un conjunt convex compacte i Y un conjunt convex qualsevol subconjunts d'un espai vectorial normat. Si $f : X \times Y \rightarrow \mathbb{R}$ és una funció contínua que és còncaua respecte a $x \in X$ i convexa respecte a $y \in Y$, llavors:*

$$\min_{y \in Y} \max_{x \in X} f(x, y) = \max_{x \in X} \min_{y \in Y} f(x, y).$$

L'aplicació del Teorema Minimax als jocs de suma zero és la següent:

Teorema 4.2.2 (Teorema del Minimax en Jocs de Suma Zero). *En un joc de suma zero, sempre existeix una estratègia òptima per a cada jugador que permet:*

- *Maximitzar el mínim guany possible, independentment de l'estratègia de l'oponent.*
- *Minimitzar la pèrdua màxima possible, prevenint que l'oponent obtingui un benefici elevat.*

En termes generals, la seva expressió matemàtica és:

$$v = \max_{s_1} \min_{s_2} U(s_1, s_2) = \min_{s_2} \max_{s_1} U(s_1, s_2,$$

on:

- $U(s_1, s_2)$: és la funció de pagament en funció de les estratègies s_1 i s_2 ,
- s_1 : són les estratègies disponibles per al jugador 1,
- s_2 : són les estratègies disponibles per al jugador 2,
- v : és el valor del joc.

El teorema minimax implica que:

- Cada jugador selecciona una estratègia que li garanteix el millor resultat possible en el cas més desfavorable,
- Aquest enfocament ajuda a prendre decisions en escenaris on un jugador intenta evitar la pèrdua en el pitjor cas.

En altres paraules, es tracta d'anticipar la resposta de l'oponent i triar una estratègia que ofereixi un equilibri entre seguretat i resultats òptims.

Exemple: Considerem el joc de pedra-paper-tisora com un joc de suma zero amb la següent matriu de pagaments per al jugador 1:

$$A = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix},$$

on:

- Les files representen les estratègies del jugador 1: pedra (1), paper (2), tisora (3),
- Les columnes representen les estratègies del jugador 2: pedra (1), paper (2), tisora (3),
- A_{ij} és el guany del jugador 1 quan aquest tria l'estratègia i i el jugador 2 tria l'estratègia j .

Segons el Teorema Minimax, busquem un valor v i estratègies mixtes òptimes per als jugadors; $p = (p_1, p_2, p_3)$ per al jugador 1 i $q = (q_1, q_2, q_3)$ per al jugador 2, on:

$$\begin{aligned} p_1 + p_2 + p_3 &= 1, & p_i &\geq 0 \quad \forall i, \\ q_1 + q_2 + q_3 &= 1, & q_j &\geq 0 \quad \forall j. \end{aligned}$$

El valor del joc v satisfà:

$$v = \max_p \min_q p^\top A q = \min_q \max_p p^\top A q.$$

Per aquest joc, es pot demostrar que:

- Les estratègies òptimes per a tots dos jugadors són mixtes, amb $p = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ i $q = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.
- El valor del joc és $v = 0$, ja que cap jugador té un avantatge sobre l'altre si tots dos segueixen les estratègies òptimes.

Demostració del Teorema Minimax

Provarem el teorema en $\mathbb{R}^n$ [1].

Siguin $\phi(x, y)$ una funció real de dues variables $x, y \in \mathbb{R}^n$ i $A, B \subseteq \mathbb{R}^n$ conjunts convexos. La variable x denota l'estratègia del primer jugador, i la variable y l'estratègia del segon jugador.

Un dels jugadors escull estratègies (punts) en A amb l'objectiu de maximitzar $\phi(x, y)$ (o equivalentment, de minimitzar $-\phi(x, y)$). Aquest jugador s'anomena el jugador maximitzador.

L'altre jugador escull estratègies (punts) de B amb l'objectiu de minimitzar $\phi(x, y)$ (o equivalentment, de maximitzar $-\phi(x, y)$). Aquest jugador s'anomena jugador minimitzador.

La funció $\phi(x, y)$ és la funció de pagaments. El valor $\phi(x_0, y_0)$ representa simultàniament el guany del jugador maximitzador i la pèrdua del jugador minimitzador quan s'escullen les estratègies x_0 i y_0 . Aquesta simultànitat ve donada pel fet que en un joc de suma zero, el guany d'un jugador és la pèrdua de l'altre.

Provem el següent resultat, que és una versió més general del Teorema 4.2.1.

Teorema 4.2.3. *Siguin $A, B \in \mathbb{R}^n$ amb $A, B \neq \emptyset$ i B compacte. Suigui $\phi : A \times B \rightarrow \mathbb{R}$ tal que :*

- (i) $\phi(\cdot, y)$ és còncava per tot $y \in B$. És a dir, si $t_1, \dots, t_n \in [0, 1]$ amb $\sum_{i=1}^n t_i = 1$, aleshores

$$\phi(t_1 x_1 + \dots + t_n x_n, y) \geq t_1 \phi(x_1, y) + \dots + t_n \phi(x_n, y)$$

per tot $x_1, \dots, x_n \in A$.

- (ii) $\phi(x, \cdot)$ és convexa per tot $x \in A$. És a dir, amb la mateixa notació que (i), tenim ara que

$$\phi(x, t_1 y_1 + \dots + t_n y_n) \leq t_1 \phi(x, y_1) + \dots + t_n \phi(x, y_n)$$

per tot $y_1, \dots, y_n \in B$.

- (iii) $\phi(x, \cdot)$ és semicontínua inferior a B per tot $x \in A$. És a dir, els conjunts

$$\{(y, t) \in B \times \mathbb{R} : t \geq \phi(x, y)\}$$

són tancats a $\mathbb{R}^{n+1}$.

Aleshores se satisfà que

$$\inf_{y \in B} \sup_{x \in A} \phi(x, y) = \sup_{x \in A} \inf_{y \in B} \phi(x, y).$$

Demostració. Veiem les dues desigualtats:

$\geq$: Aquesta desigualtat és sempre certa: per tot $x \in A, y \in B$ tenim que

$$\inf_{y \in B} \phi(x, y) \leq \phi(x, y).$$

Agafant el suprem en $x \in A$, tenim que

$$\sup_{x \in A} \inf_{y \in B} \phi(x, y) \leq \sup_{x \in A} \phi(x, y)$$

per tot $y \in B$. En particular, agafant l'ínfim en B , ja tenim la desigualtat.

$\leq$: Sigui $p = \inf_{y \in B} \sup_{x \in A} \phi(x, y)$. Sigui $\alpha < p$ qualsevol. Serà suficient veure que

$$\sup_{x \in A} \inf_{y \in B} \phi(x, y) \geq \alpha$$

ja que si $\sup_{x \in A} \inf_{y \in B} \phi(x, y) \geq \alpha$ per tot $\alpha < p$, també $\sup_{x \in A} \inf_{y \in B} \phi(x, y) \geq p$, com volem veure.

Sigu, per $x \in A$, el conjunt

$$C(x) = \{y \in B : \phi(x, y) \leq \alpha\}.$$

Aquest és un subconjunt tancat (per l'hipòtesi (ii) de $\phi(x, y)$) dins el compacte B , per tant, és també compacte. Com que $\alpha < \sup_{x \in A} \phi(x, y)$ per tot $y \in B$, tenim que, per tot $y \in B$ existeix $x \in A$ tal que $\phi(x, y) > \alpha$. És a dir, que

$$\bigcap_{x \in A} C(x) = \emptyset.$$

Aleshores, els complementaris $\bigcup_{x \in A} C(x)^C$ són un recobriment obert de B . Com que B és compacte, existeix un subrecobriment finit de B , és a dir, existeixen $x_1, \dots, x_n \in A$ tals que $B \subseteq \bigcup_{i=1}^n C(x_i)^C$. Així, per tot $y \in B$, existeix $i \in \{1, \dots, n\}$ tal que $y \notin C(x_i)$, o equivalentment, $\phi(x_i, y) > \alpha$. Aleshores,

$$\inf_{y \in B} \sup_{i=1, \dots, n} \phi(x_i, y) > \alpha.$$

Considerem ara el conjunt

$$E := \{(z, r) \in \mathbb{R}^{n+1} : \exists y \in B \text{ tal que } \phi(x_i, y) \leq r + z_i, i = 1, \dots, n\}.$$

Observem que E és convex : si $(z_1, r_1), (z_2, r_2) \in E$, aleshores

$$t(z_1, r_1) + (1 - t)(z_2, r_2) \in E,$$

per tot $t \in [0, 1]$, ja que $\phi(x, \cdot)$ és convexa (per l'hipòtesi (ii)). Essent $(z_j, r_j) \in E$, existeixen $y_1, y_2 \in B$ tals que

$$\phi(x_i, y_1) \leq r_1 + z_i^{(1)},$$

$$\phi(x_i, y_2) \leq r_2 + z_i^{(2)}$$

per $i = 1, \dots, n$. Llavors, prenent $y = ty_1 + (1 - t)y_2$, que és a B ja que B és convex, tenim que,

$$\begin{aligned}\phi(x_i, ty_1 + (1 - t)y_2) &\leq t\phi(x_i, y_1) + (1 - t)\phi(x_i, y_2) \\ &\leq t(r_1 + z_i^{(1)}) + (1 - t)(r_2 + z_i^{(2)}) \\ &= tr_1 + (1 - t)r_2 + tz_i^{(1)} + (1 - t)z_i^{(2)}.\end{aligned}$$

Per tant, $(tz_1 + (1 - t)z_2, tr_1 + (1 - t)r_2) \in E$.

Observem que $E^\circ \neq \emptyset$. En concret, $(0, 1 + \max_i \phi(x_i, y)) \in E^\circ$ per tot $y \in B$, per definició del conjunt E .

Observem també que $(0, \alpha) \notin E$, ja que per tot $y \in B$ existeix $i = 1, \dots, n$ amb $\phi(x_i, y) > \alpha$, per construcció.

Estem doncs en condicions d'aplicar el Teorema de Separació 4.2.1 al convex $E \subseteq \mathbb{R}^{n+1}$ i al punt $(0, \alpha) \notin E$: existeix $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ lineal tal que

$$F(z_1, \dots, z_n, r) \geq F(0, \dots, 0, \alpha)$$

per tot $(z_1, \dots, z_n, r) \in E$.

Equivalentment, existeixen $\lambda_1, \dots, \lambda_n, \mu \in \mathbb{R}$ tals que

$$\sum_{j=1}^n \lambda_j z_j + \mu r \geq \lambda_1 \cdot 0 + \dots + \lambda_n \cdot 0 + \mu \alpha = \mu \alpha \quad (4.2.1)$$

per tot $(z, r) \in E$.

Observem que $E + \mathbb{R}_+^{n+1} \subseteq E$, és a dir, si $(z, r) \in E$ i $z'_1, \dots, z'_n, r' \geq 0$, llavors $(z, r) + (z', r') \in E$ (això és obvi per la definició d' E).

Observem també que $\mu > 0$ ja que $(0, 1 + \max_i \phi(x_i, y)) \in E^\circ$ (si fos $\mu \leq 0$ això no valdria).

Per tant, de la desigualtat (4.2.1) tenim, dividint per $\mu > 0$,

$$\sum_{j=1}^n \frac{\lambda_j}{\mu} z_j + r \geq \alpha,$$

per tot $(z_1, \dots, z_n, r) \in E$.

Com que $(\phi(x_1, y) + r, \dots, \phi(x_n, y) + r, -r) \in E$ (obvi per definició), tenim que, en particular,

$$\sum_{j=1}^n \frac{\lambda_j}{\mu} (\phi(x_j, y) + r) - r \geq \alpha,$$

i per tant,

$$\sum_{j=1}^n \frac{\lambda_j}{\mu} \phi(x_j, y) + \left(\sum_{j=1}^n \frac{\lambda_j}{\mu} - 1 \right) r \geq \alpha$$

per tot $y \in B$ i per tot $r \in \mathbb{R}$.

Això obliga a que $\sum_{j=1}^n \frac{\lambda_j}{\mu} - 1 = 0$ ja que si fos positiu, prenent $r \rightarrow -\infty$ la desigualtat seria falsa, i si fos negatiu, prenent $r \rightarrow \infty$ la desigualtat també seria falsa. Aleshores queda, posant $t_j = \frac{\lambda_j}{\mu}$, $\sum_{j=1}^n t_j = 1$, i per tant,

$$\sum_{j=1}^n t_j \phi(x_j, y) \geq \alpha,$$

per tot $y \in B$ i per tot $r \in \mathbb{R}$.

Agafant $\hat{x} = \sum_{j=1}^n t_j x_j$ tenim, per la concavitat en x , que

$$\phi(\hat{x}, y) = \phi\left(\sum_{j=1}^n t_j x_j\right) \geq \sum_{j=1}^n t_j \phi(x_j, y) \geq \alpha,$$

per tot $y \in B$.

Així doncs, com que $x_j \in A$ i A és convex, tenim que $\hat{x} \in A$ amb $\phi(\hat{x}, y) \geq \alpha$ per tot $y \in B$, és a dir, $\hat{x} \in A$ amb $\inf_{y \in B} \phi(\hat{x}, y) \geq \alpha$.

Això prova que $\sup_{x \in A} \inf_{y \in B} \phi(x, y) \geq \alpha$ com volíem.

□

Conclusió

El teorema Minimax és una eina poderosa per analitzar i resoldre conflictes en situacions competitives. És aplicable en àrees com la intel·ligència artificial, l'economia i l'estratègia militar, oferint un marc matemàtic per comprendre la presa de decisions en entorns de competència pura.

4.3 Teoria de conjunts i Teorema de Hahn-Banach

En aquest capítol veurem una reformulació del Teorema de Hahn-Banach en el llenguatge de teoria de conjunts, i en comentarem un parell d'aplicacions.

Recordem que els Axiomes de Zermelo-Fraenkel és un sistema axiomàtic fonamental de la teoria de conjunts. Concretament, l'Axioma de l'Elecció afirma que, donada una col·lecció de conjunts no buits, sempre és possible seleccionar simultàniament un element de cada conjunt de la col·lecció. Més formalment, l'Axioma de l'Elecció estableix que si tenim una família de conjunts no buits $\{A_i\}_{i \in I}$, on I és un conjunt indexat, llavors existeix una funció de selecció f , tal que per a cada $i \in I$, $f(i) \in A_i$. Es sabut que l'Axioma de l'Elecció (un dels Axiomes de Zermelo-Fraenkel) implica el Lema de Zorn, i com hem vist a la prova del Teorema de Hahn-Banach, el Lema de Zorn implica el Teorema de Hahn-Banach.

En aquest context, existeixen diversos treballs que intenten demostrar fets ja coneguts a partir només del Teorema de Hahn-Banach (en lloc de l'Axioma de l'Elecció o el Lema de Zorn), és a dir, veure si és possible demostrar aquests fets utilitzant únicament el Teorema de Hahn-Banach.

Passem a continuació a veure els dos resultats esmentats: l'existència de conjunts no mesurables de Lebesgue i la paradoxa de Banach-Tarsky. Les demostracions s'escapen de l'objectiu primari del treball, però m'ha semblat interessant l'aplicació del Teorema de Hahn-Banach en la teoria de conjunts. Per tant, no comentarem els resultats amb tot detall, ja que això requeriria entrar en la teoria de conjunts, que ens allargaria molt i no és l'objectiu del treball.

4.3.1 Reformulació del Teorema de Hahn-Banach en la teoria de conjunts

Començem veient la reformulació del Teorema de Hahn-Banach, que és en la que es basen els resultats del capítol. Per fer això ens cal presentar nocions bàsiques de la teoria de conjunts.

Definició 4.3.1. Una àlgebra Booleana és una estructura $(\mathcal{B}, \wedge, \vee, \neg, 0, 1)$ on:

- (i) $\mathcal{B}$ és no buit,
- (ii) $\wedge$ i $\vee$ són operacions binàries en $\mathcal{B}$,
- (iii) $\neg$ és una operació unària en $\mathcal{B}$,
- (iv) 0 i 1 són elements de $\mathcal{B}$ (representant els valors fals i cert respectivament),

i se satisfan les propietats següents:

1. *Commutativitat:*

$$a \wedge b = b \wedge a \quad i \quad a \vee b = b \vee a, \quad \forall a, b \in \mathcal{B}.$$

2. *Associativitat:*

$$(a \wedge b) \wedge c = a \wedge (b \wedge c) \quad i \quad (a \vee b) \vee c = a \vee (b \vee c), \quad \forall a, b, c \in \mathcal{B}.$$

3. *Distributivitat:*

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c) \quad i \quad a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c), \quad \forall a, b, c \in \mathcal{B}.$$

4. *Identitats:*

$$a \wedge 1 = a \quad i \quad a \vee 0 = a, \quad \forall a \in \mathcal{B}.$$

5. *Complementarietat:*

$$a \wedge \neg a = 0 \quad i \quad a \vee \neg a = 1, \quad \forall a \in \mathcal{B}.$$

Definició 4.3.2. *Sigui $\mathcal{B}$ una àlgebra Booleana. Una mesura de probabilitat finitament additiva en $\mathcal{B}$ és una aplicació $\mu : \mathcal{B} \rightarrow [0, 1]$ que compleix les condicions següents:*

- (i) $\mu(1_{\mathcal{B}}) = 1$, on $1_{\mathcal{B}}$ indica la unitat de l'àlgebra Booleana, i 1 és el nombre 1.
- (ii) $\mu(x \vee y) = \mu(x) + \mu(y)$, sempre que $x \wedge y = 0$.

A partir d'aquí, direm *mesura Booleana* a una mesura de probabilitat finitament additiva.

Un cop definits aquests conceptes passem a veure la reformulació del Teorema de Hahn-Banach en llenguatge a la teoria de conjunts, que ve donada pel següent teorema:

Teorema 4.3.1. *([5, Theorems 7.3, 8.3]) Són equivalents:*

- (i) *El Teorema de Hahn-Banach.*
- (ii) *Tota àlgebra Booleana admet una mesura Booleana no trivial.*
- (iii) *Sigui $\mathcal{B}_0$ una subàlgebra d'una àlgebra Booleana $\mathcal{B}$ i sigui m_0 una mesura Booleana sobre $\mathcal{B}_0$. Aleshores, existeix una mesura Booleana m sobre $\mathcal{B}$ tal que $m = m_0$ a $\mathcal{B}_0$ i el rang de m està contingut en l'embolcall convex tancat del rang de m_0 .*

Passem doncs a comentar les dues aplicacions esmentades anteriorment.

4.3.2 Conjunts no mesurables de Lebesgue

L'existència dels conjunts no mesurables de Lebesgue no és trivial. Una de les aplicacions del Teorema de Hahn-Banach és la demostració de l'existència d'aquests conjunts sense fer ús de l'Axioma de l'Elecció. La demostració més coneguda de l'existència de conjunts no mesurables, la construcció de Vitali (que recordarem més endavant), fa servir l'Axioma de l'Elecció. En aquesta secció donarem una idea de com el Teorema de Hahn-Banach es fa servir per demostrar l'existència de conjunts no mesurables de Lebesgue.

Tot seguit, passem a recordar la mesura de Lebesgue a $\mathbb{R}^n$:

Mesura de Lebesgue a $\mathbb{R}^n$

La mesura de Lebesgue és una forma estàndard d'assignar una longitud, àrea o volum als subconjunts d'un espai euclidià. Un conjunt s'anomena Lebesgue-mesurable si és possible assignar-li una mesura.

Recordem que una família $\mathcal{F} \in \mathcal{P}(\mathbb{R}^n)$ és una σ -àlgebra si se satisfà:

- (i) $\mathbb{R}^n \in \mathcal{F}$,
- (ii) $A \in \mathcal{F}$ implica que $A^C \in \mathcal{F}$,
- (iii) $\{A_n\}_{n \geq 1} \subseteq \mathcal{F}$ implica que $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Definició 4.3.3. *Sigui $\mathcal{F} \in \mathcal{P}(\mathbb{R}^n)$ una σ -àlgebra. Una funció $m : \mathcal{F} \rightarrow [0, \infty]$ és una mesura si:*

- (i) $m(\emptyset) = 0$
- (ii) m és σ -additiva, és a dir, si $\{A_k\}_{k=1}^{\infty}$ és una família de conjunts disjunts dos a dos, aleshores

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

La propietat de la σ -additivitat és essencial per tenir després una bona teoria d'integració, que és la principal raó per la qual es va introduir el concepte de mesura.

El primer intent de definir la mesura de qualsevol conjunt $A \subseteq \mathbb{R}^n$ surt de considerar recobriments d' A per conjunts fonamentals, en aquest cas, per rectangles.

Definició 4.3.4. *Sigui $A \subseteq \mathbb{R}^n$. Definim la mesura exterior d' A com*

$$|A|^* := \inf \left\{ \sum_{k=1}^{\infty} v(I_k) \mid A \subseteq \bigcup_{k=1}^{\infty} I_k \right\},$$

on I_k són intervals oberts de $\mathbb{R}^n$ de la forma

$$I_k = (a_1^k, b_1^k) \times (a_2^k, b_2^k) \times \cdots \times (a_n^k, b_n^k), \text{ amb } a_i^k < b_i^k$$

$$i \ v(I_k) = (b_1^k - a_1^k) \cdots (b_n^k - a_n^k).$$

La mesura exterior $|A|^*$ està ben definida per tot $A \subseteq \mathbb{R}^n$ per l'axioma del suprem, ja que sempre $\sum_{k=1}^{\infty} v(I_k) \geq 0$. Per tant, per definició, $|A|^* \geq 0$.

Notem que la mesura exterior és σ -subadditiva, és a dir, que satisfà que donats $A_j \subseteq \mathbb{R}^n$, tenim que

$$\left| \bigcup_{j=1}^{\infty} A_j \right|^* \leq \sum_{j=1}^{\infty} |A_j|^*.$$

Aquesta propietat surt directament de la definició: fixem $\epsilon > 0$. Sabem que existeixen recobriments $A_j \subseteq \bigcup_{k=1}^{\infty} I_k^j$, on I_k^j són intervals oberts tal que

$$|A_j|^* \leq \sum_{k=1}^{\infty} v(I_k^j) \leq |A_j|^* + \frac{\epsilon}{2^j}. \quad (4.3.1)$$

Considerem ara

$$\bigcup_{j=1}^{\infty} \bigcup_{k=1}^{\infty} I_k^j$$

un recobriment de $\bigcup_{j=1}^{\infty} A_j$. Aleshores, utilitzant les desigualtats 4.3.1, obtenim:

$$\begin{aligned} \left| \bigcup_{j=1}^{\infty} A_j \right|^* &\leq \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} v(I_k^j) \\ &\leq \sum_{j=1}^{\infty} \left(|A_j|^* + \frac{\epsilon}{2^j} \right) \\ &= \sum_{j=1}^{\infty} |A_j|^* + \sum_{j=1}^{\infty} \frac{\epsilon}{2^j} \\ &= \sum_{j=1}^{\infty} |A_j|^* + \epsilon \sum_{j=1}^{\infty} \frac{1}{2^j} \\ &= \sum_{j=1}^{\infty} |A_j|^* + \epsilon. \end{aligned}$$

Ara, fent tendir $\epsilon \rightarrow 0$, obtenim que

$$\left| \bigcup_{j=1}^{\infty} A_j \right|^* \leq \sum_{j=1}^{\infty} |A_j|^*.$$

Tot i ser σ -subadditiva, la mesura exterior no és σ -additiva, és a dir, no satisfà la igualtat

$$\left| \bigcup_{j=1}^{\infty} A_j \right|^* = \sum_{j=1}^{\infty} |A_j|^*.$$

Un exemple d'això és el Conjunt de Vitali, explicat més endavant a la Secció 4.3.2.

Ara ja tenim totes les eines per definir la mesura de Lebesgue:

Definició 4.3.5. *Sigui $A \subseteq \mathbb{R}^n$. Diem que A és mesurable-Lebesgue si per tot $B \subseteq \mathbb{R}^n$ tenim que*

$$|A|^* = |A \cap B|^* + |A \cap B^C|^*.$$

Equivalentment, A és mesurable-Lebesgue si per tot $\epsilon > 0$ existeixen un obert U i un tancat C tal que $C \subseteq A \subseteq U$ i

$$|U \setminus C|^* < \epsilon.$$

Considerem el conjunt format per tots els conjunts mesurable-Lebesgue, que denotem per $\mathcal{M}$. Es pot demostrar que $\mathcal{M}$ és una σ -àlgebra (part del Teorema de Caratheodory), i a més, que

$$\begin{aligned} m : \mathcal{M} &\rightarrow [0, \infty] \\ A &\mapsto |A|^*. \end{aligned}$$

és una mesura, és a dir, que la mesura exterior restringida a $\mathcal{M}$ és una mesura. Per tant, la mesura de Lebesgue és exactament la mesura exterior (pels conjunts on funciona bé, és a dir, pels conjunts mesurable-Lebesgue).

Conjunts no mesurables de Lebesgue

Tot i que la intuïció ens diu que tots els conjunts haurien de ser mesurable-Lebesgue, la veritat és que existeixen els conjunts no mesurables-Lebesgue. L'exemple més conegut és el conjunt de Vitali. Recordem-ne tot seguit la construcció ([12]):

Considerem el conjunt $V \subset [0, 1]$ format per un representant de cadascuna de les classes d'equivalència de la relació

$$v \sim v' \iff v - v' \in \mathbb{Q}.$$

Observem que en aquesta construcció es fa servir l'Axioma de l'Elecció per triar simultàniament un element d'una col·lecció infinita de conjunts. Veurem més endavant que es pot fer servir només el Teorema de Hahn-Banach per construir conjunts no mesurables.

Volem veure que aquest conjunt és ben definit i no mesurable. Ho provarem per contradicció.

La relació d'equivalència $\sim$ és ben definida, ja que $\mathbb{Q}$ és un subgrup normal de $(\mathbb{R}, +)$. Podem pensar que $\mathbb{R} \setminus \mathbb{Q}$ està format per còpies traslladades disjunts de $\mathbb{Q}$: cada element diferent del conjunt quocient $\bar{r} \in \mathbb{R}/\mathbb{Q}$ dona lloc a un traslladat diferent $r + \mathbb{Q}$. Hi ha un nombre infinit no numerable d'elements a $\mathbb{R}/\mathbb{Q}$, i per tant, també a la partició de $\mathbb{R}$ donada pels $r + \mathbb{Q}$.

Cada element $r + \mathbb{Q}$ d'aquesta partició interseca l'interval $[0, 1]$, i per tant, l'Axioma de l'Elecció garanteix que existeix el conjunt $V \subset [0, 1]$ format per exactament un representant de cada classe.

Tot plegat, V és no numerable i $u - v \in \mathbb{R} \setminus \mathbb{Q}$ si $u, v \in V, u \neq v$.

Per veure que V no és mesurable considerem primer el conjunt auxiliar $\mathbb{Q} \cap [-1, 1]$. Com que és numerable, el podem expressar de la forma $\mathbb{Q} \cap [-1, 1] = \{q_k\}_{k=1}^{\infty}$. Definim $V_k = q_k + V$ i observem que

$$(i) \quad [0, 1] \subset \bigcup_{k=1}^{\infty} V_k \subset [-1, 2],$$

(ii) els conjunts $V_k, k \geq 1$ són disjunts dos a dos. Efectivament, si hi ha un element $x \in V_k \cap V_j, k \neq j$, tindrà la forma

$$x = q_k + v = q_j + u, \quad u, v \in V.$$

Llavors, $u - v = q_k - q_j \in \mathbb{Q}$, i per tant $u = v$, ja que hem triat un únic element de cada classe d'equivalència.

Si V fos mesurable tindríem que $|V_k| = |V|$ per tot $k \geq 1$ i per tant, per (i) i per la σ -additivitat, tindríem que

$$3 = |[-1, 2]| \geq \left| \bigcup_{k=1}^{\infty} V_k \right| = \sum_{k=1}^{\infty} |V_k| = \sum_{k=1}^{\infty} |V|.$$

La convergència d'aquesta última sèrie obligaria a que $|V| = 0$, i per tant, que

$$\left| \bigcup_{k=1}^{\infty} V_k \right| = 0,$$

però això contradiu (i). Per tant, el conjunt V no és mesurable.

Demostració de l'existència de conjunts no mesurable-Lebesgue

En tot aquest apartat suposem els Axiomes de Zermelo-Fraenkel (ZF), sense cap suposició sobre l'Axioma d'Elecció.

Posem el problema en un context més general, prenent espais X (en lloc de subconjunts de $\mathbb{R}^n$) i grups que actuen sobre X . Així, un espai de mesura (X, μ) serà un espai X amb una mesura Booleana μ a $\mathcal{P}(X)$.

Definició 4.3.6. *Sigui G un grup actuant sobre un espai X . L'expressió $IM(X, G)$ vol dir que "existeix una mesura Booleana G -invariant a $\mathcal{P}(X)$ ", és a dir, tal que*

$$\mu(gA) = \mu(A),$$

per tot $g \in G$ i per tot $A \in \mathcal{P}(X)$.

Definició 4.3.7. *Un grup G és "amenable" si existeix una mesura Booleana en $\mathcal{P}(G)$ tal que*

$$\mu(Ag) = \mu(A),$$

per tot $g \in G$ i per tot $A \in \mathcal{P}(G)$.

Suposant el Teorema de Hahn-Banach, es pot veure que molts grups són amenable (com els grups finits, els grups resolubles, les seves extensions, etc).

L'exemple segurament més senzill de grup no amenable és el grup lliure d'ordre 2 $\mathbb{F}_2$. Recordem que $\mathbb{F}_2$ és el grup format per totes les paraules sobre els generadors a i b i les seves inverses, sense cap altra relació entre ells. Cada element de $\mathbb{F}_2$ és una cadena de la forma $a^{\epsilon_1} b^{\epsilon_2} \cdots a^{\epsilon_n}$, on $\epsilon_i \in \{\pm 1\}$.

Definició 4.3.8. *Sigui G un grup. El grup G actua lliurement sobre X si per tot $g \in G$ i per tot $x \in X$ se satisfà que*

$$gx = x \text{ implica que } g = 1,$$

o equivalentment,

$$gx \neq x \text{ si } g \neq 1.$$

Sigui ara $\mathbb{S}^2$ l'esfera unitat a $\mathbb{R}^3$ i $SO(3) = \{R \in \mathbb{R}^{3 \times 3} : R^T R = I \text{ i } \det(R) = 1\}$ el grup de rotacions.

El punt de partida és el resultat següent:

Proposició 4.3.1. *[3, Proposition 3] Supposem $IM(\mathbb{S}^2, SO(3))$. Aleshores existeix una acció lliure de $\mathbb{F}_2$ en algun espai de mesura (X, μ) que preserva la mesura Booleana.*

Passem finalment al teorema principal, el qual es demostra usant només el Teorema de Hahn-Banach (no són necessaris ni l'Axioma de l'Elecció ni el Lema de Zorn).

Teorema 4.3.2. *[3, Theorem 4] Sigui G un grup actuant lliurement i preservant la mesura sobre un espai de mesura (X, μ) . Aleshores G és amenable.*

Aquest teorema, juntament amb el fet de que $\mathbb{F}_2$ no és amenable i la Proposició 4.3.1, diuen en particular, aplicat a $\mu = m$ mesura de Lebesgue, que no és cert $IM(\mathbb{S}^2, SO(3))$, i per tant, hi ha algun subconjunt de l'esfera $\mathbb{S}^2$ que no és mesurable.

4.3.3 Paradoxa de Banach-Tarsky

La paradoxa de Banach-Tarsky afirma que, donats dos subconjunts acotats de $\mathbb{R}^3$ (com una bola petita i una gran), existeixen descomposicions dels dos objectes en un numero finit de peces no solapades de manera que, juntant les peces d'una forma diferent, cada una pot ser reassemblada en l'altre. Aquest reassemblatge de les peces consisteix únicament en rotacions i desplaçaments de les peces, sense canviar-ne la forma. Una forma més informal que s'utilitza amb freqüència per explicar aquesta paradoxa és: un pèsol pot desmuntar-se i reassemblar-se per formar el sol. Donem ara l'enunciat formal:

Teorema 4.3.3. *[10, Theorem 1.1] Siguin X i Y subconjunts acotats de $\mathbb{R}^3$ amb interiors no buits. Aleshores existeixen $n \in \mathbb{N}$ i particions $\{X_j : 1 \leq j \leq n\}$, $\{Y_j : 1 \leq j \leq n\}$ de X i Y respectivament, tal que X_j és congrüent a Y_j per tot j (mirar Definició 4.3.10).*

Una altre versió d'aquest teorema és la següent: donada una bola en un espai tridimensional, existeix una descomposició de la bola en un numero finit de peces no solapades de manera que, juntant les peces d'una forma diferent, podem obtenir dues còpies idèntiques a la bola original.



La raó per la que aquest teorema es considera una paradoxa és perquè contradiu la intuïció geomètrica bàsica, ja que dividir la bola en parts i fer desplaçaments i

rotacions són operacions que conserven el volum (si tractem amb conjunts dels que se'n pot calcular el volum, és a dir, mesurables).

Al contrari que la majoria de teoremes de geometria, aquest depèn de forma crítica de l'Axioma de l'Elecció de la teoria de conjunts.

Hem vist prèviament que el Teorema de Hahn Banach implica l'existència de conjunts no mesurables.

Volem veure que, usant únicament el Teorema de Hahn-Banach, podem demostrar aquesta paradoxa, sense fer servir l'Axioma de l'Elecció. Com abans, posarem el problema en un context més general d'accions de grups G actuant sobre espais X . El grup que actuarà sobre $\mathbb{R}^3$ és el grup de moviments rígids, que presentem a continuació.

Grup de les rotacions i translacions

Definició 4.3.9. *Un moviment rígid (també anomenat transformació euclídea) és una aplicació*

$$\begin{aligned} r : \mathbb{R}^3 &\rightarrow \mathbb{R}^3 \\ x &\mapsto r(x) = \rho(x) + a, \end{aligned}$$

on $\rho \in SO(3)$ és una rotació i $a \in \mathbb{R}^3$ és un vector de translació.

Definició 4.3.10. *Diem que el conjunt A és congruent al conjunt B si existeix un moviment rígid ρ tal que $\rho(A) = B$.*

Per tant, per provar la paradoxa de Banach-Tarsky hem de veure que els subconjunts acotats X i Y de $\mathbb{R}^3$ poden dividir-se en un mateix nombre finit de trosos congruents un amb l'altre (equivalentment, com veurem més endavant, que són equidecomposables amb el grup G format únicament per rotacions i desplaçaments). Veiem primer que, efectivament, el conjunt format per les rotacions i les translacions de $\mathbb{R}^3$ té una estructura de grup.

Teorema 4.3.4. *[10, Theorem 1.11] El conjunt de moviments rígids $\mathcal{R}$ formen un grup.*

Demostració. Per veure que el conjunt de moviments rígids tenen estructura de grup cal veure que:

- (i) $\mathcal{R}$ és tancat,
- (ii) conté un element neutre,
- (iii) Cada element $r \in \mathcal{R}$ té un element invers.

Pasem a provar cada apartat:

- (i) Siguin $r(x) = \rho(x) + a$ i $s(x) = \tau(x) + b$ dos moviments rígids on ρ, τ són rotacions i a, b són punts fixats de $\mathbb{R}^3$. Hem de provar que $r \circ s \in \mathcal{R}$, però això és immediat:

$$(r \circ s)(x) = r(s(x)) = \rho(\tau(x) + b) + a = \rho(\tau(x)) + \rho(b) + a.$$

Observem que $\rho \circ \tau \in SO(3)$, ja que $SO(3)$ és un grup. A més, notem que la rotació d'un punt fix és un punt fix també, per tant $\rho(b)$ és algun punt fix de $\mathbb{R}^3$.

- (ii) Considerem el moviment rígid següent donat per $\rho = Id$ i $a = 0$:

$$e(x) = Id(x) + 0.$$

Aleshores, donat qualsevol $x \in \mathbb{R}^3$,

$$e(x) = Id(x) + 0 = x + 0 = x,$$

i per tant, aquest element és l'element neutre.

- (iii) Sigui $r(x) = \rho(x) + a \in \mathcal{R}$ qualsevol, aleshores l'invers de $r(x)$ és

$$r^{-1}(x) = \rho^{-1}(x) - \rho^{-1}(a),$$

ja que

$$\begin{aligned} (r \circ r^{-1})(x) &= r(r^{-1}(x)) \\ &= r(\rho^{-1}(x) - \rho^{-1}(a)) \\ &= \rho(\rho^{-1}(x) - \rho^{-1}(a)) + a \\ &= \rho(\rho^{-1}(x)) - \rho(\rho^{-1}(a)) + a \\ &= x - a + a \\ &= x. \end{aligned}$$

Observem que $r^{-1}(x) = \rho^{-1}(x) - \rho^{-1}(a)$ és efectivament un moviment rígid, ja que $\rho^{-1} \in SO(3)$.

□

Posem ara el problema de la paradoxa en un context més general. Introduïm els grups i els subconjunts paradoxals:

Grups i subconjunts paradoxals

Definició 4.3.11. *Sigui G un grup actuant en un conjunt X . Un subconjunt $E \subset X$ és paradoxal si existeixen dues famílies de subconjunts disjunts dos a dos $A_1, \dots, A_n$ i $B_1, \dots, B_m$ en E i existeixen $g_1, \dots, g_n, h_1, \dots, h_m$ en G tal que*

$$E = \bigcup_{i=1}^n g_i A_i = \bigcup_{j=1}^m h_j B_j.$$

Es diu que l'acció d'un grup G en un conjunt X és paradoxal (o G -paradoxal) si el conjunt X és paradoxal. Un grup G es diu que és paradoxal si l'acció de G sobre ell mateix per l'esquerra és paradoxal.

Exemple: Considerem el grup lliure de dos generadors $\mathbb{F}_2$. Siguin $a, b \in \mathbb{F}_2$ els generadors. Denotem per $\omega(x)$ el conjunt de totes les paraules de $\mathbb{F}_2$ que comencen amb x . Aleshores el grup pot ser descomposat en conjunts disjunts dos a dos de la següent manera:

$$\mathbb{F}_2 = \{e\} \cup \omega(a) \cup \omega(a^{-1}) \cup \omega(b) \cup \omega(b^{-1}).$$

Observem que $\mathbb{F}_2 \setminus \omega(x) = x\omega(x^{-1})$ per tot $x \in \{a, a^{-1}, b, b^{-1}\}$, on $\mathbb{F}_2 \setminus \omega(x)$ són les paraules que no comencen amb x . Tenim doncs dues descomposicions de $\mathbb{F}_2$:

$$\mathbb{F}_2 = \omega(a) \cup a\omega(a^{-1}) = \omega(b) \cup a\omega(b^{-1}).$$

Amb algunes suposicions addicionals, es pot transferir una descomposició paradoxal d'un grup a un conjunt sobre el qual actua. Aquest és exactament el punt on es necessita l'Axioma de l'Elecció.

Teorema 4.3.5. [4, Theorem 0.1.2] *Un grup G és paradoxal si, i només si, existeix una acció lliure en un conjunt X tal que l'acció de G sobre X és paradoxal.*

Demostració. Sigui $G = \bigcup_{i=1}^n g_i A_i = \bigcup_{j=1}^m h_j B_j$ una descomposició paradoxal de G . Per l'Axioma de l'Elecció, podem prendre un subconjunt M de X tal que contingui exactament un element de cada òrbita de G . Tenim aleshores que $\bigcup_{g \in G} gM$ és una partició disjunta de M . De fet, si $gx = hy$ per alguns $g, h \in G$ i $x, y \in X$, aleshores $x = y$ per l'elecció de M . Com l'acció és lliure, tenim que $g = h$.

Definim ara

$$\hat{A}_i = \bigcup_{g \in A_i} gM \quad \text{i} \quad \hat{B}_j = \bigcup_{h \in B_j} hM.$$

Clarament, aquests conjunts són també disjunts per construcció, i satisfan que

$$X = \bigcup_{i=1}^n g_i \hat{A}_i = \bigcup_{j=1}^m h_j \hat{B}_j.$$

Per mostrar el recíproc, considerem una òrbita $\mathcal{O}$ d'un punt en X . Llavors, l'acció de G sobre $\mathcal{O}$ és exactament una acció sobre l'espai de classes laterals per algun subgrup $H < G$ (és a dir, el conjunt format per les classes d'equivalència de G/H). Així, una descomposició paradoxal de $\mathcal{O}$ implica una descomposició paradoxal de G .

□

Aquest resultat implica, en particular, que totes les accions lliures de $\mathbb{F}_2$ admeten descomposicions paradoxals. Cal observar que, si l'acció de G sobre X és transitiva, llavors no és necessari utilitzar l'Axioma de l'Elecció.

La Paradoxa de Hausdorff

Atès que totes les accions lliures del grup lliure de dos generadors donen lloc a accions paradoxals, la idea de descomposició paradoxal en l'espai euclidià es basa en l'existència de subgrups lliures en el grup de rotacions en l'espai euclidià tridimensional $SO(3)$.

És un fet ben conegut que $SO(3)$ conté moltes còpies del grup lliure de dos generadors. De fet, una afirmació més forta és certa: si considerem $SO(3) \times SO(3)$ amb la topologia producte, llavors el conjunt de totes les parelles que generen $\mathbb{F}_2$ és dens en $SO(3) \times SO(3)$. La primera construcció explícita del subgrup lliure en $SO(3)$ es remunta a Hausdorff [4].

Teorema 4.3.6. [4, Theorem 0.2.1] *Existeixen dues rotacions en $SO(3)$ que generen el grup lliure de dos generadors.*

Cal remarcar que aquestes dues rotacions es poden donar explícitament.

Aquest resultat dona lloc a l'anomenada paradoxa de Hausdorff:

Teorema 4.3.7. [4, Theorem 0.2.2] *Existeix un subconjunt numerable en $\mathbb{S}^2$ tal que el seu complementari en $\mathbb{S}^2$ és $SO(3)$ -paradoxal.*

Conjunts equidecomposables

La noció que generalitza les descomposicions que trobem a la paradoxa de Banach-Tarski és la següent:

Definició 4.3.12. *Sigui G un grup actuant sobre el conjunt X . Dos subconjunts A, B són equidecomposables si existeixen subconjunts disjunts dos a dos $A_1, \dots, A_n \subset A$, subconjunts disjunts dos a dos $B_1, \dots, B_n \subset B$ i $g_1, \dots, g_n \in G$ tals que*

- (i) $A = \bigcup_{i=1}^n A_i$,
- (ii) $B = \bigcup_{j=1}^n B_j$,
- (iii) $g_i(A_i) = B_i$ per tot $1 \leq i \leq n$.

És fàcil provar que la equidecomposibilitat és una relació d'equivalència.

Observem que, en aquest llenguatge, la paradoxa de Banach-Tarsky (Teorema 4.3.3) diu que dos conjunts acotats no buits de $\mathbb{R}^3$ són equidecomposables per l'acció del grup format pels moviments rígids.

Proposició 4.3.2. [4, Proposition 0.3.2] *Sigui G un grup actuant en un conjunt X i $E, F \subset X$ subconjunts. Si F és paradoxal i F és equidecomposable amb E , aleshores E és paradoxal.*

Demostració. Com que F és paradoxal, tenim amb la notació de la Definició 4.3.11, que

$$F = \bigcup_{i=1}^n g_i A_i = \bigcup_{j=1}^m h_j B_j.$$

Tenim que F i E són equidecomposables, per tant, per transitivitat, tenim que E és equidecomposable tant amb $\bigcup_{i=1}^n g_i A_i$ com amb $\bigcup_{j=1}^m h_j B_j$ (ja que l'equidecomposibilitat és una relació d'equivalència, i per tant, transitiva). Això implica que E és paradoxal. $\square$

El següent resultat és molt semblant ja a la paradoxa de Banach-Tarski:

Proposició 4.3.3. *[4, Proposition 0.3.3] Sigui D un subconjunt numerable de $\mathbb{S}^2$. Aleshores $\mathbb{S}^2$ i $\mathbb{S}^2 \setminus D$ són $SO(3)$ -equidecomposables.*

Hi ha un corol·lari directe que surt d'aquest resultat i de la paradoxa de Hausdorff:

Corol·lari 4.3.1. *[4, Corollary 0.3.4] $\mathbb{S}^2$ és $SO(3)$ -paradoxal.*

Ara ja estem a punt de provar la paradoxa de Banach-Tarski.

Prova de la paradoxa de Banach-Tarsky

Pel Teorema 4.3.3), és suficient provar el següent lema.

Lema 4.3.1. *[6, Lemma] Sigui G un grup lliure de rang 2. Si G actua lliurement sobre X , aleshores X és G -paradoxal.*

Idea de la Demostració: Per tot $x \in X$, considerem l'òrbita de G

$$Gx = \{gx : g \in G\}.$$

Prenem ara $\mathcal{P}(Gx)$, el conjunt de les parts de les òrbites de G . Aleshores, definim l'àlgebra booleana

$$\mathcal{B} = \prod_{x \in X} \mathcal{P}(Gx),$$

on $\prod_{x \in X}$ és el producte Booleà lliure.

Com hem vist abans, pel Teorema 4.3.1 (és a dir, pel Teorema de Hahn-Banach reformulat en termes d'àlgebres i mesures Booleanes), podem assegurar l'existència d'una mesura booleana μ en $\mathcal{B}$ tal que $\mu(1_{\mathcal{B}}) = 1$. Considerem a continuació les següents particions explícites:

Siguin a_1, a_2 els generadors $\mathbb{F}_2$ i $a_3 = a_1^{-1}, a_4 = a_2^{-1}$. Considerem $A_i \in \mathbb{F}_2$ amb $i = 1, 2, 3, 4$, on A_i és el conjunt de les paraules que comencen amb a_i . Considerem també la permutació $\sigma(1, 2, 3, 4) = (3, 4, 1, 2)$. Aleshores tenim que:

- (i) A_i són disjunts dos a dos,
- (ii) $a_i^{-1} A_i \cup A_{\sigma(i)} = \mathbb{F}_2$.

Definirem ara les següents particions: per a $i = 1, 2, 3, 4$ siguin,

$$X_i = \{x \in X : \mu(xA_i) > 1/2\}.$$

Siguin també

$$Y_1 = X \setminus (X_1 a_1 \cup X_2 a_2),$$

$$Y_2 = X \setminus (X_3 a_3 \cup X_4 a_4).$$

Es comprova directament que tenim les següents particions disjunes

$$X = X_1 a_1 \cup X_2 a_2 \cup Y_1 = X_3 a_3 \cup X_4 a_4 \cup Y_2,$$

i per tant, X és paradoxal.

□

Bibliografia

- [1] Borwein, Jonathan M., and D. Zhuang. *On Fan's minimax theorem*. Mathematical programming 34 (1986): 232-234.
- [2] Eidelman, Yuli, Vitali D. Milman, and Antonis Tzolomitis. *Functional analysis: an introduction*. Vol. 66. American Mathematical Soc., 2004. ISBN: 978-0821836465.
- [3] Foreman, Matthew & Wehrung, Friedrich. *The Hahn-Banach theorem implies the existence of a non-Lebesgue measurable set*. Fund. Math. 138 (1991), no. 1, 13–19.
- [4] Juschenko, Kate. *Lecture 1: Paradoxical decompositions of groups and their actions*. Disponible a: <https://web.ma.utexas.edu/users/juschenko/files/Lecture1.pdf>.
- [5] Luxemburg, W. A. J. *Reduced powers of the real number system and equivalents of the Hahn-Banach extension theorem*. Applications of Model Theory to Algebra, Analysis, and Probability (Internat. Sympos., Pasadena, Calif., 1967), pp. 123–137. Holt, Rinehart and Winston, New York-Montreal, Que.-London, 1969. ISBN: 978-0037158653.
- [6] Pawlikowski, Janusz. *The Hahn-Banach theorem implies the Banach-Tarski paradox*. Fundamenta Mathematicae 138.1 (1991), 21-22.
- [7] Peng, Robert. *The Hahn-Banach separation theorem and other separation results*. Participant paper from a REU programme, <http://math.uchicago.edu/may/REU2014/REUPapers/Peng.pdf> (2014).
- [8] Rudin, Walter. *Análisis funcional*. Reverté, 2012. ISBN: 978-8429145178.
- [9] Saxe, Karen. *Beginning Functional Analysis*. Undergrad. Texts Math. Springer-Verlag, New York, 2002. ISBN: 978-0387952246.
- [10] Tremblay, Madeline. *The Banach-Tarski Paradox*. (2017). Disponible a: <https://www2.math.uconn.edu/~solomon/BTFinal.pdf>.
- [11] Wagon, Stan. *The Banach-Tarski Paradox*. No. 24. Cambridge University Press, 1993. ISBN: 978-0521457040.

- [12] Wikipedia contributors. *Vitali set*. Wikipedia, The Free Encyclopedia. Disponible a: https://en.wikipedia.org/w/index.php?title=Vitali_set&oldid=1242377808. Accedit el 24 de desembre de 2024.