TREBALL DE FI DE GRAU

Grau en Matemàtiques

Títol: Axioma d'elecció. Equivalències i aplicacions

Autor: Guim Olivé Oller

Director: Jordi Quer Bosor

Departament: Departament de Matemàtiques (749)

Convocatòria: Juny/Juliol 2016

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

M'agradaria agrair a tots els meus col·laboradors del treball tots els esforços fets, en particular per part del director del treball Jordi Quer Bosor per la seva paciència i la seva excel·lent capacitat per comprendre les dificultats i les limitacions que comporten fer recerca i el professor Josep Maria Brunat per dirigir-me el primer esbós del treball abans de la seva jubilació i pel seu èmfasi en ensenyar-nos a escriure correctament a Latex i articles de matemàtiques en general. També agraeixo la col·laboració del professor jubilat expert en la matèria Josep Pla i Carrera per informar-me de les referències més adequades per al treball. També a la família i als companys que m'han sapigut animar i donar suport per realitzar aquest projecte i al professor Oriol Valentín, que a través de converses m'ha donat molt de bagatge en lògica matemàtica.

Resum

Paraules clau: Axioma de l'elecció, lema de Zorn, Zermelo-Frankael, teoria de conjunts

MSC2000: 03-XX

En primer lloc, al Capítol 1 s'estableixen definicions bàsiques i preliminars, alguns d'ells ja vists al llarg del Grau de Matemàtiques -alguns d'ells els ometem- i d'altres que poden resultar nous, on es presenten algunes de les definicions bàsiques de teoria de conjunts necessàries per postular l'axioma de l'elecció i el lema de Zorn incloent la llista d'axiomes de Zermelo-Frankael. He donat per sabuts coneixements de teoria dels nombres ordinals, tenint en compte que el seu desenvolupament és molt llarg i tediós i és difícil de presentar de manera informal donant només quatre pinzellades; alguns resultats que s'utilitzen els vam veure a l'assignatura optativa de Lògica i Fonamentació. En canvi, he inclòs teoria de nombres cardinals (sense posar-me fer teoria de classes i considerar *classes* de cardinals), atès que per demostrar la llei de tricotomia es fa servir l'axioma de l'elecció. He decidit col·locar un lema previ molt tediós per demostrar que l'axioma de l'elecció implica el lema de Zorn (el lema de Bourbaki) i conceptes previs de teoria de l'ordre en aquesta secció per tal que la secció següent quedés més polida.

He separat en dos capítols diferents (Capítol 2 i Capítol 3) aquelles equivalències conjuntístiques que considero més *bàsiques* d'aquelles que no ho són tan al tenir un flaire més pertinent a altres branques de les matemàtiques. El Capítol 2 inclou l'equivalència entre l'axioma de l'elecció, el lema de Zorn, el principi de la bona ordenació, l'axioma de l'elecció, el principi de maximalitat de Hausdorff i el lema de Tukey. El capítol 3 inclou la seva equivalència amb l'axioma de les eleccions múltiples, el teorema d'Steinitz, el teorema de Tychonov, l'existència d'ideals maximals i el teorema de Tarski, algun d'ells amb demostracions de recíprocs i en d'altres ometent-les.

Al Capítol 4 demostrarem alguns teoremes que es dedueixen de l'axioma de l'elecció: l'existència d'inversa per la dreta d'una aplicació exhaustiva (teoria de conjunts), els teoremes de Banach-Tarski (anàlisi funcional), l'existència de conjunts no mesurables (teoria de la mesura), la paradoxa de Banach-Tarski (teoria de la mesura) i l'existència de clausura algebraica d'un cos (teoria de cossos).

Al Capítol 5 demostrarem alguna versió feble de l'axioma de l'elecció, en particular l'axioma de les eleccions dependents, que l'aplicarem per demostrar el teorema de la base de Hilbert.

Al Capítol 6 s'incorpora un apèndix que conté diverses aplicacions del teorema de Hahn-Banach a l'anàlisi funcional (en particular, versions geomètriques) i una demostració alternativa de la paradoxa de Banach-Tarski, juntament amb alguna generalització seva.

Notació

$\mathbb{N}$	Nombres naturals
$\mathbb{Z}$	Nombres enters
$\mathbb{Q}$	Nombres racionals
$\mathbb{R}$	Nombres reals
$\mathbb{C}$	Nombres complexos

Índex general

1. Introducció 1

Capítol 1. Preliminars 3
1. Definicions bàsiques conjuntístiques 3
2. Axiomes de Zermelo-Fraenkel 4
3. Relacions d'ordre 5
4. Lema de Bourbaki 7
5. Conjunts ben ordenats i teorema de bona ordenació 9
6. Cardinals 9

Capítol 2. Axioma de l'elecció (equivalències bàsiques) 15
1. Caracterització de l'axioma de l'elecció i observacions sobre el seu ús 15
2. Axioma de l'elecció, Lema de Zorn i principi de bona ordenació 16
3. Axioma de l'elecció, Principi de maximalitat de Hausdorff i lema de Tukey 18

Capítol 3. Resultats equivalents 21
1. Equivalències bàsiques addicionals de l'axioma de l'elecció 21
2. Bases d'espais vectorials 22
3. Teorema de Tychonov 27
4. Existència d'ideals maximals 29
5. Teorema de Tarski 30

Capítol 4. Aplicacions de l'axioma de l'elecció 31
1. Inversa per la dreta d'una aplicació exhaustiva 31
2. Teoremes de Hahn-Banach 32
3. Existència de conjunts no mesurables 35
4. Paradoxa de Banach-Tarski 37
5. Existència de clausura algebraica 41

Capítol 5. Versions febles de l'Axioma de l'elecció i algunes aplicacions 43
1. Axioma de les eleccions dependents. Aplicació al teorema de la base de Hilbert 43

Capítol 6. Apèndix 47
1. Conseqüències en anàlisi funcional del teorema de Hahn-Banach: versions geomètriques 47
2. Demostració alternativa de la paradoxa de Banach-Tarski (més moderna) 55

Bibliografia 61

1. Introducció

Objectiu

L'objectiu d'aquest treball és establir alguns teoremes equivalents a l'axioma de l'elecció i enunciar i demostrar algunes de les seves aplicacions.

Motivacions

Una de les motivacions que m'ha portat a investigar sobre l'axioma de l'elecció és el fet sorprenent que descobreixes al principi dels estudis, en aquells moments que necessites establir amb més o menys rigor una base axiomàtica que et defineixen les construccions i les regles permeses en un raonament matemàtic a partir de la qual se'n desprenen la resta de resultats, que malgrat la seva solidesa i rigidesa aparent, delimitar *què* és un raonament matemàtic i *què no* ho és té el seu punt de subjectivitat.

Recordem que s'ha demostrat que l'axioma de l'elecció és indecidible a l'axiomàtica de Zermelo-Fraenkel (ZF), és a dir, si ZF fos consistent, aleshores tant ZF amb l'axioma de l'elecció (abreujat ZFC, provenint de Zermelo-Fraenkel *Choice*) com ZF negant l'axioma de l'elecció serien teories consistents (és a dir, l'axioma de l'elecció és *relativament consistent* a ZF i *independent* a ZF). A més a més, l'axioma de l'elecció sempre ha generat controvèrsia entre els matemàtics pels resultats *paradoxals* que es poden obtenir a partir d'ell (com ara la paradoxa de Banach-Tarski, que de manera informal ens diu que podem "clonar esferes" a través d'isometries aplicades a un nombre finit de conjunts que fan una partició de l'original), i encara actualment és rebutjat per alguns matemàtics.

En particular, sense voler entrar en detall en qüestions subjectives d'opinió personal i només per donar-ne quatre pinzellades sense aprofundir en temes complexos de filosofia de les matemàtiques de les quals se n'ha escrit molta literatura i hi segueix existint molt de debat, crec que aquesta suposada *paradoxa* es pot "resoldre" fàcilment o convertir-la en una idea més intuïtiva si es pensa que l'aplicabilitat de les matemàtiques a la vida quotidiana (en aquest cas particular, a la física) no és la totalitat de la disciplina; per exemple, podem pensar que els conjunts no mesurables són objectes matemàtics que, tanmateix, no s'utilitzen a la física -de fet, només treballem amb conjunts borelians- i, per tant, aquestes *clonacions* d'objectes no es poden dur a terme fora del context de les matemàtiques. En general, són qüestions que a mesura que la disciplina avança i s'amplia, especialment després de les transformacions en la concepció de les matemàtiques dels segles XIX i XX, han anat *normalitzant* l'ús d'aquest axioma per la seva vasta aplicabilitat.

Per tant, part de l'interès d'establir algunes equivalències d'aquest axioma roman en el fet que la negació de l'axioma de l'elecció implicaria que aquestes formulacions equivalents de l'axioma, que apareixen em moltes branques de les matemàtiques, serien falses. En canvi, acceptant l'axioma de l'elecció com a cert, aquestes formulacions equivalents també serien certes i, de fet, la seva equivalència ens permet afirmar que es podrien agafar com a postul·lats alternatius de l'axioma. A més a més, també és una eina útil que ens permet escurçar l'axiomàtica de la teoria de conjunts de manera que se'n dedueixen les seves aplicacions o versions més febles a partir d'aquest, sense haver d'incorporar aquests nous resultats a la nova axiomàtica.

Originalment un dels temes del meu interès que m'hagués interessat abordar -i que segueixen sent del meu interès- però que difícilment són abordables sense coneixements d'assignatures més avançades de Lògica matemàtica era la demostració de la seva consistència i indecidibilitat i les tècniques de demostració d'aquest tipus de proposicions, atès que té el seu punt sorprenent respecte als resultats més habituals que s'estableixen en matemàtiques pel fet de tenir conseqüències immediates sobre la pròpia fonamentació de les matemàtiques, com si aquesta disciplina *s'autoalimentés* recursivament en forma fractal a partir dels seus propis fonaments. Tanmateix, atès que es tracta d'un treball del Grau de Matemàtiques on el tipus de raonaments i demostracions segueixen un estil *semiformal* -o el que popularment se'n diuen "matemàtiques per a un *working mathematician*"- i les aplicacions que en fem de l'axioma són a nivell no tan fonamental com en d'altres branques més rigoroses de la lògica matemàtica, hem optat per ometre aquesta part i seguir un estil de presentació lleugerament més informal (demostracions semiformals).

També m'ha semblat interessant veure alguns exemples de versions de l'axioma més febles que (potser) alguns matemàtics reticents a acceptar l'axioma de l'elecció admetrien per tal de no renunciar a la totalitat

de totes les seves conseqüències i veure una mica un petit reticle del seu diagrama d'implicacions. Com que n'hi ha moltes i aquest treball no pot abordar-les totes, m'he limitat a fer-ne una selecció basada en criteris subjectius d'importància o d'interès personal.

M'ha semblat interessant haver pogut treballar amb temes de disciplines tan diverses, atès que és un axioma que s'utilitza en branques molt diverses (àlgebra, topologia, teoria de conjunts, anàlisi funcional, etc).

Continguts

En primer lloc, al Capítol 1 s'estableixen definicions bàsiques i preliminars, alguns d'ells ja vists al llarg del Grau de Matemàtiques -alguns d'ells els ometem- i d'altres que poden resultar nous, on es presenten algunes de les definicions bàsiques de teoria de conjunts necessàries per postular l'axioma de l'elecció i el lema de Zorn incloent la llista d'axiomes de Zermelo-Frankael. He donat per sabuts coneixements de teoria dels nombres ordinals, tenint en compte que el seu desenvolupament és molt llarg i tediós i és difícil de presentar de manera informal donant només quatre pinzellades; alguns resultats que s'utilitzen els vam veure a l'assignatura optativa de Lògica i Fonamentació. En canvi, he inclòs teoria de nombres cardinals (sense posar-me fer teoria de classes i considerar *classes* de cardinals), atès que per demostrar la llei de tricotomia es fa servir l'axioma de l'elecció. He decidit col·locar un lema previ molt tediós per demostrar que l'axioma de l'elecció implica el lema de Zorn (el lema de Bourbaki) i conceptes previs de teoria de l'ordre en aquesta secció per tal que la secció següent quedés més polida.

He separat en dos capítols diferents (Capítol 2 i Capítol 3) aquelles equivalències conjuntístiques que considero més *bàsiques* d'aquelles que no ho són tan al tenir un flaire més pertinent a altres branques de les matemàtiques. El Capítol 2 inclou l'equivalència entre l'axioma de l'elecció, el lema de Zorn, el principi de la bona ordenació, l'axioma de l'elecció, el principi de maximalitat de Hausdorff i el lema de Tukey. El capítol 3 inclou la seva equivalència amb l'axioma de les eleccions múltiples, el teorema d'Steinitz, el teorema de Tychonov, l'existència d'ideals maximals i el teorema de Tarski, algun d'ells amb demostracions de recíprocs i en d'altres ometent-les.

Al Capítol 4 demostrarem alguns teoremes que es dedueixen de l'axioma de l'elecció: l'existència d'inversa per la dreta d'una aplicació exhaustiva (teoria de conjunts), els teoremes de Banach-Tarski (anàlisi funcional), l'existència de conjunts no mesurables (teoria de la mesura), la paradoxa de Banach-Tarski (geometria i teoria de la mesura) i l'existència de clausura algebraica d'un cos (teoria de cossos).

Al Capítol 5 demostrarem alguna versió feble de l'axioma de l'elecció, en particular l'axioma de les eleccions dependents, que l'aplicarem per demostrar el teorema de la base de Hilbert.

Al Capítol 6 s'incorpora un apèndix que conté diverses aplicacions del teorema de Hahn-Banach a l'anàlisi funcional (en particular, versions geomètriques) i una demostració alternativa de la paradoxa de Banach-Tarski, juntament amb alguna generalització seva.

Capítol 1
Preliminars

En aquest capítol s'estableixen definicions bàsiques i preliminars, alguns d'ells ja vists al llarg del Grau de Matemàtiques -alguns d'ells els ometem- i d'altres que poden resultar nous.

1. Definicions bàsiques conjuntístiques

En aquest apartat es presenten algunes de les definicions bàsiques necessàries per postular l'axioma de l'elecció, que més endavant farem servir per establir l'equivalència entre tres de les seves formulacions: famílies de conjunts, particions i productes cartesians amb funcions d'elecció.

Definim, des del punt de vista de la teoria de conjunts, la següent noció:

Definició. Una *parella ordenada* de coordenades a i b és un conjunt de la forma $\{\{a\},\{a,b\}\}$. Aquest conjunt el denotarem per (a,b), i direm que a és la *primera coordenada* i b la *segona coordenada*.

Remarca. Observem que podem reescriure les coordenades a i b de la parella ordenada $z := (a,b)$ en termes únicament de z i les operacions bàsiques de la teoria de conjunts (la reunió i la intersecció).

Per a la primera coordenada, tenim:

$$\cap z = \{a\};$$
$$\cap\cap z = a.$$

Per tant, $a := \cap\cap z$ en termes de z. Per a la segona coordenada, tenim:

$$(\cap\cup z)\cup((\cup\cup z)\setminus(\cup\cap z)) = (\cap\{a,b\})\cup((a\cup b)\setminus a) = (a\cap b)\cup(b\setminus a) = b.$$

Per tant, $b := (\cap\cup z)\cup((\cup\cup z)\setminus(\cup\cap z))$ en termes de z.

Definició. Una *partició* $\mathcal{P}\subseteq\mathcal{P}(A)$ d'un conjunt A és un conjunt $\mathcal{P}$ de I subconjunts no buits de A disjunts dos a dos i tals que la reunió de tots ells és A. Simbòlicament, una partició de A és un conjunt $\mathcal{P}\subseteq\mathcal{P}(A)$ tal que:

(1) $P\neq\emptyset$ per a tot $P\in\mathcal{P}$;
(2) $P_i\cap P_j=\emptyset$ per a tot $i,j\in I$, $i\neq j$;
(3) $A=\bigcup_{P\in\mathcal{P}} P$.

Definició. Sigui I un conjunt d'índexs, i sigui $\{A_i\}_{i\in I}$ una família arbitrària de conjunts. El *producte cartesià* d'aquesta família, que denotarem per $\prod_{i\in I}A_i$ o $\prod\{A_i|\ i\in I\}$, és el conjunt de totes les funcions f amb domini I tals que $f_i := f(i)\in A_i$ per tot $i\in I$. Simbòlicament:

$$\prod_{i\in I}A_i := \left\{ f\colon I\to\bigcup_{i\in I}A_i \mid f(i)\in A_i \text{ per a tot } i\in I\right\}.$$

Diem que f_i és la *coordenada i-èsima* de la funció f.

Diem que una funció $f\colon \{A_i\}_{i\in I} \to \bigcup_{i\in I} A_i$ és una *funció d'elecció* si satisfà que $f(A_i) \in A_i$ per a tot $i \in I$. Qualsevol de les funcions $f \in \prod_{i\in I} A_i$ es pot identificar, a través de la bijecció $h\colon I \to \{A_i\}_{i\in I}$ definida per $h(i) = A_i$, amb una funció d'elecció de la família $\{A_i\}_{i\in I}$. A partir d'ara anomenarem funció d'elecció qualsevol de les dues formulacions.

2. Axiomes de Zermelo-Fraenkel

Els axiomes que se solen agafar per a la teoria de conjunts, que ens donen una idea intuïtiva de què és un conjunt a partir de proposicions que ens diuen *com es comporten* els conjunts (sense donar una definició de conjunt de forma explícita, ja que s'agafa com un concepte bàsic).

Escrivint la llista d'axiomes de Zermelo-Fraenkel-Choice (ZFC) -un dels subconjunts dels axiomes que ens estableixen la teoria, alguns d'ells es poden intercanviar per d'altres de manera que s'obtingui el mateix- tan en el llenguatge de la lògica de primer ordre com en llenguatge col·loquial com al Jech (1997), tenim:

(1) Axioma 1 (Axioma d'extensionalitat):

$$\forall u(u \in X \equiv u \in Y) \Rightarrow X = Y.$$

Si cada element d'un conjunt M és també un element de N i viceversa i, per tant, $M \subseteq N$ i $N \subseteq M$, aleshores $M = N$. Més breument, cada conjunt està determinat pels seus elements.

(2) Axioma 2 (Axioma dels Aparellaments):

$$\forall a\ \forall b\ \exists c\ \forall x\ (x \in c \equiv (x = a \vee x = b)).$$

Per a tot a, b existeix un conjunt $\{a, b\}$ que conté exactament els elements a i b.

(3) Axioma 3 (Axioma dels subconjunts / Axioma de comprensió / Axioma de separació):

$$\forall X\ \forall p\ \exists Y\ \forall u(u \in Y \equiv (u \in X) \wedge \varphi(u, p)).$$

Si φ és una propietat (amb paràmetre p), aleshores per a tot X i p existeix un conjunt $Y = \{u \in X \mid \varphi(u, p)\}$ que conté tots aquells $u \in X$ que satisfan la propietat φ.

(4) Axioma 4 (Axioma de la Unió):

$$\forall X\ \exists Y\ \forall u(u \in Y \equiv \exists z(z \in X \wedge u \in z).$$

Per a tot conjunt X, existeix un conjunt $Y := \cup X$, la unió de tots els elements de X.

(5) Axioma 5 (Axioma del conjunt potència):

$$\forall X\ \exists Y\ \forall u(u \in Y \equiv u \subseteq X).$$

Per a tot X existeix un conjunt $Y := P(X)$, el conjunt format per tots els subconjunts de X.

(6) Axioma 6 (Axioma de l'infinit):

$$\exists S\ (\emptyset \in S[\wedge(\forall x \in S)[x \cup \{u\} \in S]].$$

(7) Axioma 7 (Axioma de substitució):

$$\forall x\ |\forall y\ |\forall z[\varphi(x, y, z) \wedge \varphi(x, z, p) \Rightarrow y = z]$$

$$\Rightarrow \forall X\ \exists Y\ \forall y[y \in Y \equiv (\exists x \in X)|\ \varphi(x, y, p)]$$

Si F és una funció, aleshores per a tot X existeix un conjunt $Y = F[X] = \{F(x) \mid x \in X\}$.

(8) Axioma 8 (Axioma de Fonamentació / Regularitat):

$$\forall S\ [S \neq \emptyset] \Rightarrow (\exists x \in S\ S \cap x = \emptyset.$$

Tot conjunt no buit té un $\in$-element minimal.

(9) Axioma 9 (Axioma de l'elecció): Per a tota família de conjunts no buits existeix una funció d'elecció.

$$\forall x \in a\ \exists A(x, y) \Rightarrow \exists y \forall x \in a \mid A(x, y(x)).$$

El sistema d'axiomes formats per 1-8 corresponen als axiomes de la teoria de conjunts de Zermelo-Fraenkel (sense l'axioma de l'elecció) i es denota per "ZF". El sistema d'axiomes 1-8 llevat de l'axioma de substitució s'anomena teoria de conjunts de Zermelo, anomenada "Z". El conjunt d'axiomes 1-9 corresponent a la teoria de Zermelo-Fraenkel amb l'axioma de l'elecció, i es denota "ZFC".

Hi ha algunes discrepàncies sobre què s'hauria de considerar per teoria de conjunts de Zermelo. Per exemple, el llibre del Mendelson (1997), no inclou ni l'axioma de l'elecció ni l'axioma de fonamentació, mentre que sí que inclou l'axioma de substitució. El llibre de l'Enderton (1997) inclou els axiomes de l'elecció i el de fonamentació, però no inclou el de substitució.

A partir de l'axioma dels subconjunts (3) i l'axioma de l'infinit (6) es pot deduir el que s'anomena l'Axioma del conjunt buit, que diu que existeix un conjunt que no té cap element:

$$\exists X \forall y (!y \in x),$$

utilitzant que $\exists X(X = X)$ i $\emptyset := \{u \mid u \neq u\}$.

Alguns llibres afegeixen l'Axioma del conjunt buit a la seva llista malgrat la seva redundància per remarcar la seva importància. En el cas d'incorporar l'Axioma del conjunt buit, tindrem que l'Axioma dels subconjunts es pot deduir de la resta d'axiomes (veure Keith Devlin).

3. Relacions d'ordre

En aquest apartat es defineixen els conceptes bàsics de teoria de l'ordre necessaris per enunciar el lema de Zorn.

Definició.

Sigui E un conjunt. Una *relació d'ordre (parcial)* sobre el conjunt E, que denotarem per $\leq$, és un subconjunt no buit $\emptyset \neq R \subseteq E \times E$ que satisfà les propietats següents:

- $\leq$ és reflexiva. Simbòlicament:

$$\text{per a tot } a \in E,\ (a, a) \in R.$$

- $\leq$ és antisimètrica. Simbòlicament:

$$\text{per a tot } a, b \in E,\ \text{si } (a, b) \in R \text{ i } (b, a) \in R \text{ aleshores } a = b.$$

- $\leq$ és transitiva. Simbòlicament:

$$\text{per a tot } a, b, c \in E,\ \text{si } (a, b) \in R \text{ i } (b, c) \in R \text{ aleshores } (a, c) \in R.$$

Notarem $a \leq b = aRb := (a, b) \in R$. També definim $a < b := a \leq b$ i $a \neq b$.

Un *conjunt ordenat* és una parella $(E, \leq)$ on E és un conjunt i $\leq$ una relació d'ordre definida a E. Per abús de llenguatge, usualment denotarem un conjunt ordenat $(E, \leq)$ només pel conjunt E.

Remarca. En un conjunt totalment ordenat, qualsevol subconjunt finit té màxim.

Remarca. L'enunciat del lema de Zorn, que es troba a 15, no requereix el fet que les relacions d'ordre parcials necessàriament hagin de satisfer la propietat antisimètrica. Tanmateix, tenint en compte que en la majoria de situacions es treballa amb relacions d'ordre parcials antisimètriques, considerarem aquesta propietat com a intrínseca a la definició.

Definició. Sigui E un conjunt ordenat. Diem que $\leq$ és una relació d'ordre *total* si es compleix la propietat següent: per a tot $x, y \in E$ aleshores $x \leq y$ o $y \leq x$. Informalment, diem que qualsevol parell d'elements de E són *comparables*.

Definició. Sigui E un conjunt ordenat. Aleshores, direm que $\mathcal{C} \subseteq E$ és una *cadena* de E si és un conjunt totalment ordenat amb la relació $\leq$.

Remarca. Es pot comprovar que si E és el conjunt de subgrups d'un grup G i $\mathcal{C} \subseteq E$ és una cadena, aleshores la reunió $\cup_{H\in\mathcal{C}} H$ també és un subgrup de G.

També es pot comprovar que si E és el conjunt d'ideals d'un anell commutatiu A i $\mathcal{C} \subseteq E$ és una cadena, aleshores la reunió $\cup_{I\in\mathcal{C}} I$ també és un ideal de A.

També podem comprovar que si E és el conjunt de subespais vectorials d'un espai vectorial F i $\mathcal{C} \subseteq E$ és una cadena, aleshores la reunió $\cup_{H\in\mathcal{C}} H$ també és un subespai vectorial de F.

Més generalment, hi ha una sèrie de classes d'objectes complint unes certes propietats (que aquí no definirem) que, si les dotem d'algun ordre, "es comporten bé"quan calculem la reunió d'elements d'una cadena d'aquestes classes d'objectes, és a dir, que aquestes reunions també seran elements de les classes satisfent aquestes propietats.

Definicions. Sigui $(E, \leq)$ un conjunt ordenat, $A \subseteq E$ un subconjunt ordenat de $(E, \leq)$ i $x \in E$ un element.

Diem que x és una *fita superior* de A si $x \geq a$ per a tot $a \in A$.

Diem que x és un *màxim* de A si x és un fita superior de A i $x \in A$. Anàlogament, diem que x és un *primer element* o *mínim* de A si x és un fita inferior de A i $x \in A$.

Diem que x és un *suprem* de A si qualsevol fita superior de A és més gran o igual que x. Anàlogament, diem que x és un *ínfim* de A si qualsevol fita inferior de A és més petita o igual que x.

Diem que $a \in A$ és un *element maximal* de A si no existeix cap element $x \in A$ tal que $a < x$.

Remarca. Observem que si $(E, \leq)$ és un conjunt totalment ordenat i $A \subseteq (E, \leq)$ és un subconjunt ordenat, aleshores, si existeix un element $s \in E$ que és suprem de A, aquest és únic. En tal cas, denotarem $s := \sup A$. Anàlogament, si existeix un element $i \in E$ que és ínfim de A, aquest és únic. En tal cas, denotarem $i := \inf A$.

Exemple. Sigui E un conjunt. Considerem $(\mathcal{P}(E), \subseteq) := (\{A \mid A \subseteq E\}, \subseteq)$ el conjunt de les parts de E amb la inclusió. Observem que $\subseteq$ és una relació d'ordre a $\mathcal{P}(E)$:

(1) $\subseteq$ és reflexiva, ja que $A \subseteq A$ per a tot $A \in \mathcal{P}(E)$.
(2) $\subseteq$ és antisimètrica, ja que si A i B subconjunts de $\mathcal{P}(E)$ són tals que $A \subseteq B$ i $B \subseteq A$, aleshores $A = B$.
(3) $\subseteq$ és transitiva, ja que si A, B i C subconjunts de $\mathcal{P}(E)$ són tals que $A \subseteq B$ i $B \subseteq C$, aleshores $A \subseteq C$.

Per tant, hem vist que $\subseteq$ és un ordre parcial. Tanmateix, aquest ordre no és un ordre total, ja que si $|E| \geq 2$ i si a i b són dos elements de E tals que $a \neq b$ (que existeixen per ser E de cardinal més gran o igual que 2), els conjunts $\{a\}$ i $\{b\}$ de $\mathcal{P}(E)$ són tals que $\{a\} \nsubseteq \{b\}$ i $\{b\} \nsubseteq \{a\}$ i, per tant, no comparables.

Si E és numerable, $E = \{a_1, a_2, \ldots, a_n, \ldots\}$, i si ens restringim al subconjunt

$$\mathcal{C} = \{\emptyset, \{a_1\}, \{a_1, a_2\}, \{a_1, a_2, a_3\}, \ldots\}$$

format per una quantitat numerable de conjunts de forma que la seva construcció s'obté per adjunció d'un nou element de E diferent als del conjunt anterior, tenim que $(\mathcal{C}, \subseteq)$ és un conjunt totalment ordenat, ja que $\emptyset \subseteq \{a_1\} \subseteq \{a_1, a_2\} \subseteq \{a_1, a_2, a_3\} \subseteq \ldots$ i, per la transitivitat de la relació d'inclusió, tots els elements de $\mathcal{P}$ són comparables.

Observem que la cadena $\mathcal{C}$ té el conjunt E com a suprem, ja que la reunió de tots els elements és una fita superior (per ser tots els elements de $\mathcal{C}$ subconjunts de E) i, a més a més, és la més petita, ja que si tinguéssim un conjunt estrictament més petit que contingués tots els elements de la cadena, el conjunt format per aquest conjunt afegint-li un nou element de E diferent dels anteriors no hi estaria contingut.

Definicions. Sigui $(E, \leq)$ un conjunt ordenat. Diem que E és *inductiu* si tota cadena de E té fita superior i diem que E és *estrictament inductiu* si tota cadena de E té suprem.

Exemple. Agafem un anell commutatiu i unitari no trivial A, i considerem $(\mathcal{P}, \subseteq)$ el conjunt de tots els ideals de A ordenats amb la relació d'inclusió. Sigui $\mathcal{C} \subseteq \mathcal{P}$ una cadena d'ideals no buida arbitrària i sigui $I_0 := \bigcup \{I \mid I \in \mathcal{C}\}$.

Comprovem que I_0 és un ideal. I_0 és no buit perquè $\mathcal{C}$ és no buida. Si x, $y \in I_0$, aleshores existeixen I i J ideals tals que $x \in I \in \mathcal{C}$ i $y \in J \in \mathcal{C}$. Com que el fet que $\mathcal{C}$ sigui una cadena vol dir que sempre puc comparar els elements de la cadena, tenim que $I \subseteq J$ o $J \subseteq I$. Podem suposar, sense pèrdua de generalitat, que $I \subseteq J$. Aleshores, x,$y \in J$ i, com que J és un ideal, $x + y \in J$, i si és de J també és de la reunió; $x + y \in J \subseteq I_0$. A més a més, si $x \in I \subseteq \mathcal{C}$ i $a \in A$, aleshores $ax \in I \subseteq I_0$. Per tant, I_0 és un ideal.

A més a més, I_0 és una fita superior de $\mathcal{C}$, ja que, per definició de I_0, tots els ideals de $\mathcal{C}$ estan continguts a I_0.

Com que hem vist que aquesta construcció del conjunt I_0 la podem fer per qualsevol cadena $\mathcal{C}$ arbitrària de $\mathcal{P}$, aleshores $(\mathcal{P}, \subseteq)$ és un conjunt ordenat inductiu.

4. Lema de Bourbaki

En aquest apartat es demostren alguns lemes tècnics previs a la demostració del lema de Zorn a partir de l'axioma de l'elecció, com ara el lema de Bourbaki.

Per tal de demostrar el lema de Zorn a partir de l'axioma de l'elecció, necessitarem utilitzar un parell de lemes tècnics. Anem-los a veure:

Lema 1. *Sigui $E \neq \emptyset$ un conjunt. Sigui $(\mathcal{P}(E), \subseteq) = \{A \mid A \subseteq E\}$ el conjunt de les parts de E parcialment ordenat per inclusió. Aleshores $(P(E), \subseteq)$ és estrictament inductiu.*

Demostració: A l'exemple 2.3 hem demostrat que, efectivament, $\subseteq$ defineix una relació d'ordre parcial al conjunt $\mathcal{P}(E)$ i, per tant, la relació d'ordre està ben definida (l'enunciat del lema té sentit).

Sigui $\mathcal{C}$ una cadena i $C_0 = \bigcup \{C \mid C \in \mathcal{C}\}$. Comprovem que és el suprem de $\mathcal{C}$. Per a veure que $C_0 \subseteq E$ només fem servir que la reunió de subconjunts de E és un subconjunt de E i, per tant, pertany a $\mathcal{P}(E)$. Per tal de comprovar que és una fita superior de la cadena $\mathcal{C}$, només cal adonar-se que $C \subseteq C_0$ per a tot $C \in \mathcal{C}$, ja que C_0 és la unió de tots ells. Finalment, per a comprovar que és minimal amb aquesta propietat (i, per tant, és suprem), suposem que existeix un altre conjunt C_1 que conté tots els elements de la cadena $\mathcal{C}$ i $C_1 \subseteq C_0$. Aleshores, com que C_0 també conté tots els elements de la cadena, també es té la inclusió $C_0 \subseteq C_1$, d'on tenim $C_0 = C_1$. Per tant, $\sup \mathcal{C} = C_0$, tal com volíem comprovar.

Per tant, tenim que $(\mathcal{P}(E), \subseteq)$ és estrictament inductiu, tal com volíem veure. □

Lema 2 (de Bourbaki). *Sigui $(X, \leq) \neq \emptyset$ un conjunt parcialment ordenat i tal que tota cadena $\mathcal{T} \subseteq X$ no buida té un suprem a X. Si $f\colon X \to X$ és una funció tal que $x \leq f(x)$ per a tot $x \in X$, aleshores existeix un $x_0 \in X$ tal que $f(x_0) = x_0$.*

Demostració: Definim que un conjunt $E \subseteq X$ no buit és *admissible* si $f(E) \subseteq E$ (és f invariant) i si el suprem de qualsevol cadena no buida a E, que existeix a X per hipòtesi, també pertany a E. Per hipòtesi, X és un subconjunt admissible de X. Si $x \in X$, aleshores la intersecció de tots els subconjunts admissibles que contenen x, que anomenarem A_x, és admissible. Això passa perquè si $\left\{A_x^i \mid i \in I\right\}$ són tots els subconjunts de X admissibles que contenen x, aleshores $A_x := \bigcap_{i \in I} A_x^i$ és tal que $f(A_x) = f(\bigcap_{i \in I} A_x^i) \subseteq \bigcap_{i \in I} f(A_x^i) \subseteq \bigcap_{i \in I} A_x^i = A_x$. Al penúltim pas hem usat que $f(A_x^i) \subseteq A_x$, i tota cadena no buida de A_x té suprem a A_x perquè conté la intersecció de tots els suprems de cadenes de $\left\{A_x^i\right\}_{i \in I}$ per a tot $i \in I$.

Per hipòtesi, $X \neq \emptyset$. Fixem un element $a \in X$ i considerem el conjunt A_a. El primer pas de la demostració consisteix en demostrar que A_a és una cadena. Una vegada demostrat això, podem fer el següent argument: Com que A_a és una cadena, el seu suprem, que denotem per x_0, pertany a X per hipòtesi i, com que A_a és

un subconjunt admissible, aleshores $x_0 \in A_a$. Com que A_a és admissible, $f(A_a) \subseteq A_a$. Per tant, $f(x_0) \in A_a$. Com que x_0 és una fita superior a A_a, $f(x_0) \leq x_0$. Per altra banda, tenim que $x_0 \leq f(x_0)$ per la propietat que tenim com a hipòtesi de f. Per tant, per la propietat antisimètrica de la relació d'ordre parcial, tenim que $f(x_0) = x_0$ i el lema queda demostrat.

Per tal de veure que A_a és una cadena, considerem el subconjunt C format pels elements $x \in A_a$ amb la propietat que existeix una cadena no buida $\emptyset \neq C_x \subseteq A_a$ que conté a i x tal que

- $a \leq y \leq x$ per a tot $y \in C_x$
- $f(C_x \setminus \{x\}) \subseteq C_x$, i
- el suprem de qualsevol subcadena no buida de C_x és de C_x, és a dir, per a tota $C_y \subseteq C_x$ cadena tal que $C_y \neq \emptyset$, aleshores $\sup C_y \in C_x$,

és a dir que $C := \{x \in A_a | \ \exists\ C_x : \emptyset \neq C_x \subseteq A\}$ on els C_x compleixen les propietats anteriors.

Observem que $a \in C$ perquè podem agafar $C_a = \{a\}$ complint les propietats dels elements de C, ja que $a \leq y \leq a$ per a tot $y \in C_a = \{a\}$, $f(C_a \setminus \{a\}) = \emptyset \subseteq C_a = \{a\}$ i el suprem de qualsevol cadena no buida de C_a és a C_a perquè C_a és l'única subcadena no buida que es pot formar, i conté a. Si $x \in C$, de forma que existeix una cadena C_x complint les propietats dels elements de C, podem usar aquestes propietats per veure que

$$A_a = A_x \cup C_x. \tag{1}$$

Tenim que $A_a \supseteq C_x$ per definició, ja que al definir C_x diem que $C_x \subseteq A_a$. A més a més, $A_a \cap A_x$ és un conjunt admissible que conté x i, per tant, conté A_x, per ser A_x la intersecció de tots els conjunts admissibles que contenen x, d'on tenim que $A_a \supseteq A_x$. Per tant, $A_a \supseteq A_x \cup C_x$.

Per demostrar la inclusió inversa, és suficient demostrar que $A_x \cup C_x$ és un subconjunt admissible de X que conté a ja que, en tal cas, la intersecció A_a de tots els conjunts admissibles que contenen a estarà contingut al conjunt admissible $A_x \cup C_x$ concret. L'element $a \in C_x$ i, per tant, $a \in A_x \cup C_x$. Per veure que és admissible, cal veure que $f(A_x \cup C_x) \subseteq A_x \cup C_x$ i que el suprem de qualsevol cadena no buida de $A_x \cup C_x$ pertany a $A_x \cup C_x$. Com que $x \in A_x$, $A_x \cup C_x = A_x \cup (C_x \setminus \{x\})$ i $f(A_x \cup C_x) = f(A_x) \cup f(C_x \setminus \{x\}) \subseteq A_x \cup C_x$, per ser A_x un conjunt admissible i $f(C_x \setminus \{x\}) \subseteq C_x$ per construcció dels elements $x \in C$ (segon quadradet).

Per veure que el suprem de qualsevol cadena no buida de $A_x \cup C_x$ pertany a $A_x \cup C_x$, agafem una cadena no buida $T \subseteq A_x \cup C_x$ i denotem per $u \in X$ el seu suprem (que pertany a X per hipòtesi del lema). Com que A_a és admissible, $u \in A_a$ per definició d'admissibilitat. Observem que

$$y \leq x \text{ i } x \leq z \quad \text{sempre que } y \in C_x \text{ i } z \in A_x. \tag{2}$$

La primera desigualtat prové de la construcció dels elements x del conjunt C, mentre que la segona prové del fet que el conjunt de tots els $z \in X$ tals que $x \leq z$ és un subconjunt admissible de X que conté x i, per tant, intersecant-lo amb els altres subconjunts admissibles, $x \leq z$ per a tot $z \in A_x$.

Si la cadena T té suprem a A_x, aleshores (2) implica que $x \leq u$ i, per tant, el conjunt d'elements de la cadena que pertanyen a A_x forma una cadena a A_x amb suprem u, ja que per (2) i la transitivitat de la relació d'ordre els elements de C_x són més petits o iguals que els de A_x i, per tant, el suprem de les cadenes T i $A_x \cap T$ coincideixen. Com que A_x és admissible, $u \in A_x$. Altrament, la cadena té tots els seus elements a C_x, i aleshores $u \in C_x$ per la tercera de les propietats dels elements $x \in C$.

Amb tot això, acabem de demostrar (1). Encara que no ho necessitem, observem que (1) i (2) ens dóna $A_x \cap C_x = \{x\}$ sempre que C_x existeixi. En aquest cas, $C_x = (A_a - A_x) \cup \{x\}$ és un conjunt completament determinat per les propietats exigides en la seva construcció.

Tornant al fet que C és el subconjunt dels elements de A_a tals que C_x existeix, anem a veure que C és un conjunt admissible que conté a. En cas de ser cert, es dedueix que $C \supseteq A_a$ i, com que per construcció $C \subseteq A_a$, tenim que $C = A_a$. Aquest fet, combinat amb (1) i (2), demostra que A_a és una cadena: si x i y són de A_a, (1) ens diu que $y \in A_x$ o $y \in C_x$, i (2) ens diu que $x \leq y$ en el primer cas i $y \leq x$ en el segon cas,

que coincideix amb la definició de ser cadena (pel fet de tenir una relació d'ordre total). Per tant, quedaria demostrat el teorema. Per tant, la demostració acabarà quan hàgim vist que C és admissible i conté a. Hem vist anteriorment que conté a. Ara ens cal veure que $f(C) \subseteq C$ i que el suprem de qualsevol cadena no buida de C pertany a C. Per la primera part, anem a veure que si $x \in C$, aleshores $C_{f(x)}$ existeix i es pot prendre de forma que $C_{f(x)} = C_x \cup \{f(x)\}$. La propietat (1) ens demostra que $C_x \cup \{f(x)\}$ satisfà la propietat $a \leq y \leq f(x)$ per a tot $y \in C_x \cup \{f(x)\}$ per ser $f(x) \in A_{f(x)}$ i el fet que $x \leq f(x)$, i la propietat que $f(C_x) \subseteq C_x \cup \{f(x)\}$ es dedueix del fet que per C_x es compleix la propietat $f(C_x \setminus \{x\}) \subseteq C_x$ i, afegint x al domini, tenim $f(C_x) \subseteq C_x \cup \{f(x)\}$. Qualsevol cadena no buida de $C_x \cup f(x)$ o està continguda a C_x i, en aquest cas, té la seva fita superior a x, o conté $f(x)$ com a element i té $f(x)$ com a fita superior (ja que $y \leq x \leq f(x)$ per a tot $y \in C_x$ i, per tant, $f(x)$ és el màxim de $C_x \cup \{f(x)\}$). Per tant, efectivament, podem prendre que $C_{f(x)}$ sigui $C_x \cup f(x)$.

Per acabar, sigui $X_\alpha \subseteq C$ una cadena no buida qualsevol, i sigui u el seu suprem, que necessàriament pertany a A_a. Construïm el conjunt $(\bigcup_\alpha C_{x_\alpha}) \cup \{u\}$. Observem que aquest conjunt satisfà les tres propietats de C_u. Per veure la primera propietat, veiem que $a \in (\bigcup_\alpha C_{x_\alpha}) \cup \{u\}$ perquè $a \in C_{x_\alpha}$ per a tot α, $a \leq y$ perquè $a \leq u$ i $a \leq y$ per a tot α i per a tot $y \in C_{x_\alpha}$, per ser $a \leq x_\alpha \leq y$ i $y \leq u$ per a tot $y \in (\bigcup_\alpha C_{x_\alpha}) \cup \{u\})$ per ser u el seu suprem. Per la segona propietat, $f((\bigcup_\alpha C_{x_\alpha}) \cup \{u\} \setminus \{u\}) = f(\bigcup_\alpha C_{x_\alpha}) = \bigcup_\alpha f(C_{x_\alpha}) \subseteq \bigcup_\alpha C_\alpha \subseteq (\bigcup_\alpha C_\alpha) \cup \{u\}$, on al penúltim pas hem usat la segona propietat dels C_{x_α}. Per a la tercera propietat, si T és una cadena de $(\bigcup_\alpha C_{x_\alpha}) \cup \{u\}$, aleshores pot ser que aquesta cadena tingui l'element u i, per tant, u sigui el seu suprem que pertany a $(\bigcup_\alpha C_{x_\alpha}) \cup \{u\}$ o bé tots els seus elements siguin de C_{x_α} i, com que per la tercera propietat de cadascun d'ells qualsevol cadena continguda conté el seu suprem, aleshores el suprem de T també pertanyerà a $\bigcup_\alpha C_{x_\alpha} \subseteq (\bigcup_\alpha C_{x_\alpha}) \cup \{u\}$. Per tant, podem prendre $C_u = (\bigcup_\alpha C_{x_\alpha}) \cup \{u\}$. Per tant, $u \in C$, d'on concloem que C conté el suprem de qualsevol cadena de C. Per tant, C és un conjunt admissible que conté a. □

5. Conjunts ben ordenats i teorema de bona ordenació

A continuació, presentem la definició de conjunt ben ordenat i enunciem el teorema de bona ordenació (també anomenat teorema de Zermelo), que a l'apartat següent demostrarem que també és equivalent a l'axioma de l'elecció i al lema de Zorn.

Remarca.

Recordem que si $(E, \leq) \neq \emptyset$ és un conjunt totalment ordenat, $x_0 \in E$ és un *primer element* de E si per a tot $x \in E$ es té $x_0 \leq x$.

Definició.

Sigui $(E, \leq)$ un conjunt parcialment ordenat. Diem que la relació d'ordre $\leq$ és un *bon ordre* si tot subconjunt no buit $S \subseteq E$ té un primer element. En tal cas, direm que $(E, \leq)$ és un conjunt *ben ordenat*. Si per a un conjunt E podem trobar una relació $\leq$ que satisfà que $(E, \leq)$ és un conjunt ben ordenat, direm que E *admet un bon ordre* o que *es pot ben ordenar*.

Enunciat 3 (teorema de bona ordenació). *Sigui E un conjunt. Aleshores existeix una relació d'ordre $\leq$ tal que $(E, \leq)$ és un conjunt ben ordenat.*

En altres paraules, tot conjunt es pot ben ordenar.

6. Cardinals

En aquest apartat, enunciarem i demostrarem algunes propietats dels nombres cardinals (on, en alguna de les demostracions, farà falta utilitzar l'axioma de l'elecció) que ens permetran demostrar la llei de tricotomia dels nombres cardinals i, al capítol següent, que dues bases d'un espai vectorial no trivial tenen el mateix

cardinal. També afegirem algun lema sobre cardinals que necessitarem a la demostració de l'existència de clausura algebraica d'un cos.

6.1. Definició i resultats de cardinals. Anem a definir nocions bàsiques de cardinalitat i alguns teoremes necessaris per a la proposició 25:

Definició. Diem que dos conjunts A i B són *equipotents* si existeix una funció $f\colon A \to B$ bijectiva. En aquest cas, notarem $A \simeq B$.

Remarca. La relació $\simeq$ és una *relació d'equivalència*, és a dir, compleix les propietats:

(1) $A \simeq A$ (propietat *reflexiva*);
(2) $A \simeq B \Rightarrow B \simeq A$ (propietat *simètrica*);
(3) $A \simeq B$ i $B \simeq C \Rightarrow A \simeq C$ (propietat *transitiva*).

La relació està ben definida, ja que si $A \simeq A'$ amb $b_1\colon A \to A'$ bijectiva, $B \simeq B'$ amb $b_2\colon B \to B'$ bijectiva i $f_1\colon A \to B$ és bijectiva, aleshores $b_2 \circ f_1 \circ b_1^{-1}\colon A' \to B'$ també és bijectiva i $A' \simeq B'$.

La reflexivitat es dedueix d'agafar aplicació $f = id_A$, la propietat simètrica d'agafar l'aplicació inversa de la bijecció entre A i B (que també és bijectiva) i la transitivitat d'agafar la composició de les aplicacions bijectives de A a B i B a C (que també és bijectiva, perquè la composició d'aplicacions bijectives és bijectiva).

A partir d'aquesta relació d'equivalència, es defineix la noció de *cardinal*. En aquesta breu presentació dels nombres cardinals no s'entrarà en detalls tècnics sobre la seva construcció axiomàtica. De fet, en alguns moments s'usarà algun concepte que pot induir a paradoxes com ara la paradoxa de Russell, com ara el *conjunt de tots els conjunts equipotents a* A amb A un conjunt donat o considerar el *conjunt de tots els nombres cardinals* per poder-los dotar d'una relació d'ordre. Aquestes paradoxes conjuntistes es poden evitar fent servir la distinció entre el concepte de *classe* i el de *conjunt*. Així doncs, faig servir el terme *conjunt* per referir-me a col·leccions d'objectes que pot ser que siguin classes en comptes de conjunts i que caldria reemplaçar convenientment, així com els termes *funció* i *relació* correspondrien a conceptes similars per a classes.

Definició. Sigui A un conjunt. Definim *cardinal de* A com la classe d'equivalència $\overline{\overline{A}} := \{B \text{ conjunts} \mid A \simeq B\}$ formada per tots els conjunts que són equipotents a A.

Exemple. Per a comparar els cardinals de $\mathbb{Q}$ i $\mathbb{R}$, la prova de Cantor de la no numerabilitat de $\mathbb{R}$ ens assegura que el cardinal de $\mathbb{Q}$ és diferent del cardinal de $\mathbb{R}$, ja que $\mathbb{Q} \simeq \mathbb{N}$ mentre que $\mathbb{R} \not\simeq \mathbb{N}$ i, per transitivitat, $\mathbb{R} \not\simeq \mathbb{N}$. Però per a poder-los comparar necessitem definir una relació d'ordre entre els dos cardinals.

Definició. Per a qualsevol parell de cardinals $\overline{\overline{A}}, \overline{\overline{B}}$, definim la relació $\leq$ de la forma següent: $\overline{\overline{A}} \leq \overline{\overline{B}}$ si i només si existeix una aplicació injectiva $f\colon A \to B$. Observem que en aquest cas f ens dóna una bijecció $f\colon A \to f(A) \subseteq B$. Direm que $\overline{\overline{A}} < \overline{\overline{B}}$ si i només si $\overline{\overline{A}} \leq \overline{\overline{B}}$ i $\overline{\overline{A}} \neq \overline{\overline{B}}$

A partir d'ara, sempre que comparem dos cardinals amb la relació $\leq$ ens referirem a aquesta relació en concret.

Remarca. Observem que la relació $\leq$ està ben definida per a qualsevol parell de cardinals $\overline{\overline{A}}$ i $\overline{\overline{B}}$. Per fer-ho, podem comprovar que no depèn dels representants escollits de les classes $\overline{\overline{A}}$ i $\overline{\overline{B}}$. Siguin $A \simeq A'$ i $B \simeq B'$ dos parells de representants de cada classe. Sigui g_1 una aplicació bijectiva $g_1\colon A \to A'$ i g_2 una aplicació bijectiva $g_2\colon B \to B'$ (que existeixen per definició de $\simeq$). Aleshores, si $\overline{\overline{A}} \leq \overline{\overline{B}}$, per definició de $\leq$ es té que existeix una aplicació $f\colon A \to B$ injectiva. Aleshores, l'aplicació $g_2 \circ f \circ g_1^{-1}\colon A' \to B'$ és injectiva (per ser composició d'aplicacions injectives). Per tant, la relació $\leq$ només depèn de les classes, i està ben definida.

Observem que $\leq$ és una relació reflexiva, ja que donat un cardinal $\overline{\overline{A}}$ es té que l'aplicació $id_A\colon A \to A$ és una aplicació bijectiva (i, en particular, és injectiva) i, per tant, $\overline{\overline{A}} \leq \overline{\overline{A}}$.

També tenim que $\leq$ és transitiva. Si $\overline{\overline{A}} \leq \overline{\overline{B}}$ i $\overline{\overline{B}} \leq \overline{\overline{C}}$, aleshores existeixen aplicacions injectives $f\colon A \to B$ i $g\colon B \to C$ i, com que $g \circ f$ és composició de funcions injectives i, per tant, injectiva, tenim que $\overline{\overline{A}} \leq \overline{\overline{C}}$.

Si $\leq$ fos una relació d'ordre parcial, caldria veure que és una relació antisimètrica, que és justament l'enunciat del teorema de Schröder-Bernstein que demostrem a continuació:

Proposició 4 (teorema de Schröder-Bernstein). *Siguin $\overline{\overline{A}}$ i $\overline{\overline{B}}$ dos cardinals. Aleshores si $\overline{\overline{A}} \leq \overline{\overline{B}}$ i $\overline{\overline{B}} \leq \overline{\overline{A}}$ es té que $\overline{\overline{A}} = \overline{\overline{B}}$. En altres paraules, la relació $\leq$ és antisimètrica.*

Demostració:

Per hipòtesi, existeixen $f\colon A \to B$ i $g\colon B \to A$ injectives. Si f és bijectiva o g és bijectiva, aleshores $\overline{\overline{A}} = \overline{\overline{B}}$ i ja hem acabat.

Suposem, ara, que ni f ni g són bijectives. Aleshores es té que $f(A) \neq B$ i $g(B) \neq A$. Defineixo l'aplicació $\phi\colon \mathcal{P}(A) \to \mathcal{P}(A)$ tal que $\phi(U) = A \setminus (g(B \setminus f(U)))$ per a tot $U \in \mathcal{P}(A)$. Podem comprovar que l'aplicació ϕ és creixent. En efecte, si $U \subseteq V$ amb $U, V \in \mathcal{P}(A)$, tenim:

$$f(U) = f(V) \Rightarrow B \setminus f(V) \subseteq B \setminus f(U) \Rightarrow g(B \setminus f(V)) \subseteq g(B \setminus f(U)) \Rightarrow$$
$$A \setminus (g(B \setminus f(U)) \subseteq A \setminus (g(B \setminus f(V)) \Rightarrow \phi(U) \subseteq \phi(V).$$

Defineixo la família de conjunts $\mathcal{D} := \{U \subseteq A \mid U \subseteq \phi(U)\}$. Observem que $\mathcal{D} \neq \emptyset$, ja que, en particular, $\emptyset \in \mathcal{D}$ per tenir-se $\emptyset \subseteq \phi(U)$ i $\emptyset \subseteq A$ (el conjunt buit és subconjunt de qualsevol conjunt).

Defineixo $D := \bigcup_{U \in \mathcal{D}} U$. Anem a veure que $D = \phi(D)$:

- $D \subseteq \phi(D)$. Per a tot $U \in \mathcal{D}$, es tenen les inclusions $U \subseteq \phi(U)$ i $U \subseteq D$. Aleshores, com que ϕ és una aplicació creixent i obtenim que $\phi(U) \subseteq \phi(D)$, també s'obté que $U \subseteq \phi(U) \subseteq \phi(D)$ per a tot $U \in \mathcal{D}$. Llavors, $D = \bigcup_{U \in \mathcal{D}} U \subseteq \bigcup_{U \in \mathcal{D}} \phi(U) \subseteq \phi(D)$.
- $D \supseteq \phi(D)$. Com que ϕ és creixent i es compleix $D \subseteq \phi(D)$, també es compleix $\phi(D) \subseteq \phi(\phi(D))$, d'on per definició de $\mathcal{D}$ obtenim que $\phi(D) \in \mathcal{D}$. Per tant, tenim la inclusió $\phi(D) \subseteq \bigcup_{U \in \mathcal{D}} U = D$.

Definim l'aplicació a trossos $h\colon A \to B$ tal que $h(x) = f(x)$ si $x \in D$ i $h(x) = g^{-1}(x)$ si $x \notin D = \phi(D)$ per a tot $x \in A$. Observem que $x \notin D$ és equivalent a dir que $x \in A \setminus \phi(D) = A \setminus (A \setminus (g(B \setminus f(D)))) = g(B \setminus f(D))$.

Observem que h està ben definida, ja que f té A com a domini i g^{-1} està definida a $g(B \setminus f(D))$. Observem que h és injectiva perquè f és injectiva i g^{-1} és injectiva a $g(B \setminus f(D))$ (per ser g una aplicació ben definida). Observem, també, que h és exhaustiva. La funció f és exhaustiva sobre D a $f(D)$ (perquè tota funció és exhaustiva a la seva imatge) i g^{-1} és exhaustiva sobre $g(B \setminus f(D))$ perquè envia tots els elements de $g(B \setminus f(D))$ a $B \setminus f(D)$ (i també és exhaustiva perquè aquesta és la seva imatge). Per tant, com que $B = f(D) \cup (B \setminus f(D))$, veiem que efectivament h és exhaustiva. Per tant, h és bijectiva i tenim $\overline{\overline{A}} = \overline{\overline{B}}$. □

Observem que en la demostració anterior no s'utilitza l'axioma de l'elecció.

Corol·lari 5. *La relació $\leq$ sobre els nombres cardinals és una relació d'ordre.*

Anem a veure una generalització del teorema que ens permet afirmar que el cardinal de $\mathbb{R}$ (que es pot demostrar que és un conjunt equipotent a $\mathcal{P}(\mathbb{N})$) és estrictament major que el cardinal de $\mathbb{N}$.

Proposició 6. *Sigui A un conjunt i $\mathcal{P}(A)$ el seu conjunt potència (el conjunt format per tots els subconjunts de A). Aleshores $\overline{\overline{A}} < \overline{\overline{\mathcal{P}(A)}}$.*

Demostració:

Suposem per reducció a l'absurd que $\overline{\overline{A}} = \overline{\overline{\mathcal{P}(A)}}$ i arribem a contradicció. En aquest cas, existeix una funció $g\colon A \to \mathcal{P}(A)$ bijectiva i, en particular, exhaustiva. Defineixo el conjunt $B := \{y \in A \mid y \notin g(y)\}$. Per construcció, $B \subseteq A$. Com que g és exhaustiva, existeix $z \in A$ tal que $g(z) = B \subseteq A$. Si $z \in B = g(z)$,

aleshores $z \notin g(z) = B$ i si $z \notin B = g(z)$, aleshores $z \in g(z) = B$. Per tant, el conjunt B està mal definit, i arribem a una contradicció que prové de suposar que $\overline{\overline{A}} = \overline{\overline{\mathcal{P}(A)}}$. □

Definició. Direm que un conjunt A és *enumerable* si $\overline{\overline{A}} = \overline{\overline{\mathbb{N}}} := \aleph_0$. Diem que A és *numerable* si és finit o enumerable.

Remarca. Observem que la reunió d'una quantitat numerable de conjunts numerables és numerable. Si tots ells són finits i fem una reunió d'un nombre finit de conjunts, aleshores aquesta reunió és finita. Si un dels conjunts d'una família numerable de conjunts és enumerable, aleshores la seva reunió és enumerable (i, en particular, també és numerable).

LEMA 7. *Tot conjunt infinit té un subconjunt enumerable.*

Demostració: Sigui E un conjunt infinit. Sigui $a_1 \in E$. Aleshores, per ser E infinit, $E \setminus \{a_1\} \neq \emptyset$. Sigui $a_2 \in E \setminus \{a_1\}$. Repetint el procés, existeix $a_3 \in E \setminus \{a_1, a_2\} \neq \emptyset$. En general, definits $a_1, \ldots, a_n \in E$ elements diferents, tenim que $E \setminus \{a_1, \ldots, a_n\} \neq \emptyset$. Aleshores existeix $a_{n+1} \in E \setminus \{a_1, \ldots, a_n\}$. Per inducció sobre $\mathbb{N}$, tenim que la successió $(a_n)_{n\in\mathbb{N}} \subseteq E$ és un subconjunt enumerable de E. □

COROL·LARI 8. *Si A és infinit, aleshores $\aleph_0 \leq \overline{\overline{A}}$*

Demostració: Si B és un subconjunt enumerable (de cardinal $\aleph_0$) que podem trobar a A pel lema 7, aleshores l'aplicació inclusió $i\colon B \to A$ és injectiva i, per tant, $\aleph_0 \leq \overline{\overline{A}}$. □

LEMA 9. *Tot conjunt A enumerable admet una partició en dos conjunts enumerables.*

Demostració: Si A és enumerable, aleshores existeix $f\colon \mathbb{N} \to A$ bijectiva (per definició d'enumerabilitat). Considero la partició $\mathcal{P}_1$ de $\mathbb{N}$ definida com $\mathcal{P}_1 := \{A, B\}$ on $A := \{n \in \mathbb{N} \mid n \text{ és parell}\}$ i $B := \{n \in \mathbb{N} \mid n \text{ és senar}\}$ (efectivament és una partició perquè A i B són no buits, recobreixen $\mathbb{N}$ i són disjunts). Aleshores, $P_2 := \{f(A), f(B)\}$ és una partició de A (com que f és bijectiva, les imatges dels conjunts de $\mathcal{P}_1$ són no buides, recobreixen A i són disjuntes). □

PROPOSICIÓ 10. *Tot conjunt infinit admet una partició en conjunts enumerables.*

Demostració: Sigui A un conjunt infinit. Defineixo la família

$$\mathcal{S} := \{\mathcal{P} \mid \mathcal{P} \text{ és una partició d'un subconjunt de } A \text{ en parts enumerables}\}.$$

Aleshores $\mathcal{S} \neq \emptyset$, ja que pel lema 7 existeix un subconjunt enumerable B de A, i aquest subconjunt admet particions enumerables (la partició trivial si considero $\mathcal{P} = \{A\}$ o una altra no trivial si utilitzo el lema 9).

A partir d'ara, considero la família de subconjunts $(\mathcal{S}, \subseteq)$ ordenada per inclusió.

Anem a veure que $(\mathcal{S}, \subseteq)$ és un conjunt inductiu amb aquesta relació d'ordre. Per definició, hem de veure que tota cadena $\mathcal{C} \subseteq \mathcal{S}$ té una fita superior. En efecte, si $\mathcal{C} := \{\mathcal{P}_k \mid k \in K\}$ és una cadena de $\mathcal{S}$ (que podem suposar no buida perquè $\mathcal{S}$ és no buit), aleshores puc definir la col·lecció $\mathcal{P} := \bigcup_{k\in K} \mathcal{P}_k$, és a dir, el conjunt format per la reunió de totes les particions $\mathcal{P}_k \subseteq \mathcal{C}$.

Anem a veure que $\mathcal{P}$ és una partició: Com que $\mathcal{C} \subseteq \mathcal{S}$ és no buida, $\mathcal{P}$ també és no buit. Cada $U \in \mathcal{P}$ pertany a algun $\mathcal{P}_k$ i, per tant, és no buit i enumerable. Si $U_i \in \mathcal{P}_i$ i $U_j \in \mathcal{P}_j$ són dos elements diferents de $\mathcal{P}$, aleshores, com que per ser $\mathcal{C}$ una cadena es té que $\mathcal{P}_i \subseteq \mathcal{P}_j$ o $\mathcal{P}_j \subseteq \mathcal{P}_i$, han de pertànyer a una mateixa partició i, conseqüentment, han ser disjunts. Per tant, $\mathcal{P}$ és una partició del conjunt $\bigcup \{U \mid U \in \mathcal{P}_k \text{ per a algun } k \in K\} \subseteq A$ i, per definició de $\mathcal{S}$, es té $\mathcal{P} \in \mathcal{S}$.

A més a més, el conjunt $\mathcal{P}$ és una fita superior de la cadena $\mathcal{C}$, ja que totes les particions $P_k \in \mathcal{C}$ estan contingudes a la reunió $\mathcal{P}$ (en altres paraules, $\mathcal{P}$ és *més fina* que les particions $\mathcal{P}_k \in \mathcal{C}$) i $\mathcal{P} \in \mathcal{S}$. Aleshores, com que $(\mathcal{S}, \subseteq)$ és un conjunt inductiu, pel lema de Zorn sabem que té un element maximal $\mathcal{M}$.

Defineixo $F := \bigcup_{U\in\mathcal{M}} U$ la reunió de tots els conjunts de $\mathcal{M}$. Si el conjunt $A\setminus F$ fos infinit, aleshores el lema 7 ens permetria afirmar que existeix $Z \subseteq A \setminus F$ un subconjunt enumerable. Aleshores el conjunt $\mathcal{M} \cup \{Z\}$ seria una partició d'un subconjunt de A en parts enumerables tal que $\mathcal{M} \subsetneq \mathcal{M} \cup \{Z\}$, ja que per a tot $U \in \mathcal{M}$ tenim $Z \cap U \subseteq (A \setminus F) \cap U = \emptyset$. Per tant, arribaríem a una contradicció amb la maximalitat de $\mathcal{M}$. Per tant, $A \setminus F$ és finit.

Hem vist abans que $\mathcal{S}$ és no buit i, per tant, $\mathcal{M}$ té algun element T. Aleshores el conjunt $T' := T \cup (A \setminus F)$ és enumerable (ja que la reunió d'un conjunt enumerable i un conjunt finit és enumerable per 6.1), i $(\mathcal{M} \setminus \{T\}) \cup \{T'\}$ és la partició de A en conjunts enumerables disjunts que buscàvem. □

Lema 11. *Sigui S un conjunt infinit i E un conjunt. Suposem que per a tot $s \in S$, es tenen $E_s \subseteq E$ no buits i E_s numerables de forma que $E = \bigcup_{s\in S} E_s$. Aleshores $\overline{\overline{S}} \geq \overline{\overline{E}}$*

Demostració: Com que, pel teorema 10, el conjunt infinit S admet una partició $\mathcal{P}$ en conjunts enumerables, aleshores, per a cada $P \in \mathcal{P}$, el conjunt $E_P = \bigcup_{s\in P} E_s$ és numerable per ser reunió numerable de conjunts numerables. Per tant, tant P com E_P són numerables. Aleshores, existeix una funció $f_P \colon P \to E_P$ exhaustiva.

A partir d'aquesta funció exhaustiva f_P, construïm l'aplicació $f \colon S \to E = \bigcup_{s\in S} E_s$ definida com $f(s) := f_P(s)$. Aquesta aplicació és exhaustiva per ser-ho f_P per a tot $P \in \mathcal{P}$. Per tant, usant el corol·lari 36, tenim que $\overline{\overline{E}} \leq \overline{\overline{S}}$, tal com volíem demostrar. □

6.2. Llei de tricotomia dels cardinals. En aquest apartat, demostrem a partir del lema de Tukey la llei de tricotomia dels nombres cardinals, que ens diu que la relació d'ordre dels nombres cardinals definida a 1.3.6.1 és una relació d'ordre *total*, és a dir, els cardinals de qualsevol parell de conjunts A i B són comparables amb aquesta relació.

Proposició 12 (llei de tricotomia dels nombres cardinals). *Siguin A i B dos conjunts. Aleshores tenim que o bé $\overline{\overline{A}} < \overline{\overline{B}}$, o bé $\overline{\overline{A}} = \overline{\overline{B}}$ o bé $\overline{\overline{B}} < \overline{\overline{A}}$.*

Demostració:

Si $A = \emptyset$ o $B = \emptyset$, el resultat és clar. Suposem $A \neq \emptyset \neq B$.

Definim $\mathcal{F} := \{f \colon U \to B \mid f \text{ és injectiva i } U \subseteq A\}$. Defineixo la relació $\subseteq$ sobre $\mathcal{F}$ de forma que $f \subseteq g$ si i només si g és una extensió de f. Ordenem, doncs, la família de conjunts $\mathcal{F}$ per la inclusió $\subseteq$. Com que $A \neq \emptyset$ i $B \neq \emptyset$, existeixen $x \in A$ i $y \in B$. Considero l'aplicació $d \colon \{x\} \to B$ tal que $d(x) := y$, que la puc pensar com $\{(x, y)\}$ per la definició conjuntística d'aplicació. En aquest cas, tinc que $\{(x, y)\} \in \mathcal{F} \neq \emptyset$.

Volem veure que $\mathcal{F}$ és una família de caràcter finit. Sigui $f \in \mathcal{F}$ i sigui $g := \{(x_1, y_1), \ldots, (x_n, y_n)\} \subseteq f$, per a cert $n \in \mathbb{N}$, qualsevol subconjunt finit de f (ambdues aplicacions f i g pensades com a conjunts). Com que f és injectiva i f és una extensió de g, aleshores g també és injectiva. Per tant, $g \in \mathcal{F}$.

Sigui, ara, $f \subseteq A \times B$ tal que tot subconjunt finit de f és de $\mathcal{F}$. Anem a veure que $f \in \mathcal{F}$. Definim $U := \{x \in A \mid \text{existeix } y \in B \text{ que compleix } (x, y) \in f\}$, és a dir, el domini de f. Si per a un $x \in A$ existeixen $y_1, y_2 \in B$ tals que $(x_1, y_1), (x_2, y_2) \in f$, aleshores $\{(x, y_1), (x, y_2)\}$ és un subconjunt finit de f. Per tant, $\{(x, y_1), (x, y_2)\}$ és aplicació i, per tant, $y_1 = y_2$. Com que aquesta propietat es compleix per a qualsevol $x \in A$ arbitrari, aleshores f és una aplicació. Si per a un $y \in B$ existeixen $x_1, x_2 \in A$ tals que $(x_1, y), (x_2, y) \in f$, aleshores $\{(x_1, y), (x_2, y)\}$ és un subconjunt finit de f i, per hipòtesi, és aplicació injectiva. Per tant, per definició, $x_1 = x_2$. Com que aquesta propietat es compleix per a qualsevol $y \in B$ arbitrària, aleshores f és injectiva. Per tant, $f \in \mathcal{F}$.

Aleshores acabem de demostrar que $\mathcal{F} \neq \emptyset$ és una família de caràcter finit. Pel lema de Tukey, existeix $h \colon D \to B$ per a cert $D \subseteq A$ tal que $h \in \mathcal{F}$ és maximal amb la relació d'inclusió $\subseteq$. Anem a veure que o bé $U = A$ i, aleshores, $\overline{\overline{A}} = \overline{\overline{U}} \leq \overline{\overline{B}}$ o bé $\mathrm{Im}(h) = B$ i, aleshores, $\overline{\overline{A}} \geq \overline{\overline{U}} = \overline{\overline{B}}$. Si no es compleix cap de les dues

coses, existeixen $x \notin U$ i un $y \notin \mathrm{Im}(h)$. Aleshores $h \subseteq h \cup \{(x,y)\} \in \mathcal{F}$, fet que contradiu la maximalitat de h. La desigualtat $\overline{\overline{U}} \leq \overline{\overline{B}}$ es veu aplicant la injectivitat de l'aplicació h mentre que la desigualtat $\overline{\overline{A}} \geq \overline{\overline{U}}$ es veu aplicant el corol·lari 5.1.36.

Per tant, tenim que o bé $\overline{\overline{A}} < \overline{\overline{B}}$, o bé $\overline{\overline{A}} = \overline{\overline{B}}$ o bé $\overline{\overline{B}} < \overline{\overline{A}}$, tal com volíem demostrar. □

Anem a veure un lema sobre cardinals d'unions i interseccions que ens serà útil a la demostració de l'existència de clausura algebraica de cossos:

LEMA 13. *Sigui A un cos infint expressat com $A = B \cup C$ amb $B \cap C = \emptyset$. Aleshores:*

(1) o bé $|B| \leq |C|$ o $|C| \leq |B|$;
(2) si $|B| \leq |C|$, aleshores $|B \cup C| = |C|$;
(3) si $|C| \leq |B|$, aleshores $|B \cup C| = |B|$.

Demostració:

(1) El primer apartat del lema és simplement una aplicació de la llei de tricotomia 12 en aquest cas particular.
(2) Suposem que $|B| \leq |C|$. En tindrem prou veient que $|C \times \{0,1\}| = |C|$, atès que aleshores tindrem $|C| \leq |B \cup C| \leq |C \times \{0,1\}| = |C|$ i n'obtindrem la igualtat que es vol demostrar. Per demostrar-ho, sigui $\mathcal{D}$ el conjunt de totes les parelles (D, f) tals que $D \subseteq C$ i $f\colon D \times \{0,1\} \to D$ és una funció bijectiva. Com que C ha de ser infinit, podem seleccionar un subconjunt infinit numerable $D' = \{c_0, c_1, \ldots\} \subseteq C$. Definint $f'\colon D' \times \{0,1\} \to D'$ a partir de la fórmula $f'(c_n, i) = c_{2n+i}$, observem que $(D', f') \in \mathcal{D}$ de manera que $\mathcal{D} \neq \emptyset$. Com a la demostració de 12, ordenem parcialment el conjunt $\mathcal{D}$ per la relació $\leq$ definida per $(D_1, f_1) \leq (D_2, f_2)$ si i només si $D_1 \subseteq D_2$ i f_2 és una extensió de f_1. De manera anàloga veiem que és un conjunt inductiu i, per tant, podem aplicar el lema de Zorn per obtenir un element maximal (D, f). Aleshores $f\colon D \times \{0,1\} \to D$ és una bijecció i, per tant, $|D \times \{0,1\}| = |D|$. Per acabar la demostració, només ens falta veure que $|D| = |C|$. Com que C és infinit i $D \subseteq C$, n'hi ha prou amb veure que $C \setminus D$ és finit per tal d'establir la igualtat. Suposem per tal d'arribar a contraadicció que $C \setminus D$ és infinit i escollim un subconjunt finit enumerable $\bar{D} = \{c_0, c_1, \ldots\} \subseteq C \setminus D$. Com abans, hi ha una bijecció $\bar{f}\colon \bar{D} \times \{0,1\} \to \bar{D}$ que podem combinar amb la bijecció $f\colon D \times \{0,1\} \to D$ per obtenir una bijecció $g\colon (D \cup \bar{D}) \times \{0,1\} \to (D \cup \bar{D})$. Però aquest fet contradiu la maximalitat de (D, f) a $\mathcal{D}$. Per tant, $C \setminus D$ és finit i obtenim el que volíem demostrar. □
(3) Anàleg a (2) canviant els papers dels conjunts B i C.

Capítol 2
Axioma de l'elecció (equivalències bàsiques)

En aquest apartat enunciarem i demostrarem les equivalències més bàsiques i conegudes de l'axioma de l'elecció.

1. Caracterització de l'axioma de l'elecció i observacions sobre el seu ús

Intuïtivament, segons Bertrand Russell, l'axioma de l'elecció ens diu que "f selecciona un element de cadascun dels conjunts".

Anem a introduir aquí la primera caracterització de l'axioma de l'elecció i agafarem qualsevol dels enunciats equivalents següents com a formulacions de l'axioma:

PROPOSICIÓ 14 (caracterització de l'axioma de l'elecció). *Són equivalents:*

(e1): *Si $\mathcal{H} \neq \emptyset$ és una família no buida de conjunts no buits de E, aleshores existeix $f\colon \mathcal{H} \to E$ tal que $f(H) \in H$ per a tot $H \in \mathcal{H}$.*
(e2): *Si $\mathcal{P}$ és una partició de E, aleshores existeix un conjunt $X \subseteq E$ que conté exactament un element de cada $P \in \mathcal{P}$.*
(e3): *Si $\{E_i |\ i \in I\}$ és una família no buida de conjunts no buits, aleshores $\prod_{i\in I} E_i \neq \emptyset$.*

Demostració: (e1)⇒(e2). Agafem $\mathcal{H} = \mathcal{P}$ (que és una família no buida de conjunts no buits, per la primera condició de ser $\mathcal{P}$ una partició). Per **(e1)**, tenim que existeix $f\colon \mathcal{P} \to E$ tal que $f(P) \in P$ per a tot $P \in \mathcal{P}$. Llavors $X = \{f(P) \mid P \in \mathcal{P}\}$ té un element de cada P ja que es compleix per a tot $P \in \mathcal{P}$. A més a més, X té un únic element de cada conjunt de la partició, ja que la condició $f(P) \in P$ ens assegura que estem escollint exactament un element per cada conjunt P.

(e2)⇒(e3). Sigui $\{E_i \mid i \in I\}$ una família no buida de conjunts no buits. Per cada $i \in I$, definim ${E_i}' := \{(i, x) \mid x \in E_i\} \neq \emptyset$ no buits per hipòtesi. Observem que són disjunts dos a dos. Per tant, si agafem $E = \bigcup_{i\in I} {E_i}'$, aleshores $\left\{{E_i}' \mid i \in I\right\}$ és una partició de E. Per **(e2)**, existeix $X' \subseteq E'$ que conté exactament un element (i, x_i) de cada ${E_i}'$. Aleshores, l'aplicació $f\colon I \to E$ tal que $f(i) = x_i \in E_i$ és una funció d'elecció, és a dir, $f \in \prod_{i\in I} E_i \neq \emptyset$.

(e3)⇒(e1). Sigui $\mathcal{H} \neq \emptyset$ una família no buida de conjunts no buits de E. Aleshores, per **(e3)**, existeix un element $f \in \prod_{H\in\mathcal{H}} H \neq \emptyset$ tal que $f\colon \mathcal{H} \to \bigcup_{H\in\mathcal{H}} H \subseteq E$ (ja que la reunió de subconjunts H de E és un subconjunt de E) és una funció tal que $f(H) \in H$ per a tot $H \in \mathcal{H}$ (per definició de producte cartesià). □

Anomenarem *axioma de l'elecció* qualsevol de les tres formulacions anteriors.

Remarca. Existeixen molts casos particulars en els quals es disposa d'un criteri explícit *a priori* per poder definir funcions d'elecció.

Per exemple, sigui C la col·lecció de tots els subconjunts no buits de $\mathbb{N}$. Aleshores el principi de bona ordenació (equivalent al principi d'inducció), que ens diu que tot subconjunt no buit dels nombres naturals té un primer element, ens permet definir una funció $f\colon C \to \mathbb{N}$ definida com $f(S) = \min S$. Efectivament, f és una funció d'elecció. Per poder construir aquesta funció no hem fet ús de l'axioma de l'elecció, ja que l'hem pogut definir explícitament. En realitat, aquest exemple es pot estendre a col·leccions C de tots els subconjunts no buits d'un conjunt numerable indexat, agafant com a funció d'elecció aquella que a cada conjunt li assigna l'element del conjunt amb l'índex més petit.

Un altre exemple: Sigui ara C la col·lecció de tots els intervals no buits S de $\mathbb{R}$ de llargada finita, que és una col·lecció no numerable. Aleshores la funció $f\colon C \to \mathbb{R}$ definida com $f(S) = \frac{a_S + b_S}{2}$ que assigna a cada interval el seu punt mitjà és una altra funció d'elecció. En general, si donem col·leccions de subconjunts de $\mathbb{R}$ seguint algun criteri de definició, les condicions que ens determinen aquests conjunts poden definir-nos una regla (possiblement més complicada) per definir funcions d'elecció sense usar l'axioma.

En canvi, si C és la col·lecció de tots els subconjunts de $\mathbb{R}$, no queda gens clar com determinar un criteri per escollir una funció d'elecció f (una de les dificultats prové del fet que C és un conjunt no numerable). Donant una definició prou precisa de *trobar un criteri*, es pot demostrar fent servir raonaments de teoria de models que una tal funció f no es pot trobar amb l'axiomàtica de Zermelo-Fraenkel (ZF). Aleshores, com que l'axioma de l'elecció és indecidible a ZF, és imprescindible estendre l'axiomàtica a Zermelo-Fraenkel *Choice* (ZFC) assumint com a cert l'axioma de l'elecció dintre d'aquesta teoria per tal de poder assegurar l'existència de funcions d'elecció per al conjunt C.

Més generalment: Donada una col·lecció C de conjunts S no buits, les regles de la lògica formal de ZF només ens permeten fer un nombre finit n de *tries arbitràries* de conjunts S (utilitzant la definició de $S \neq \emptyset$). En el cas que $|C| < \infty$, aquestes $n = |C|$ *tries* ens defineixen una funció d'elecció sobre C. En canvi, si $|C| = \infty$, és necessari usar l'axioma de l'elecció per tal de poder fer infinites *tries arbitràries*, però podrien haver-hi altres maneres de construir funcions d'elecció. La cita següent de Bertrand Russell (a *Introduction to Mathematical Philosophy*) reflecteix aquest aspecte:

≪Per poder escollir un mitjó de cada parella d'una col·lecció infinita de parelles de mitjons és necessari l'axioma de l'elecció, però si en comptes de mitjons tenim sabates aleshores no és necessari≫.

Una possible funció d'elecció per a les sabates podria ser, per exemple, "de cada parell de sabates escullo la que correspon al peu esquerre". Per a mitjons, en canvi, aquest criteri no es pot adoptar perquè assumim que són indistingibles.

2. Axioma de l'elecció, Lema de Zorn i principi de bona ordenació

En aquest apartat, veurem que l'axioma de l'elecció és equivalent al lema de Zorn i al teorema de bona ordenació.

En primer lloc, enunciem el lema de Zorn:

ENUNCIAT 15 (Lema de Zorn). *Sigui $(E \leq)$ un conjunt parcialment ordenat. Si E és un conjunt inductiu, aleshores E conté (com a mínim) un element maximal.*

En altres paraules, tot conjunt ordenat que satisfà la propietat que tota cadena té una fita superior, té un element maximal.

En aquest punt, ja estem preparats per a demostrar l'equivalència entre l'axioma de l'elecció, el lema de Zorn i el teorema de bona ordenació.

PROPOSICIÓ 16. *Són equivalents:*

(1) L'axioma de l'elecció.
(2) El lema de Zorn.
(3) El teorema de la bona ordenació.

Demostració:

- (1) $\Rightarrow$ (2) Sigui E un conjunt ordenat inductiu. Volem demostrar que té un element maximal. Prenem $\mathcal{H} := \mathcal{P}(E) \setminus \{\emptyset\} = \{A \mid A \subseteq E\} \setminus \{\emptyset\}$. Per l'axioma de l'elecció, existeix $F\colon \mathcal{H} \to E$ tal que $F(H) \in H$ (ben definida, perquè l'antiimatge per F de tot element de E és un conjunt no buit, és a dir, $\{a\} \in F^{-1}(a)$ per a tot $a \in E$). Definim una funció $f\colon E \to E$ tal que:

$$f(x) = \begin{cases} x & \text{si } x \text{ és maximal;} \\ F(\{y \in E \mid y > x\}) \subseteq E & \text{si } x \text{ no és maximal.} \end{cases}$$

  Si x no és maximal, denotem $H := \{y \in E \mid y > x\}$. Observem que, en tal cas, $F(H) > x$, per ser $F(H) \in H = \{y \in E \mid y > x\}$ per definició de F. Aleshores, per la propietat reflexiva de la relació d'ordre $\leq$, per a tot $x \in E$ es té $x \leq f(x)$. Observem que l'aplicació f satisfà les hipòtesis del lema de Bourbaki (2), ja que $(\mathcal{P}(E) \setminus \{\emptyset\}, \subseteq)$ és un conjunt estrictament inductiu (utilitzant el lema 2.2.1) i $f\colon E \to E$ és tal que per a tot x es compleix $x \leq f(x)$. Aleshores, existeix un $x_0 \in E$ tal que $f(x_0) = x_0$. Per definició de la funció, f té com a punts fixos els elements maximals. Per tant, x_0 és un element maximal.
- (2) $\Rightarrow$ (3) Sigui E un conjunt, que podem suposar diferent del buit ($E = \emptyset$ està ben ordenat amb qualsevol relació d'ordre). Volem demostrar que es pot ben ordenar. Sigui

$$\mathcal{B} := \{(B, \leq) \text{ conjunt ben ordenat amb } B \subseteq E\}$$

  la família formada per tots els subconjunts de E que es poden ben ordenar. Aleshores:
  $\mathcal{B}$ és no buit: Si $x \in E$, aleshores $\{x\} \in \mathcal{B}$ es pot ben ordenar (qualsevol relació d'ordre serveix, ja que $\{x\}$ és l'únic subconjunt no buit de $\{x\}$ i té x com a primer element). Per tant, $\mathcal{B} \neq \emptyset$.
  Per a ordenar $\mathcal{B}$, definim la següent relació d'ordre: $(B_1, \leq_1) \leq (B_2, \leq_2)$ si $B_1 \subseteq B_2$ i $\leq_1 \subseteq \leq_2$. A partir d'ara diré que $\leq_1$ és *restricció* de $\leq_2$ a B_1 si $\leq_1 \subseteq \leq_2$.
  Comprovem que efectivament és una relació d'ordre:
  $\leq$ és reflexiva, ja que $B_1 \subseteq B_1$ i $\leq_1$ és restricció de $\leq_1$ (sempre és cert) implica, per definició de $\leq$, que $(B_1, \leq_1) \leq (B_1, \leq_1)$.
  $\leq$ és antisimètrica, perquè si tenim $(B_1, \leq_1) \leq (B_2, \leq_2)$ i $(B_2, \leq_2) \leq (B_1, \leq_1)$, aleshores $B_1 \subseteq B_2$, $B_2 \subseteq B_1$, $\leq_1$ és restricció de $\leq_2$ i $\leq_2$ és restricció de $\leq_1$, fet que implica que $(B_1, \leq_1) = (B_2, \leq_2)$.
  $\leq$ és transitiva. Si $(B_1, \leq_1) \leq (B_2, \leq_2)$ i $(B_2, \leq_2) \leq (B_3, \leq_3)$, aleshores $B_1 \subseteq B_2 \subseteq B_3$ i $\leq_1$ és restricció de $\leq_2$ que és restricció de $\leq_3$. Per tant, $B_1 \subseteq B_3$ i $\leq_1$ és restricció de $\leq_3$, que per definició de $\leq$ implica $(B_1, \leq_1) \leq (B_3, \leq_3)$.
  Vegem, ara, que $(\mathcal{B}, \leq)$ és un conjunt inductiu, és a dir, que tota cadena de $(\mathcal{B}, \leq)$ té una fita superior. Observem que totes les cadenes de $(\mathcal{B}, \leq)$ són de la forma $C = \{(B_i, \leq_i) | \ i \in I\}$ amb un conjunt d'índexs I tal que els conjunts B_i s'ordenen per inclusió i les relacions d'ordre $\leq_i$ són totes elles tals que $\leq_i$ és restricció de $\leq_j$ per tot $i, j \in I$ tals que $i \leq j$. Aquest fet prové de com hem construït la relació $\leq$. Per tant, per tal de veure la inductivitat de $(\mathcal{B}, \leq)$, només ens fa falta trobar una fita superior per a les cadenes C d'aquesta forma. Anem a veure que, si $B := \bigcup_{i \in I} B_i$, aleshores $C_0 := (\bigcup_{i \in I} B_i, \bigcup_{i \in I} \leq_i) = (B, \leq)$ és una fita superior de $C = \{(B_i, \leq_i) \mid i \in I\}$. Com que per a tot $x, y \in B$, tenim que $x \leq y$ si i només si existeix $i \in I$ tal que $x \leq_i y$, i com que per a tot $i \in I$ es té que $(B_i, \leq_i)$ són conjunts ben ordenats, aleshores $(B, \leq)$ també és un conjunt ben ordenat. Per tant, $(B, \leq) \in (\mathcal{B}, \leq)$. Observem, a més a més, que per a tot $i \in I$ es té que $(B_i, \leq_i) \leq (B, \leq)$, per ser $B_i \subseteq B$ i $\leq_i$ restricció de $\leq$. Per tant, $(B, \leq)$ és una fita superior de C.
  Utilitzant el lema de Zorn a la família parcialment ordenada i inductiva $(\mathcal{B}, \leq)$, tenim que existeix un element maximal $(M, \leq) \in (\mathcal{B}, \leq)$, que és ben ordenat per ser de $(\mathcal{B}, \leq)$. Si $M \neq E$, aleshores existeix $x \in E \setminus M$, d'on tenim que $M \cup \{x\}$ es podria ben ordenar, definint $m \leq x$ per qualsevol $m \in M$ i M un conjunt ben ordenat. Però llavors es tindria que $M \subsetneq M \cup \{x\}$, en contra de la hipòtesi de ser M maximal.
  La contradicció prové de suposar que $E \neq M$. Per tant, $(M, \leq) = (E, \leq)$, que és un conjunt ben ordenat.
- (2) $\Rightarrow$ (1) Sigui $\mathcal{H} \neq \emptyset$ una família arbitrària de subconjunts d'un conjunt E. Pel teorema de bona ordenació, E es pot ben ordenar (existeix una relació d'ordre $\leq$ tal que $(E, \leq)$ és un conjunt ben ordenat). Sigui $f\colon \mathcal{H} \to E$ la funció que envia conjunts $H \in \mathcal{H}$ al seu mínim, és a dir, $f(H) = \min H$. Observem que $f(H) = \min H \in H$, ja que, per ser $\leq$ un bon ordre, tot subconjunt H de E té un

primer element, que per definició és $\min H$, i aquest pertany a H per definició. Per tant, f és una funció d'elecció sobre $\mathcal{H} \neq \emptyset$ família arbitrària de subconjunts de E. □

3. Axioma de l'elecció, Principi de maximalitat de Hausdorff i lema de Tukey

En aquest apartat, veurem que l'axioma de l'elecció és equivalent al principi de maximalitat de Hausdorff i al lema de Tukey.

A continuació, presentem un altre resultat equivalent a l'axioma de l'elecció: el principi de maximalitat de Hausdorff. Aquest versió equivalent ens resultarà especialment útil a la demostració del teorema de la subbase d'Alexander, lema que ens servirà per a demostrar el teorema de Tychonov.

Definició. Una *cadena maximal* $(T, \leq)$ d'un conjunt $(S, \leq)$ és una cadena $(T, \leq) \subseteq (S, \leq)$ tal que si $(C, \leq) \subseteq (S, \leq)$ és una altra cadena tal que $T \subseteq C$, aleshores $T = C$.

Enunciat 17 (Principi de maximalitat de Hausdorff). *El principi de maximalitat de Hausdorff postul·la que tot conjunt parcialment ordenat conté una cadena maximal.*

Ara, enunciem un altre resultat equivalent a l'axioma de l'elecció: el lema de Tukey. Per a fer-ho, necessitarem definir la noció de família de caràcter finit. Aquesta versió equivalent de l'axioma ens servirà per a demostrar que tots espais vectorials no trivials tenen una base.

Definició. Diem que una família de conjunts $\mathcal{F}$ és de *caràcter finit* si es compleix:

(1) Si $A \in \mathcal{F}$ i $F \subseteq A$ és finit, aleshores $F \in \mathcal{F}$.
(2) Si tot conjunt $F \in A$ subconjunt finit és de $\mathcal{F}$, aleshores $A \in \mathcal{F}$.

Lema 18 (de Tukey). *Tota família de conjunts $\mathcal{F} \neq \emptyset$ de caràcter finit ordenada per inclusió té un element maximal.*

Proposició 19. *La reunió dels conjunts d'una cadena d'una família de caràcter finit $\mathcal{F}$ és de $\mathcal{F}$.*

Demostració: Sigui $\mathcal{C}$ una cadena arbitrària de $(\mathcal{F}, \subseteq)$ i $T := \bigcup_{C \in \mathcal{C}} C$ la seva reunió. Per definició de família de caràcter finit, per tal de veure que $T \in \mathcal{F}$ en tenim prou amb veure que qualsevol subconjunt finit de T és de $\mathcal{F}$. Sigui $F := \{x_1, \ldots, x_n\} \subseteq T$, doncs, un subconjunt finit qualsevol. Aleshores, per definició de reunió, existeixen $B_1, \ldots, B_n \in \mathcal{C}$ tals que $x_j \in B_j$ per a tot $j \in [n]$. Com que $\mathcal{C}$ és una cadena, existeix $j_0 \in [n]$ tal que $B_j \subseteq B_{j_0}$ per a tot $j \in [n]$. Per tant, $F \subseteq B_{j_0} \in \mathcal{C} \subseteq \mathcal{F}$. Com que $\mathcal{F}$ és de caràcter finit, aleshores $F \in \mathcal{F}$, i tenim el que volíem demostrar. □

A continuació, demostrem l'equivalència entre el lema de Zorn, el lema de Tukey i el principi de maximalitat de Hausdorff.

Proposició 20. *Són equivalents:*

(1) El lema de Zorn.
(2) El lema de Tukey.
(3) El principi de maximalitat de Hausdorff.

Demostració:

- (1) ⇒ (2) Anem a veure que si per a qualsevol conjunt parcialment ordenat en el qual tota cadena té una fita superior existeix un element maximal (lema de Zorn), aleshores per a qualsevol família de caràcter finit $\mathcal{F}$, i per a qualsevol conjunt $X \in \mathcal{F}$, existeix un conjunt $Y \in \mathcal{F}$ tal que $X \subseteq Y$ i Y és maximal respecte la relació $\subseteq$ (lema de Tukey). Sigui E un conjunt i $\mathcal{F} := \{F_i\}_{i \in I} \subseteq \mathcal{P}(E)$ una família

de caràcter finit no buida. Pel lema de Zorn és suficient demostrar que qualsevol cadena de $\mathcal{F}$ té una fita superior. Observem que podem ordenar $\mathcal{F}$ per inclusió, és a dir, $(\mathcal{F}, \subseteq)$ és un conjunt parcialment ordenat amb la relació d'ordre $\subseteq$. Amb aquesta relació d'ordre, podem veure que, per a tot element $B \in \mathcal{F}$, el conjunt $\{B\}$ és una cadena (per la propietat reflexiva, $\{B\} \subseteq \{B\}$), que té B com a fita superior i $B \in \mathcal{F}$.
Més generalment: Sigui $C = (F_i)_{i \in J \subseteq I} \subseteq \mathcal{F}$ una cadena. Anem a veure que aquesta cadena té una fita superior, $\bigcup_{i \in J \subseteq I} F_i$, i que aquesta fita superior pertany a $\mathcal{F}$. Sigui D un subconjunt finit de $\bigcup_{i \in J \subseteq I} F_i$. Aleshores qualsevol element $d \in D$ satisfà $d \in F_i$ per a algun $i \in I$. Com que D és finit i $(F_i)_{i \in J \subseteq I}$ és una cadena, existeix $k \in J$ tal que $D \subseteq F_k$. Però F_k és un element de la família de caràcter finit i, per tant, tots els seus subconjunts finits són de la família. Aleshores $D \in \mathcal{F}$. Per tant, tot subconjunt finit de $\bigcup_{i \in J \subseteq I} F_i$ és de $\mathcal{F}$. Per ser $\mathcal{F}$ família de caràcter finit, tots els conjunts tals que tots els seus subconjunts finits pertanyen a $\mathcal{F}$ són de $\mathcal{F}$. Aleshores tenim que $\bigcup_{i \in J \subseteq I} F_i \in \mathcal{F}$. En realitat, veure que aquesta reunió pertany a la família $\mathcal{F}$ és una aplicació de la proposició 19.
Com que $\bigcup_{i \in J \subseteq I} F_i$ és una fita superior de la nostra cadena arbitrària, concloem que totes les cadenes de la família tenen una fita superior i, pel lema de Zorn, $\mathcal{F}$ té un element maximal.

- (2) $\Rightarrow$ (3) Sigui $(A, \leq)$ un conjunt parcialment ordenat. Anem a veure que A té una cadena maximal. Sigui $\mathcal{C}$ la família de tots els subconjunts de A totalment ordenats (cadenes).
Veiem que $\mathcal{C}$ és una família de caràcter finit: Sigui $B \in \mathcal{C}$ una cadena. Aleshores tot subconjunt finit de B també és una cadena (perquè tot subconjunt d'un conjunt totalment ordenat és totalment ordenat) i, per tant, també pertany a $\mathcal{C}$. Sigui, ara, B un conjunt de A tal que tots els seus subconjunts finits són de $\mathcal{C}$, és a dir, són cadenes. Si B és finit, aleshores $B \subseteq B \in \mathcal{C}$ és una cadena. Si B és infinit, agafem $b_1, b_2 \in B$ dos elements qualsevol. Aleshores $S = \{b_1, b_2\} \subseteq \mathcal{C}$ és un subconjunt finit de $\mathcal{C}$ que, per hipòtesi, és una cadena. Per tant, $b_1 \leq b_2$ o $b_2 \leq b_1$. Com que aquesta propietat es compleix per a qualsevol parell d'elements de B, aleshores $B \in \mathcal{C}$ és una cadena. Per tant, $\mathcal{C}$ és una família de caràcter finit.
Aplicant el lema de Tukey a la família $\mathcal{C}$, tenim que $\mathcal{C}$ té un element maximal. Com que $\mathcal{C}$ l'hem definit com la família de totes les cadenes de A, tenim que aquest element maximal és, precisament, una cadena maximal.
- (3) $\Rightarrow$ (1) Anem a veure que si tot conjunt parcialment ordenat $(A, \leq)$ té una cadena maximal, aleshores tot conjunt inductiu té un element maximal. Sigui $(A, \leq)$ un conjunt inductiu, és a dir, tal que tota cadena de A té una fita superior. Pel principi de maximalitat de Hausdorff, existeix una cadena maximal $B \subseteq A$. Per hipòtesi, existeix una fita superior $c \in A$ de B. Anem a veure que c és un element maximal de A. Suposem, per reducció a l'absurd, que existeix $d \in A$ tal que $d > c$. Aleshores $B \cup \{d\}$ és una cadena de A, ja que B és una cadena i $d > b$ per a tot $b \in B$ (d és comparable amb tot element de B) per transitivitat de la relació d'ordre amb c. Però $B \cup \{d\}$ és una cadena de A que conté B estrictament (per ser $d \notin B$), en contra de la maximalitat de B. Per tant, c és un element maximal de A. $\square$

Capítol 3
Resultats equivalents

En aquest apartat enunciarem i demostrarem algunes equivalències bàsiques de l'axioma de l'elecció. La primera equivalència és de teoria de conjunts i teoria de l'ordre i ens permet reescriure l'axioma de l'elecció com un axioma que s'anomena l'axioma de les eleccions múltiples. A continuació, demostrem la versió general (per a espais vectorials infinits) del teorema d'Steinitz (àlgebra lineal). Tot seguit, demostrem l'equivalència de l'axioma de l'elecció i el teorema de Tychonov (topologia). Després demostrarem que l'axioma de l'elecció implica l'existència d'idaels maximals (sabem que el recíproc és cert, però n'ometrem la seva demostració). Finalment, com a curiositat, enunciarem sense demostrar un teorema de nombres cardinals que va portar certa polèmica, el teorema de Tarski.

1. Equivalències bàsiques addicionals de l'axioma de l'elecció

En aquest apartat, anem a veure algunes equivalències addicionals de l'axioma de l'elecció que no hem vist anteriorment i que tenen a veure directament amb teoria de conjunts i teoria de l'ordre. Anem a veure que és equivalent a l'axioma de les eleccions múltiples i al principi de les anticadenes.

L'axioma de les eleccions múltiples es pot enunciar de la manera següent:

Proposició 21. *(Axioma de les eleccions múltiples) Per a tota família $\mathcal{F}$ de conjunts no buits, existeix una funció f en $\mathcal{F}$ tal que per a tot $X \in \mathcal{F}$, $F(X)$ és un subconjunt no buit de X ($\emptyset \neq F(X) \subseteq X$).*

Anem a donar la definició d'anticadena:

Definició. Una *anticadena* en un conjunt parcialment ordenat és una cadena tal que tots els seus elements són incomparables.

Proposició 22. *Les següents proposicions són equivalents:*

(1) L'Axioma de l'elecció.
(2) L'Axioma de les eleccions múltiples.
(3) El principi de l'anticadena: Tot conjunt parcialment ordenat té una anticadena maximal.
(4) Tot conjunt linealment ordenat admet un bon ordre.
(5) El conjunt potència d'un conjunt ben ordenat pot ser ben ordenat.

Demostració:

- (1) $\Rightarrow$ (2) Sigui $f\colon \mathcal{F} \to \cup_{X\in\mathcal{F}} X$ una funció d'elecció (és a dir, $f(X) \in X$ per a tot $X \in \mathcal{F}$). Sigui ara $F\colon \mathcal{F} \to \{Y \neq \emptyset \mid \exists X \in \mathcal{F}\ Y \subseteq X\}$ tal que $F(X) = \{f(x)\}$. Aleshores F satisfà les hipòtesis de l'axioma de les eleccions múltiples, atès que $F(X) = \{f(X)\}$ és un conjunt no buit i, com que $f(X) \in X$, aleshores $F(X) = \{f(X)\} \subseteq X$. Per tant, es compleix l'axioma de les eleccions múltiples.

- (2) ⇒ (3) Sigui $(P, \leq)$ un conjunt parcialment ordenat. Aleshores per l'axioma de les eleccions múltiples existeix una funció $f\colon \mathcal{P}(P) \to \{Y \neq \emptyset \mid \exists X \in \mathcal{P}(P)\ Y \subseteq X\}$ tal que per a tot $X \subseteq P$, $X \neq \emptyset$, $f(X)$ és un subconjunt finit no buit de X.
  Definim $g(X)$ per a tot $X \subseteq P$ com $g(X) := \{x \in f(X) \mid x \text{ és minimal amb } \leq\}$. Aleshores $g(X)$ és una anticadena no buida finita, atès que cap d'aquests elements minimals són comparables (altrament, algun d'ells no seria minimal).
  Ara anem a construir una anticadena maximal utilitzant el principi de recursió transfinita. Sigui $A_0 := g(P)$ i, si $\alpha \in \mathrm{Ord}$ és un ordinal, aleshores $A_\alpha := g(X)$, on

  $$X := \{x \in P \mid x \text{ no és comparable amb cap element de } \cup_{\beta \in \mathrm{Ord}} \{A_\beta \mid \beta < \alpha\}\}.$$

  Finalment anem a veure que $A := \cup_{\alpha \in \mathrm{Ord}} A_\alpha$ és una anticadena maximal de P. En efecte, per construcció, és una anticadena, atès que cap dels seus elements són comparables. Suposem per reducció a l'absurd que existís algun element $a \in P$ tal que a no fos comparable amb cap element de A. Aleshores també existeix algun element de P que és comparable amb a, minimal amb la relació $\leq$, i que no pertany a A. Però per construcció de A, $a \in A_\alpha$ per a algun $\alpha \in \mathrm{Ord}$, fet que és una contradicció. Per tant, A és una anticadena maximal de P.
- (3) ⇒ (4) Sigui $(Q, \leq)$ un conjunt ordenat. Pel principi de bona ordenació, sabem que Q es podrà ben ordenar si podem trobar una funció d'elecció del conjunt potència $\mathcal{P}(Q)$.
  Sigui $P := \{(X, x) \mid X \subseteq Q,\ x \in X\}$. Defineixo l'ordre parcial $\leq_P$ a P per $(X, x) \leq_P (Y, y)$ si i només si $X = Y$ i $x \leq y$. Pel principi de l'anticadena, $(P, \leq_P)$ té una anticadena maximal A i A és, clarament, una funció d'elecció que a cada conjunt de Q li assigna un element d'aquest conjunt, atès que si existís algun $X \subseteq Q$ no buit tal que $(X, X) \cap A = \emptyset$, aleshores afegint qualsevol element de la forma (X, b) amb $b \in X$ (no buit) a A arribaríem a una contradicció amb la maximalitat de A i, per tant, la funció està ben definida a $\mathcal{P}(Q)$ i efectivament li assigna un element d'aquest conjunt.
- (4) ⇒ (5) Sigui X un conjunt ben ordenat amb $\leq$. Considerem en $\mathcal{P}(X)$ l'ordre lexicogràfic definit per $A, B \in \mathcal{P}$ satisfan $A \leq_{\mathrm{lex}} B$ si i només si al pensar $A, B \subseteq X$ com a subconjunts de X totalment ordenats amb l'ordre induït com a subconjunt i escrivint-los com $A := (a_i)_{i \in \alpha}$ i $B := (b_i)_{i \in \beta}$ amb $\alpha, \beta \in \mathrm{Ord}$ ordinals amb $\alpha \leq \beta$ sense pèrdua de generalitat, o bé $a_i = b_i$ per a tot $i \in \alpha$ o bé existeix algun ordinal $k < \alpha$ tal que $a_i = b_i$ per a tot $i < k$ i $a_k \leq b_k$. Observem, doncs, que tot parell d'elements de $\mathcal{P}(X)$ són comparables. El bon ordre del conjunt X juntament amb el bon ordre dels nombres ordinals ens permeten comprovar que $(\mathcal{P}(X), \leq_{\mathrm{lex}})$ és un conjunt ben ordenat.
- (5) ⇒ (1) Assumim que el conjunt potència de tot conjunt ben ordenat és un conjunt ben ordenat. Aleshores només fa falta demostrar que per a qualsevol ordinal límit α, la família V_α de tots els conjunts de rang inferior a α pot ser ben ordenat.
  Sigui α un ordinal límit fixat, i sigui κ l'ordinal més petit tal que existeix una funció bijectiva de κ a V_α. Aleshores per hipòtesi el conjunt potència de κ pot ser ben ordenat.
  Definim W_β on $V_\beta := \bigcup_{\gamma < \beta} V_\gamma$ com un bon ordre de V_β tal que $\beta_1 < \beta_2$ implica $V_{\beta_1} < V_{\beta_2}$ per a tot $\beta < \alpha$ (és a dir, si $X, Y \in W_\beta$, aleshores $X < Y$ si $\mathrm{rang}(X) \leq \mathrm{rang}(Y)$) amb $W_0 = 0$ (ben ordenat) i β un ordinal límit. Ara, si $\beta = \gamma + 1$, V_β és el conjunt potència de V_γ. V_γ és un conjunt ben ordenat per W_γ (per hipòtesi d'inducció) i per tant hi ha una correspondència bijectiva amb un ordinal $\lambda < \kappa$. Per tant, tenim una bona ordenació W_β del conjunt potència de V_γ, de manera que tots els seus conjunts estan ben ordenats. Per tant, atès que qualsevol conjunt X pertany a algun V_γ i, per transitivitat, $X \subseteq V_\gamma$ ben ordenat, qualsevol conjunt es pot ben ordenar, fet que implica l'axioma de l'elecció. □

2. Bases d'espais vectorials

En aquest apartat es recorden algunes nocions bàsiques dels espais vectorials i es demostra l'existència de bases dels espais vectorials no trivials i que dues bases d'un espai vectorial tenen el mateix cardinal.

Definim, en primer lloc, els conceptes de combinació lineal, sistema de generadors, conjunts linealment independents i bases. Donem per sabuts coneixements sobre alguns conceptes i propietats bàsiques d'espais vectorials.

Definició.

Sigui K un cos, E un K-espai vectorial i $u \in E$ un vector. Direm que u és una *combinació lineal* d'un conjunt finit de vectors $u_1, \ldots, u_n \in E$, amb $n \in \mathbb{N}$, si existeixen escalars $\lambda_1, \ldots, \lambda_n \in K$ tals que $u = \sum_{i=1}^{n} \lambda_i u_i$.

Definició.

Sigui K un cos. Sigui E un K-espai vectorial. Diem que un conjunt de vectors S és un *conjunt generador* de E si per a qualsevol vector $u \in E$ existeix una combinació lineal de u formada per un subconjunt finit de vectors de S, és a dir, existeixen $n \in \mathbb{N}$, $(v_i)_{i\in[n]} \subseteq S$ vectors de S i $(\lambda_i)_{i\in[n]} \subseteq K$ escalars tals que $u = \sum_{i=1}^{n} \lambda_i u_i$. En tal cas, diem que S *genera* l'espai vectorial E.

Definició.

Sigui K un cos i E un K-espai vectorial. Diem que un conjunt de vectors L és un *conjunt independent* de E si per a tot $n \in \mathbb{N}$, $k_1, \ldots, k_n \in K$ escalars i $u_1, \ldots, u_n \in L$ vectors, es té que $\sum_{i=1}^{n} k_i u_i = 0$ implica $k_i = 0$ per a tot $i \in [n]$.

Definició.

Sigui K un cos i sigui E un K-espai vectorial. Diem que $B \subseteq E$ és una *base* (de Hamel) de E si és un conjunt independent i és un conjunt generador de E.

Remarca.

Sigui K un cos, E un K-espai vectorial i $S := \{e_1, \ldots, e_n\} \subseteq E$. Defineixo la funció $f\colon K^n \to E$ com aquella que envia vectors $(\lambda_1, \ldots, \lambda_n)$, amb $\lambda_1, \ldots, \lambda_n$ escalars, a $\sum_{i=1}^{n} \lambda_i e_i$. Aleshores:

(1) f és exhaustiva $\Leftrightarrow \{e_1, \ldots, e_n\}$ és un sistema de generadors.
(2) f és injectiva $\Leftrightarrow \{e_1, \ldots, e_n\}$ és un sistema de vectors linealment independents.

Anem a veure, utilitzant el lema de Zorn, l'existència de bases de K-espais vectorials E no trivials ($E \neq \{0\}$).

PROPOSICIÓ 23. *Sigui K un cos. Tot K-espai vectorial no trivial E té una base.*

Demostració: Sigui $\mathcal{S}$ el conjunt format per tots els subconjunts de vectors linealment independents de E ordenats amb la inclusió, és a dir, el conjunt $\mathcal{S} := \{C \subseteq E \mid C \text{ és linealment independent}\}$. Aleshores $\mathcal{S} \neq \emptyset$, ja que $\emptyset \in \mathcal{S}$ és un conjunt linealment independent. De fet, també ho podem veure observant que com que E és no trivial, aleshores existeix $x \in E$ tal que $x \neq 0$ i el conjunt $\{x\}$ és independent. Sigui $\mathcal{C} := \{T_j\}_{j\in J} \subseteq \mathcal{S}$ una cadena de $\mathcal{S}$, i sigui $U := \bigcup_{j\in J} T_j$ la unió de tots els elements de la cadena. Si $T \in \mathcal{C}$, aleshores $T \subseteq U$.

Veurem que U és independent. Sigui $\{x_1, \ldots, x_n\}$ un subconjunt finit de U i suposem $v_1x_1 + \ldots + v_nx_n = 0$. Per a cada x_i, existeix $T_i \in \mathcal{C}$ tal que $x_i \in T_i$. Com que en una cadena tot conjunt finit té màxim (remarca 3), existeix $T = \max\{T_1, \ldots, T_n\}$. Aleshores $x_1, \ldots, x_n \in T$, que és independent. Per tant, $r_1 = \ldots = r_n = 0$. Per tant, el conjunt U és linealment independent. Atès que U és un conjunt linealment independent de E que conté tots els elements de $\mathcal{C}$, podem concloure que $U \in \mathcal{S}$ és una fita superior de la cadena $\mathcal{C}$.

Aplicant el lema de Zorn, sabem que $\mathcal{S}$ té un element maximal M. Anem a veure que M forma una base: Si M no fos una base, aleshores, com que per ser un element de $\mathcal{S}$ és linealment independent, no pot generar E. Si $x \notin \langle M \rangle$ és un element que no està generat pels elements de M, aleshores $M \cup \{x\}$ és un conjunt linealment independent de $\mathcal{S}$ que té M com a subconjunt propi, fet que contradiu la maximalitat de M. Per tant, M és una base. □

De fet, aquest teorema es pot generalitzar al resultat següent:

PROPOSICIÓ 24. *Sigui K un cos. Sigui E un K espai vectorial no trivial, G un sistema de generadors de E i $L \subseteq G$ un conjunt linealment independent. Aleshores existeix una base B de E tal que $L \subseteq B \subseteq G$.*

Demostració: La demostració consisteix en repetir l'argument anterior agafant com a conjunt $S := \{C \subseteq V \mid C \text{ és linealment independent i } L \subseteq C \subseteq G\}, \subseteq)$ ordenat per inclusió, demostrant que S és una família de conjunts inductiva amb $\subseteq$ i aplicant el lema de Zorn. □

Remarca. Si $E = 0$ és l'espai vectorial trivial sobre un cos K amb la condició del remarca **??**, aleshores E no admet cap base (continguda a E). Observem que $0 \in E$ és l'únic vector de E que genera E. Però qualsevol escalar $k \in K$ satisfà $k0 = 0$ (en particular algun $k \neq 0$) i, per tant, 0 és linealment dependent i no pot formar una base. Per tant, la hipòtesi sobre la no trivialitat de l'espai vectorial E és necessària per a demostrar l'existència de bases. Tanmateix, alguns autors assumeixen a la definició de base d'espai vectorial que el conjunt buit és una base de l'espai vectorial $\{0\}$, malgrat no ser un conjunt generador. En qualsevol cas, és un conveni que només ens afecta al cas base i no té més rellevància que aquesta a l'hora de demostrar els teoremes generals.

Anem a veure que qualsevol parell de bases d'un K espai vectorial E no trivial tenen el mateix cardinal.

PROPOSICIÓ 25. *Sigui K un cos i E un K espai vectorial. Si A i B són dues bases de E, aleshores A i B tenen el mateix cardinal.*

Demostració: Si E és finit, el teorema és cert pel teorema de Steinitz d'àlgebra lineal, que diu que el nombre d'elements (finit) d'un conjunt de vectors linealment independents és més petit o igual que el nombre d'elements d'un sistema de generadors de E. Aquest fet permet demostrar que dues bases d'un espai vectorial de dimensió finita tenen el mateix cardinal, simplement aplicant les desigualtats obtingudes a les nostres bases ($\overline{\overline{A}} \leq \overline{\overline{B}}$ i $\overline{\overline{B}} \leq \overline{\overline{A}}$) i aplicant l'antisimetria de $\leq$ (el teorema de Schröder-Bernstein, 4).

Suposem, ara, que E és infinit. El mateix argument fet anteriorment per al cas E finit ens serveix pel cas que E admeti un sistema de generadors finit ja que, en tal cas, $\overline{\overline{A}} = \overline{\overline{B}}$.

Suposem que A i B són, ambdues, bases infinites (si una d'elles fos finita, seria un sistema de generadors finit de E i estaríem en el cas anterior). Notem $A := \{a_i\}_{i \in I}$ i $B := \{b_j\}_{j \in J}$. Cada $a \in A$ és combinació lineal dels vectors d'un conjunt finit $B_a \subseteq B$, per ser B una base de E i la definició de combinació lineal (sempre amb un nombre finit de sumands). Sigui $C := \bigcup_{a \in A} B_a$. Clarament, $C \subseteq B$, d'on obtenim la desigualtat $\overline{\overline{C}} \leq \overline{\overline{B}}$. Volem demostrar que $C = B$.

Agafem un element $b \in B$ tal que $b \neq 0$. Aleshores $\langle B \setminus \{b\}\rangle \neq E$ ja que $b \in B \subseteq E$. Sigui $v \in E \setminus \langle B \setminus \{b\}\rangle$ un vector ($v \neq 0$) que no és de $\langle B \setminus \{b\}\rangle$. Aleshores v és combinació lineal de la base de A, i el podem escriure de la forma $v = \lambda_1 a_1 + \ldots + \lambda_n a_n$ per a certs $\lambda_1, \ldots, \lambda_n \neq 0$ i $a_1, \ldots, a_n \in A$. Si per a tot a_i es tingués que $a_i \in \langle B \setminus \{b\}\rangle$, aleshores tindríem que $v \in \langle B \setminus \{b\}\rangle$, en contra del que hem suposat. Per tant, existeix $i \in [n]$ tal que $a := a_i \notin \langle B \setminus \{b\}\rangle$. Aleshores, $a = \mu_1 b_1 + \ldots + \mu_k b_k + \mu b$ amb $b_i \in B$ i $\mu_i, \mu \neq 0$ per a cert $k \in [m]$ i per a tot $i \in [k]$. Llavors, $b \in B_a \subseteq C$ implica que $B \subseteq C$, ja que hem agafat un $b \in B$ arbitrari. Per tant, ajuntant les inclusions $B \subseteq C$ i $C \subseteq B$, obtenim la igualtat $B = C$ (que implica $\overline{\overline{B}} = \overline{\overline{C}}$).

Observem que els conjunts B_a són numerables per a tot $a \in A$ (perquè són conjunts finits), A és un conjunt infinit i $C = \bigcup_{a \in A} B_a$ compleixen les hipòtesis del lema 11 i, per tant, tenim que $\overline{\overline{C}} \leq \overline{\overline{A}}$. Per tant, com que $\overline{\overline{B}} = \overline{\overline{C}}$, obtenim $\overline{\overline{B}} \leq \overline{\overline{A}}$.

Repetint el mateix argument intercanviant els papers de A i B, obtenim que $\overline{\overline{A}} \leq \overline{\overline{B}}$.

Per tant, aplicant el teorema de Schröder-Bernstein (4), obtenim la igualtat $\overline{\overline{A}} = \overline{\overline{B}}$. □

A continuació anem a enunciar i demostrar una generalització d'aquest teorema:

PROPOSICIÓ 26. *Siguin $\{u_i\}_{i \in I}$ una família independent i $\{v_j\}_{j \in J}$ una família generadora. Aleshores, es poden substituir vectors de la família generadora pels vectors de la família independent de tal manera que la família que resulta segueixi sent generadora.*

Demostració: Siguin $I' \subseteq I$ i $J' \subseteq J$ dos subconjunts dels conjunts d'índexs I i J arbitraris. Aleshores denotarem per $\{u_i\}_{i\in I'} \sqcup \{v_j\}_{j\in J'}$ la família de vectors que té com a conjunt d'índexs la reunió disjunta $I' \sqcup J'$, és a dir, la família formada pels vectors u_i i v_j corresponents als índexos de I' i J', respectivament.

L'enunciat del teorema assegura que existeix una aplicació injectiva $f\colon I \hookrightarrow J$ tal que la família $\{u_i\}_{i\in I} \sqcup \{v_j\}_{j\in J\setminus f(i)}$ és una família generadora.

Aplicant el principi de bona ordenació, es pot suposar que els conjunts d'índexs I i J tenen bons ordres, amb els quals treballarem a partir d'ara. Sigui $\{I_n\}_{n\geq 1}$ una successió de conjunts amb $I_n \subseteq I$.

Anem a construir inductivament aplicacions injectives $f_n\colon I \hookrightarrow J$ amb imatges $J_n := f_n(I_n)$ tals que el lema es compleixi per a la subfamília de $\{u_i\}$ indexada pel subconjunt I_k de la manera següent:

Prenem $I_0 = \emptyset$ i f_0 la inclusió del conjunt buit en J. Suposem que ja s'ha construït un subconjunt $I_n \subseteq I$ i una aplicació injectiva f_n tals que la família $\{u_i\}_{i\in I_n} \sqcup \{v_j\}_{j\in J\setminus J_n}$ és una família generadora, i que $I_n \neq I$. En aquest cas, per a cada índex $i \in I \setminus I_n$, el vector u_i és combinació lineal de vectors

$$u_i = \sum x_r u_{i_r} + \sum y_s v_{j_s},$$

amb $i_r \in I_n$ i $j_s \in J \setminus J_n$.

Per ser $\{u_i\}$ una família de vectors independent, qualsevol combinació lineal formada per aquests vectors ha de tenir alguns vectors v_j amb coeficients no nuls. Defineixo $\phi_n(i)$ com el mínim índex $j \in J \setminus J_n$ tal que existeix una combinació lineal de la forma anterior en al qual tots els vectors v_{j_s} que apareixen amb coeficient no nul tenen índexs $j_s \leq j$. Podem garantir l'existència de $\phi_n(i)$ gràcies al fet que J tingui un bon ordre. Sigui $T_n := \{\phi_n(i) \mid i \in I \setminus I_n\}$ el subconjunt de $J \setminus J_n$ format per tots aquests mínims. Aquesta aplicació ϕ_n no és necessàriament injectiva. Per a cada $j \in T_n$ considerem la també considerem la més petita de les seves antiimatges i, d'aquesta manera, definim el conjunt $S_n := \left\{\min\left\{\phi_n^{-1}(j)\right\} \in I \setminus I_n \mid j \in T_n\right\}$, que també podem garantir que està ben definit gràcies a la bona ordenació del conjunt I. Per construcció, l'aplicació $\phi_n\colon S_n \to T_n$ és una aplicació bijectiva entre dos subconjunts $S_n \subseteq I \setminus I_n$ i $T_n \subseteq J \setminus J_n$. Ara, definim:

$$I_{n+1} = I_n \sqcup S_n,\ f_{n+1}(i) = \begin{cases} f_n(i), & si \quad i \in I_n, \\ \\ \phi_n(i), & si \quad i \in S_n, \end{cases} \qquad J_{n+1} = f_{n+1}(I_{n+1}) = J_n \sqcup T_n.$$

L'aplicació f_{n+1} és injectiva, atès que f_n ho és per construcció i ϕ_n també. Aleshores, la nova família que obtenim per construcció d'aquesta aplicació, amb índexs al conjunt $I_{n+1} \sqcup (J \setminus J_{n+1})$, també és una família generadora. Anem-ho a veure: Per fer-ho, hem de demostrar que $\langle u_i\rangle_{i\in I_n} + \langle v_j\rangle_{j\in J\setminus J_n} = \langle u_i\rangle_{i\in I_{n+1}} + \langle v_j\rangle_{j\in J\setminus J_{n+1}}$, veient que cadascun dels vectors a un costat de la igualtat és combinació lineal dels vectors de l'altre costat, veient les dues inclusions de conjunts. En efecte, els vectors de la dreta que no són a l'esquerra són els u_i per a $i \in S_n$ i la inclusió $\supseteq$ se'n dedueix del fet que cadascun d'aquests vectors és una combinació lineal de vectors de la família de l'esquerra. Els vectors de l'esquerra que no són a la dreta són els v_j per a $j \in T_n$ i per veure la inclusió $\subseteq$ podem raonar per reducció a l'absurd. En efecte, suposem que algun d'aquests vectors v_j amb $j \in T_n$ no fos combinació lineal dels vectors de la família de la dreta. Sigui j l'índex més petit tal que això es compleix, que existeix per estar J ben ordenat. Aleshores existeix un índex $i := \min\left\{\phi^{-1}(j)\right\} \in S$ i la combinació lineal de u_i com

$$u_i = \sum_{i_r\in I_n} x_r u_{i_r} + x_j v_j + \sum_{j_s<j} y_s v_{j_r}$$

amb tots els coeficients no nuls, on tots els vectors u_{i_k} que apareixen pertanyen a la família de la dreta i tots els vectors v_{j_s} són combinacions lineals de vectors d'aquesta família per minimalitat de j. Per tant, aïllant v_j en aquesta expressió, deduïm que v_j també és combinació lineal de vectors d'aquesta família, arribant a contradicció. Per tant, queda demostrada la igualtat.

Considerem el subconjunt $I' := \cup_{n\geq 0} I_n \subseteq I$ i l'aplicació injectiva $f\colon I' \to J$ que estén totes les aplicacions f_n, amb $f(I') = J' = \cup_{n\geq 0} J_n \subseteq J$. Es vol demostrar que la família construïda a partir d'aquesta aplicació $\{u_i\}_{i\in I'} \sqcup \{v_j\}_{j\in J\setminus J'}$ també és generadora. En efecte, n'hi ha prou si veiem que tot vector v_j està generat per

vectors d'aquesta família, atès que els u_i amb $i \in I'$ ja sabem que hi són. Suposem per reducció a l'absurd que j és el mínim valor per al qual això no es compleix (ben definit pel bon ordre de J). Aleshores $j \in J'$, ja que si $j \in J \setminus J'$ el vector v_j pertanyeria a la família. Siguin i, n tals que $j = \phi_n(i)$. Aleshores es té la igualtat $u_i = \sum x_r u_{i_r} + \sum y_s v_{j_s} + x_j v_j$ amb $i_r \in I_n$, $j_s \in J \setminus J'$ amb $j_s < j$ i amb coeficients no nuls, utilitzant $j = \phi_n(i)$, d'on aïllant v_j arribem a contradicció. Per tant, seguirà sent una família generadora, tal com volíem veure.

Per acabar la demostració, volem veure que $I = I'$. Observem que si un índex i no pertany ni a I_n ni a I_{n+1}, aleshores $\phi_n(i) > \phi_{n+1}(i)$. En efecte, el valor $j = \phi_n(i)$ és igual a $f_n(i')$ per a algun índex i' i el vector v_j resulta ser substituït pel vector $u_{i'}$ en el pas següent, i com que aquest nou vector $u_{i'}$ és combinació lineal de vectors u_i per a $i \in I_{n+1}$ i vectors v_{j_s} amb $j_s < j$, el nou màxim $\phi_{n+1}(i)$ ha de ser estrictament inferior a l'anterior. Aleshores, si existís un índex $i \in I \setminus I'$ obtindríem una successió estrictament descendent infinita $\phi_1(i) > \phi_2(i) > \phi_3(i) > \cdots$ al conjunt ordenat I, cosa que és impossible. □

Anem a demostrar el recíproc de l'equivalència, és a dir, anem a veure que el teorema d'Steinitz implica l'axioma de l'elecció.

Per fer-ho, necessitarem fer ús de l'equivalència de l'axioma de les eleccions múltiples i l'axioma de l'elecció vista anteriorment.

PROPOSICIÓ 27. *L'axioma de l'elecció es pot deduir de l'afirmació que tot espai vectorial té una base.*

Demostració: Sigui $\{X_i \mid i \in I\}$ una família de conjunts no buits. Volem demostrar que tenim l'axioma de les eleccions múltiples, que per 21 implicarà l'axioma de l'elecció. En efecte, volem trobar una família $\{F_i \mid i \in I\}$ de conjunts finits no buits tals que $F_i \subseteq X_i$. Assumim, sense pèrdua de generalitat, que els conjunts X_i són tots ells disjunts dos a dos. Sigui $X := \bigcup_{i \in I} X_i$ i sigui k un cos (arbitrari). Aleshores podem considerar el cos de fraccions format per $k(X)$ que resulta d'adjuntar els elements de X com a "variables". Per a tot $i \in I$, defineixo el grau i d'un monomi com la suma dels exponents dels membres de X_i en aquest monomi. Una funció racional es diu *i-homogènia de grau d* si és el quocient de dos polinomis tals que tots els monomis del denominador tenen el mateix i-grau n mentre que tots els monomis del numerador tenen el mateix i-grau $n + d$. Les funcions racionals i-homogènies de grau 0 per a tot $i \in I$ constitueix un subcos K de $k(x)$. Per tant, $k(X)$ és un espai vectorial sobre K, i sigui V el subespai generat pel conjunt X.

Per hipòtesi, el K-espai vectorial V té una base. Fixem una base B i la utilitzem per definir explícitament aquests conjunts finits F_i. Per a tot $i \in I$ i $x \in X_i$, podem expressar x com a combinació K-lineal finita d'elements de B:

$$x = \sum_{b \in B(x)} \alpha_b(x) b,$$

on $B(x)$ és un subconjunt finit de B i $\alpha_b(x)$ és, per a $b \in B(x)$, un element no nul de K. Si y és un altre element del mateix X_i que x, aleshores tenim que per altra banda

$$y = \sum_{b \in B(y)} \alpha_b(y) b.$$

Ara, si multipliquem l'expressió de la combinació lineal de x per l'element $y/x \in K$, aleshores tenim

$$y = \sum_{b \in B(x)} (y/x)\alpha_b(x) b.$$

Comparant aquestes dues expressions de y i utilitzant el fet que B és una base, inferim que $B(x) = B(y)$ i $\alpha_b(y) = (y/x)\alpha_b(x)$. Això significa que el subconjunt finit $B(x)$ de B i els elements $\alpha_b(x)/x$ de $k(X)$ depenen únicament de i, i no de l'element particular $x \in X_i$. Per tant, podem anomenar $B_i := B(x)$ i $\alpha_b(x)/x := \beta_{bi}$. Observem que, atès que $\alpha_b(x) \in K$, β_{bi} és i-homogeni de grau -1 (i j-homogeni de grau 0 per $j \neq i$), de manera que, quan β_{bi} l'escrivim com a quocient de polinomis en forma reduïda, algunes variables de X_i han d'aparèixer al denominador. Defineixo F_i com el conjunt de tots els elements de X_i que apareixen al denominador de β_{bi} (en forma reduïda) per a algun $b \in B_i$. Aleshores F_i és un subconjunt finit no buit de X_i, tal com volíem veure. □

3. Teorema de Tychonov

Demostrarem el teorema de Tychonov fent servir l'axioma de l'elecció i després demostrarem que el teorema de Tychonov implica l'axioma de l'elecció per tal d'establir la seva equivalència. En aquest apartat, donarem per coneguts els resultats bàsics i conceptes d'un primer curs de topologia.

PROPOSICIÓ 28 (Teorema de la subbase d'Alexander). *Sigui $(X, \mathcal{T})$ un espai topològic i $\mathcal{S}$ una subbase de la topologia de X. Aleshores, X és compacte si i només si tot recobriment $\mathcal{U}$ de X format per elements de S admet un subrecobriment finit.*

Demostració: La demostració de $\Rightarrow$) és trivial, ja que si X és compacte tot recobriment $\mathcal{U}$ de X admet un subrecobriment finit; en particular, $\mathcal{U} = \mathcal{S}$ és un recobriment de X que admet un subrecobriment finit.

Per tal de demostrar $\Leftarrow$), sigui $\mathcal{S}$ una subbase de l'espai topològic $(X, \mathcal{T})$ que satisfà la propietat que tot $\mathcal{S}$-recobriment admet un subrecobriment finit. Hem de veure que X és compacte.

Suposem que X no és compacte, és a dir, que hi ha recobriments oberts de X que no admeten subrecobriments finits. Notem $\mathcal{R}$ el conjunt de tots aquests recobriments. La idea de la demostració és la següent: Si $\mathcal{U}$ és un element maximal de $\mathcal{R}$, aleshores $\mathcal{U} \cap \mathcal{S}$ és un recobriment obert i, com que està format per elements de $\mathcal{S}$, admet un subrecobriment finit, en contradicció amb l'elecció dels elements de $\mathcal{R}$.

Si considerem $(\mathcal{R}, \subseteq)$ la família de conjunts $\mathcal{R}$ ordenada parcialment per inclusió (recordem que la inclusió sempre és una relació d'ordre parcial), aleshores, pel principi de maximalitat de Hausdorff, existeix una cadena maximal $\mathcal{C} \subseteq \mathcal{R}$. Sigui $\mathcal{U}$ el recobriment obert de X format per tots els elements de la cadena $\mathcal{C}$, és a dir, $\mathcal{U} = \{U \in \mathcal{T} \mid U \in C_\alpha,\ C_\alpha \in \mathcal{C}\}$.

Observem que $\mathcal{U}$ és un recobriment obert de X que, per construcció, no admet cap subrecobriment finit. El fet que $\mathcal{C}$ sigui cadena maximal de $\mathcal{R}$ i $\mathcal{U}$ tingui tots els oberts de totes les famílies de la cadena ens garanteix que $\mathcal{U}$ és un recobriment obert maximal amb la propietat de no tenir subrecobriments finits de X. D'altra banda, per a tot obert $V \subseteq X$ tal que $V \notin \mathcal{U}$, el recobriment obert $\mathcal{U}_V = \{V\} \cup \mathcal{U}$, sí que admet un recobriment finit, com se segueix de la maximalitat de $\mathcal{U}$. Sigui $\mathcal{U}' := \mathcal{U} \cap \mathcal{S}$, és a dir, $\mathcal{U}'$ és el conjunt format pels oberts de $\mathcal{U}$ que alhora són de $\mathcal{S}$. És suficient veure que $\mathcal{U}'$ és un recobriment de X, perquè aleshores admet un subrecobriment finit per estar format per elements de $\mathcal{S}$, cosa que contradiu l'elecció de $\mathcal{U}$. Anem a veure que, efectivament, $\mathcal{U}'$ és un recobriment:

Suposem que $\mathcal{U}'$ no recobreix X. Aleshores, existeix $x \in X \setminus \bigcup_{U \in \mathcal{U}'} U$. Com que $\mathcal{U}$ és un recobriment, existeix un obert $W \in \mathcal{U}$ tal que $x \in W$ i, com que és una subbase, existeixen $m \in \mathbb{N}$ i $S_1, \ldots, S_m \in \mathcal{S}$ tals que $x \in S_1 \cap \ldots \cap S_m \subseteq W$. Per l'elecció de x, es té que per a tot $i \in [m]$, $S_i \notin \mathcal{U}'$ (ja que $x \in S_i$). Per tant, per a tot $i \in [m]$, els recobriments $\mathcal{U}_{S_i}$ admeten subrecobriments finits, és a dir, $X = S_i \cup U_i$ per a tot $i \in [m]$, on cada U_i és una reunió finita d'oberts de $\mathcal{U}$. Però, aleshores $X = U_1 \cup \ldots \cup U_m \cup (\cap_{i=1}^m S_i) \subseteq U_1 \cup \ldots \cup U_m \cup W$, que és un subrecobriment finit de $\mathcal{U}$, fet que és una contradicció. Per tant, X és compacte. □

PROPOSICIÓ 29 (Teorema de Tychonov). *El producte cartesià d'espais topològics d'una família és compacte amb la topologia producte si, i només si, tots els espais de la família són compactes.*

Demostració: Sigui $\{X_i \mid i \in I\}$ una família d'espais topològics. Aleshores, hem de veure que

$$\prod_{i \in I} X_i \text{ compacte amb la topologia producte } \Leftrightarrow \text{per a tot } i \in I,\ X_i \text{ és compacte.}$$

La implicació $\Rightarrow$) és immediata, ja que la continuïtat de les projeccions π_α sobre les components de $X = \prod_{i \in I} X_i$ fa que $\pi_\alpha(X) = X_\alpha$ també siguin compactes (la imatge d'un compacte per una aplicació contínua és un compacte).

Pel que fa a la implicació $\Leftarrow$), suposem que els espais X_α són compactes per a tot $\alpha \in J$. Per definició de la topologia producte, tenim que

$$\mathcal{S} = \left\{\pi_\alpha^{-1}(U_\alpha) \mid U_\alpha \subseteq X_\alpha \text{ obert}, \alpha \in J\right\},$$

és una subbase de X i, pel teorema de la subbase d'Alexander (proposició 28) només és necessari demostrar que tot recobriment de $\mathcal{S}$ admet un subrecobriment finit (per un lema conseqüència de la definició de compacitat per al conjunt X i la definició de subbase d'una topologia). Suposem, per reducció a l'absurd, que existeix un recobriment $\mathcal{U}$ de X format per elements de $\mathcal{S}$ que no admet subrecobriments finits.

Per a cada $\alpha \in J$, definim els conjunts

$$\mathcal{D}_\alpha := \left\{U_\alpha \mid U_\alpha \subseteq X_\alpha \text{ obert}, \pi_\alpha^{-1}(U_\alpha) \in \mathcal{U}\right\}.$$

Anem a veure que no existeix cap $\alpha \in J$ tal que la família d'oberts $\mathcal{D}_\alpha$ de X_α sigui un recobriment. Com que X_α és compacte, si $\mathcal{D}_\alpha$ és un recobriment de X_α, aleshores $\mathcal{D}_\alpha$ admet un subrecobriment finit $\mathcal{D}_{\alpha_0} := (U_{\alpha_0 k})_{k\in K}$. Aleshores, el conjunt $\mathcal{C}_{\alpha_0} := \left\{\pi_\alpha^{-1}(U_{\alpha_0 k}) \mid U_{\alpha_0 k} \in \mathcal{D}_{\alpha_0}\right\}$ format per les antiimatges de la projecció π_α dels oberts d'aquest subrecobriment formarien un subrecobriment finit de $\mathcal{U}$ (per veure-ho, cal utilitzar la definició de topologia producte d'un producte cartesià infinit d'espais topològics), en contradicció amb l'elecció de $\mathcal{U}$.

Així doncs, per a tot $\alpha \in J$, existeix $x_\alpha \in X_\alpha$ tal que $x_\alpha \notin \bigcup_{U\in\mathcal{D}_\alpha} U$, ja que altrament la família d'oberts $\mathcal{D}_\alpha$ formaria un recobriment de X_α. Aplicant l'axioma de l'elecció, podem construir una aplicació $x\colon J \to \bigcup_{\alpha\in J} X_\alpha$ definida per $x(\alpha) = x_\alpha$. Aleshores, $x \in X$ per definició de producte cartesià de conjunts, però en canvi $x \notin \bigcup_{U\in\mathcal{U}} U$, cosa que és contradictòria amb el fet que $\mathcal{U}$ és un recobriment de X. Per tant, no pot existir un recobriment com $\mathcal{U}$ sense subrecobriments finits, que és equivalent a dir que X és compacte. □

Remarca. A la demostració del teorema de Tychonov s'utilitza dues vegades l'axioma de l'elecció. Una a l'utilitzar el teorema de la subbase d'Alexander (proposició 28) i l'altra a la mateixa demostració del teorema.

Anem a veure la demostració del seu recíproc. Per fer-ho, primer recordem (sense demostrar) un teorema vist a Topologia que ens caracteritza els conjunts compactes:

LEMA 30. *Sigui X un espai topològic. Aleshores X és compacte si i només si satisfà que per a tota família $\mathcal{F}$ de subconjunts tancats de X que satisfan la propietat que per a qualsevol subconjunt finit $(F_k)_{k\in[n]} \subseteq \mathcal{F}$ es té que $\cap_{k\in[n]} F_k \neq \emptyset$ (propietat d'intersecció finita), aleshores $\cap_{F\in\mathcal{F}} F \neq \emptyset$.*

PROPOSICIÓ 31. *El teorema de Tychonov implica l'axioma de l'elecció.*

Demostració: Sigui $(X_i)_{i\in I}$ una família arbitrària de conjunts no buits. Volem veure que es satisfà l'axioma de l'elecció (per exemple, que $\prod_{i\in I} X_i \neq \emptyset$). Sigui $a \notin \cup_{i\in I} X_i$ i defineixo $Y_i := X_i \cup \{a\}$ per a tot $i \in I$. Per a cada Y_i, defineixo una topologia de manera que els seus únics conjunts tancats siguin Y_i, X_i i els conjunts finits, simplement considerant com a oberts de la topologia els seus complementaris. Clarament, cada Y_i és compacte (tot recobriment obert de Y_i admet un subrecobriment finit), de manera que $Y = \prod_{i\in I} Y_i$ també és compacte pel teorema de Tychonov.

Per a tot $i \in I$, sigui $Z_i := \{f \in Y \mid f(i) \in X_i\}$, i observem que cada Z_i és tancat. Per tal d'aplicar el lema 30, volem veure que satisfan la propietat de la intersecció finita, és a dir, que si $\mathcal{Z} := \{Z_i \mid i \in I\}$, aleshores per a tot subconjunt finit $(Z_k)_{k\in[n]}$ tenim que la intersecció $\cap_{k\in[n]} Z_k$ és no buida. En efecte, siguin $Z_1, \ldots, Z_n \in \mathcal{Z}$. Com que cada X_i és no buit -en particular, per a $i \in [n]$ amb $Z_i = \pi^{-1}(X_i)$, hi ha algun $x_i \in X_i$ i, utilitzant aquest element, juntament amb el punt distingit a, podem construir un element $x \in \cap_{i\in[n]} Z_i$. En efecte, podem construir x coordenada a coordenada de manera que $\pi_k(x) = x_i$ si $k = i$ per a algun $i \in [n]$ i $\pi_k(x) = d$ altrament. Per tant, $x \in \cap_{k\in[n]} Z_k \neq \emptyset$. Per la compacitat de Y i pel lema 30, deduïm que

$$\prod_{i\in I} X_i = \bigcap_{i\in I} Z_i \neq \emptyset,$$

obtenint el que volíem demostrar. □

4. Existència d'ideals maximals

En aquest apartat, demostrarem que tot anell té ideals maximals. Considerarem només anells commutatius i unitaris. Només veurem un recordatori de la definició d'ideal maximal a partir d'un anell i donarem per sabudes algunes propietats de grups, anells i ideals que es facin servir.

Definició.

Sigui A un anell. Un conjunt I de A és un *ideal* si compleix les següents propietats:

(1) $I \neq \emptyset$.
(2) Si $x, y \in I$, aleshores $x - y \in I$.
(3) Si $x \in I$ i $a \in A$, aleshores $ax \in I$.

Remarca. Observem que els ideals són subgrups (abelians) additius de A (és a dir, els ideals $(I, +)$ són subgrups de $(A, +)$).

Definició.

Sigui A un anell no trivial. Un ideal $\subseteq A$ és *maximal* si compleix:

(1) $M \neq R$
(2) Si $M \subseteq I \subseteq R$, aleshores $M = I$ o $I = R$.

Veiem, utilitzant el lema de Zorn, que tot ideal propi d'un anell A (no trivial) està contingut en un ideal maximal. Observem que a l'exemple 2.3 hem demostrat que el conjunt de tots els ideals de A ordenats amb la relació d'inclusió és un conjunt ordenat inductiu, però en aquest teorema necessitem demostrar la inductivitat d'una família d'ideals lleugerament diferent (ideals propis que contenen un ideal fixat).

PROPOSICIÓ 32. *Tot ideal propi d'un anell no trivial està contingut en un ideal maximal.*

Demostració: Sigui A un anell no trivial i sigui I_0 un ideal propi de A. Sigui $\mathcal{F} := \{I \text{ ideal} \mid I_0 \subseteq I \subsetneq A\}$ la col·lecció de tots els ideals propis que contenen I_0 de l'anell A ordenada amb la relació d'inclusió. Observem que $\subseteq$ és una relació d'ordre parcial de $\mathcal{F}$, perquè la relació d'inclusió sempre és reflexiva, simètrica i transitiva.

Anem a veure que tota cadena de $\mathcal{F}$ té fita superior:

En efecte, sigui $\mathcal{C} := \{I_k \mid I_k \in \mathcal{C},\ k \in K \text{ conjunt d'índexs}\} \subseteq \mathcal{F}$ una cadena arbitrària (que podem suposar no buida). Definim $J := \bigcup_{I \in \mathcal{C}} I$.

Anem a veure que J és un ideal.

J és no buit perquè és la reunió d'ideals de $\mathcal{C}$ (que, per la primera propietat que defineix un ideal, són no buits).

Per a tot $x, y \in J$, es té que, per definició, existeixen ideals $I_x, I_y \in \mathcal{C}$ tals que $x \in I_x$ i $y \in I_y$. Com que $\mathcal{C}$ és una cadena, es té que $I_x \subseteq I_y$ o $I_y \subseteq I_x$. Aleshores, sense pèrdua de generalitat, podem suposar que $I_x \subseteq I_y$, d'on tenim que $x, y \in I_y$ i, per ser I_y un ideal, tenim que $x - y \in I_y \subseteq J$. Per tant, J satisfà la segona propietat que defineix un ideal.

Si $x \in J$, aleshores existeix $I_x \in \mathcal{C}$ tal que $x \in I_x$ i per a tot $a \in A$, es té que $ax \in I_x \subseteq J$. Per tant, J satisfà la tercera propietat que defineix ser un ideal i, per tant, com que compleix les tres, J és un ideal.

Observem que $I_0 \subseteq I_k$ per a qualsevol $k \in K$ implica que $I_0 \subseteq \bigcup_{i \in \mathcal{C}} I = J$. A més a més $J \subsetneq A$: Suposem que $J = A$. Aleshores, com que existeix $1 \in A$ element neutre pel producte, tindríem $1 \in J$. Per tant, existiria $k_0 \in K$ tal que $1 \in I_{k_0}$. Però aleshores tindríem $A = I_{k_0} \subsetneq A$, ja que I_{k_0} és un ideal propi de A que conté 1 i per a tot $a \in A$, es té que $a1 = 1a = a \in I_{k_0}$, i obtenim una contradicció que prové de suposar $J = A$.

Per tant, J és un ideal propi de A que conté I_0 (i, per tant, $J \in \mathcal{F}$).

Com que $J \in \mathcal{F}$ és tal que conté tots els ideals de $\mathcal{C}$, aleshores J és una fita superior de $\mathcal{C}$. Per tant, acabem de demostrar que $\mathcal{F}$ és un conjunt inductiu. Aplicant el lema de Zorn a $\mathcal{F}$, tenim que $\mathcal{F}$ té un ideal maximal que és un ideal propi de A que conté I_0. Per tant, I_0 està contingut a un ideal maximal, com volíem demostrar. □

5. Teorema de Tarski

Finalment, enunciem (sense demostrar cap de les dues implicacions) el teorema de Tarski.

PROPOSICIÓ 33 (Teorema de Tarski). *Per a tot conjunt infinit existeix una bijecció* $\phi\colon A \times A \to A$.

La demostració que va fer Tarski l'any 1924 del fet que l'axioma de l'elecció es pogués demostrar amb la seva hipòtesi va generar controvèrsia entre matemàtics de l'època que sostenien postures molt radicals a favor o en contra d'acceptar l'axioma de l'elecció. Tarski explicà a Jan Mycielski que quan ell va intentar publicar el seu teorema a la revista científica *Comptes rendus de l'Académie des Sciences* de París, els matemàtics Fréchet i Lebesgue van rebutjar la seva publicació. Fréchet argumentà que «una implicació entre dues proposicions ben conegudes no és un nou resultat», mentre que Lebesgue escrigué que «una implicació entre dues proposicions falses no té cap mena d'interès».

Capítol 4
Aplicacions de l'axioma de l'elecció

En aquest capítol, demostrarem alguns teoremes que es dedueixen de l'axioma de l'elecció. En particular, demostrarem l'existència d'inversa per la dreta d'una aplicació exhaustiva (teoria de conjunts), els teoremes de Banach-Tarski (anàlisi funcional), l'existència de conjunts no mesurables (teoria de la mesura), la paradoxa de Banach-Tarski (teoria de la mesura) i l'existència de clausura algebraica d'un cos (teoria de cossos). Com veiem, l'Axioma de l'elecció té aplicacions en branques de les matemàtiques molt diverses.

1. Inversa per la dreta d'una aplicació exhaustiva

En primer lloc, veiem que es pot demostrar l'existència d'una *aplicació inversa per l'esquerra* d'una aplicació injectiva f sense usar l'axioma de l'elecció:

Proposició 34 (existència d'inversa per l'esquerra d'una aplicació injectiva). *Sigui $f\colon A \to B$ una aplicació. Aleshores f és injectiva si, i només si, existeix una aplicació $g\colon B \to A$ tal que $g \circ f = id_A$.*

Demostració:

$\Leftarrow$) Suposem que existeix $g\colon B \to A$ tal que $g \circ f = id_A$. Aleshores, per a qualsevol $x, y \in A$, si $f(x) = f(y)$, aplicant g, i utilitzant que $g \circ f = id_A$, obtenim:

$$x = g(f(x)) = g(f(y)) = y.$$

Per tant, f és injectiva.

$\Rightarrow$) Suposem que f és injectiva. Per a cada $y \in f(A)$, existeix un únic $x \in A$ tal que $y = f(x)$ (per definició de f injectiva). Definim $g\colon f(A) \to A$ com $g(y) = x$ (per tal d'imposar la condició de $g \circ f = id_A$). Si $f(A) \neq B$, prenem un $x_0 \in A$ arbitrari i definim $g(y) = x_0$ per a tot $y \in B \setminus f(A)$. Per tant, es té que $(g \circ f)(x) = x = id_A(x)$ per a tot $x \in A$, i g definida per a tot $y \in (B \setminus f(A)) \cup f(A) = B$. □

Ara, veiem que per a demostrar l'existència d'una *aplicació inversa per la dreta* d'una aplicació exhaustiva g és necessari usar l'axioma de l'elecció:

Proposició 35 (existència d'inversa per la dreta d'una aplicació exhaustiva). *Sigui $g\colon B \to A$ una aplicació. Aleshores g és exhaustiva si, i només si, existeix $f\colon A \to B$ tal que $g \circ f = id_A$.*

Demostració:

$\Leftarrow$) Donat $a \in A$, considerem $f(a)$. Tenim $g(f(a)) = (g \circ f)(a) = a$, i veiem que $f(a)$ és original de a per g. Per tant, g és exhaustiva.

$\Rightarrow$) Com que g és una funció exhaustiva, aleshores tots els conjunts de $\mathcal{P} := \left\{g^{-1}(a) \mid a \in A\right\}$ són no buits (ja que tot element de A té com a mínim una antiimatge per g). A més a més, per ser g una aplicació, aquests conjunts són disjunts dos a dos perquè no hi poden haver dos elements amb la mateixa imatge per

g (si existís $b \in B$ tal que $b \in g^{-1}(a_1)$ i $b \in g^{-1}(a_2)$, aleshores tindríem que $a_1 = g(b) = a_2$ i, per tant, $g^{-1}(a_1) = g^{-1}(a_2)$). A més a més, $\bigcup_{a \in A} \{b \in B \mid g(b) = a\} = B$, ja que g està definida per a tot $b \in B$ i $\mathrm{Im}(g) = A$. Per tant, $\mathcal{P}$ és una partició de B.

Aplicant l'axioma de l'elecció a la partició $\mathcal{P}$, tenim que existeix un conjunt que té exactament un element y_a de $g^{-1}(a) \in \mathcal{P}$ per a tot $a \in A$. Aleshores $f\colon A \to B$ definida per $f(a) := y_a$ és l'aplicació cercada, ben definida i que efectivament compleix que $g \circ f = id_A$. □

Corol·lari 36. *Si $g\colon B \to A$ és exhaustiva, aleshores $\overline{\overline{A}} \leq \overline{\overline{B}}$.*

Demostració: En efecte, per la proposició 35, si $g\colon B \to A$ és exhaustiva aleshores existeix $f\colon A \to B$ tal que $g \circ f = id_A$. Per la proposició 34, si existeix $g\colon B \to A$ tal que $g \circ f = id_A$, aleshores f és injectiva. Per definició de $\leq$ als cardinals, el fet que $f\colon A \to B$ sigui injectiva ens diu que $\overline{\overline{A}} \leq \overline{\overline{B}}$, tal com volíem demostrar. □

2. Teoremes de Hahn-Banach

El teorema de Hahn-Banach és un dels resultats d'anàlisi funcional conseqüència de l'axioma de l'elecció que té diverses aplicacions a l'anàlisi funcional.

Anem a enunciar-lo, demostrar-lo i utilitzar-lo per demostrar-ne algunes aplicacions, en primer lloc, per al cas real:

Definició. Sigui X un espai vectorial sobre $\mathbb{R}$. Aleshores direm que $q\colon X \to \mathbb{R}$ és un *funcional sublineal* si verifica:

(a) $q(x + y) \leq q(x) + q(y)$ per a tot $x, y \in X$ (subadditivitat).
(b) $q(\alpha x) = \alpha q(x)$ per a tot $x \in X$ i $\alpha \geq 0$ (positivament homogènia).

Definició. Sigui X un espai vectorial sobre $\mathbb{R}$ i $q\colon X \to \mathbb{R}$ un funcional sublineal. Aleshores direm que q és un *funcional convex* si, a més a més, verifica la següent condició:

(c) $p(x) \geq 0$ per a tot $x \in X$ (positiva).

Les definicions anteriors han estat donades per al cas $K = \mathbb{R}$, però la definició següent pot ser donada per a espais vectorials sobre $K = \mathbb{R}$ o $K = \mathbb{C}$. entenent que si $\alpha \in \mathbb{C}$, aleshores $|\alpha|$ $(\in \mathbb{R})$ és la seva norma complexa:

Definició. Sigui X un espai vectorial sobre un cos $K = \mathbb{R}$ o $K = \mathbb{C}$. Una seminorma en E és una aplicació $p\colon E \to \mathbb{R}$ tal que

(a) $p(x + y) \leq p(x) + p(y)$ per a tot $x, y \in X$ (subbaditivitat).
(b) $p(\lambda x) = |\lambda| x$ per a tot $x \in E, \lambda \in \mathbb{K}$.
(c) $p(x) \geq 0$ per a tot $x \in X$ (positiva).

Observem que les seminormes (en particular, també les normes) són funcionals sublineals i també són funcionals convexos, però els recíprocs són falsos. Els funcionals sublineals poden prendre valors negatius i la condició (b) de la definició només és vàlida per valors reals $\alpha \geq 0$, mentre que en el cas de ser funcional convex segueix tenint el mateix problema amb la condició (b) per valors reals negatius.

Anem a demostrar un lema que ens serà útil per demostrar el teorema en el cas real:

Lema 37. *Sigui M un subespai propi de l'espai vectorial real X i $x_0 \in X \setminus M$. Defineixo $N :=< M \cup \{x_0\} >$ com l'espai generat per M i $\{x_0\}$ i suposem que f és un funcional lineal definit únicament a M, p és un funcional sublineal definit a X i es compleix que $f(x) \leq p(x)$ per a tot $x \in M$. Aleshores f es pot estendre a un funcional F definit sobre N amb la propietat $F(x) \leq p(x)$ per a tot $x \in N$.*

Demostració: Donat que $f(x) \leq p(x)$ a M, aleshores existeixen $y_1, y_2 \in M$ de manera que:

$$f(y_1) - f(y_2) = f(y_1 - y_2) \leq p(y_1 - y_2) = p(y_1 + x_0 - y_2 - x_0) \leq p(y_1 + x_0) + p(-y_2 - x_0).$$

Agrupant al primer membre dels termes que contenen y_2 i al segon terme els que contenen y_1, obtenim

$$-p(-y_2 - x_0) - f(y_2) \leq p(y_1 + x_0) - f(y_1).$$

Suposem que y_1 és fix mentre que y_2 recorre M. De l'expressió anterior deduïm que el conjunt de nombres reals

$$\{-p(-y_2 - x_0) - f(y_2) | y_2 \in M\}$$

té una fita superior $-y$ que també és un extrem superior. Llavors puc definir

$$a = \sup \{-p(-y_2 - x_0) - f(y_2) | y_2 \in M\}.$$

Anàlogament, podem assegurar l'existència de

$$b = \inf \{p(y_1 + x_0) - f(y_1) | y_1 \in M\}.$$

L'expressió anterior ens permet afirmar que $a \leq b$, fet que implica que existeix un nombre real c_0 tal que $a \leq c_0 \leq b$. Si estem al cas $a = b$, aleshores $a = b = c_0$.

Aleshores, per a tot $y \in M$,

$$-p(-y - x_0) - f(y) \leq c_0 \leq p(y + x_0) - f(y).$$

Donat que $x_0 \notin M$, podrem expressar qualsevol $x \in N$ com a

$$x = y + \alpha x_0,$$

essent α un escalar determinat de manera única i y un vector també únic de M. Donat que aquesta representació és única, aleshores l'aplicació

$$F \colon N \to \mathbb{R}$$

definida com

$$F(y + \alpha x_0) = f(y) + \alpha c_0$$

està ben definida i a més a més és un funcional lineal sobre el subespai N. També es pot veure que, si $y \in M$,

$$F(y) = f(y),$$

és a dir, F estén la funció f a N (ja que en aquest cas $\alpha = 0$). Per tal d'acabar la demostració, només ens falta veure que es satisfà la propietat

$$F(x) \leq p(x)$$

per a tot $x \in N$. Per fer aquesta demostració, anem a distingir els tres casos següents: Per a tot $x \in N$, $x = y + \alpha x_0$ s'ha de tenir $\alpha = 0$, $\alpha > 0$ o $\alpha < 0$.

(1) $\alpha = 0$. Observant que $F(y + \alpha x_0) = F(y) = f(y)$, només cal aplicar la hipòtesi $f(x) \leq p(x)$.
(2) $\alpha > 0$. A la desigualtat $c_0 \leq p(y+x_0) - f(y)$, substitueixo y per y/α. Aleshores ens queda la desigualtat

$$c_0 \leq p(y/\alpha + x_0) - f(y/\alpha).$$

Mulitplicant la desigualtat per α i tenint en compte que p és un funcional sublineal, arribem a la desigualtat

$$f(y) + \alpha c_0 \leq p(y + \alpha x_0),$$

que és equivalent a $F(x) \leq p(x)$ per a tot $x \in N$ tal com volíem demostrar.
(3) $\alpha < 0$. Considerem ara la desigualtat $-p(-y - x_0) - f(y) \leq c_0$ i també substituïm y per y/α. Aleshores obtenim la desigualtat

$$-p(-y/\alpha - x_0) - f(y/\alpha) \leq c_0,$$

que és equivalent a

$$-p(-y/\alpha - x_0) \leq c_0 + f(y/\alpha).$$

Multiplicant per α la desigualtat anterior, aleshores la desigualtat canvia de signe i arribem a

$$(-\alpha)p(-y/\alpha - x_0) \geq \alpha c_0 + f(y),$$

d'on veiem que per ser $-\alpha > 0$ i aplicant que p és sublineal en podem deduir $p(y + \alpha x_0) \geq \alpha c_0 + f(y)$ o, el que és el mateix, $p(x) \geq F(x)$ per a tot $x \in N$, tal com volíem demostrar. □

Ara ja podem enunciar i demostrar el teorema:

PROPOSICIÓ 38 (Teorema de Hahn-Banach, cas real). *Sigui X un espai vectorial sobre $\mathbb{R}$ i sigui p un funcional sublineal sobre X. Sigui M un subespai vectorial de X i f un funcional lineal sobre M tal que per a tot $x \in M$ es compleix*

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

Aleshores existeix un funcional lineal F definit en tot X que estén f i tal que

$$F(x) \leq p(x)$$

per a tot $x \in X$.

Demostració: Sigui S el conjunt format per tots els funcionals lineals $\{\hat{f}\}$, que estenen a f i tals que

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

tenint que $x \in D_{\hat{f}}$ essent $D_{\hat{f}}$ un subespai vectorial de X que agafem com a domini de $\hat{f}$.

El conjunt S és no buit, ja que f hi pertany ($f(x) \leq p(x)$ i està definida al domini M, que és subespai vectorial de X).

Al conjunt S, defineixo la relació $\leq$ següent: Donades dues funcions $\hat{f}_1$ i $\hat{f}_2 \in S$, aleshores

$$\hat{f}_1 \leq \hat{f}_2 \Leftrightarrow D_{\hat{f}_1} \subseteq D_{\hat{f}_2} \text{ i } \hat{f}_{2|D_{\hat{f}_1}} = \hat{f}_1,$$

on $\hat{f}_{2|D_{\hat{f}_1}}$ denota la restricció de l'aplicació $\hat{f}_2$ al domini $D_{\hat{f}_1}$.

Anem a veure que $\leq$ defineix un ordre parcial:

(1) $\leq$ és reflexiva, ja que $D_{\hat{f}_1} \subseteq D_{\hat{f}_1}$ i $\hat{f}_{1|D_{\hat{f}_1}} = \hat{f}_1$.
(2) $\leq$ és antisimètrica, ja que, si $\hat{f}_1 \leq \hat{f}_2$ i $\hat{f}_2 \leq \hat{f}_1$, aleshores $D_{\hat{f}_1} \subseteq D_{\hat{f}_2}$ i $\hat{f}_{2|D_{\hat{f}_1}} = \hat{f}_1$ i $D_{\hat{f}_2} \subseteq D_{\hat{f}_1}$ i $\hat{f}_{1|D_{\hat{f}_2}} = \hat{f}_2$, fet que implica que $D_{\hat{f}_1} = D_{\hat{f}_2}$ i $\hat{f}_1 = \hat{f}_2$.
(3) $\leq$ és transitiva, ja que si $\hat{f}_1 \leq \hat{f}_2$ i $\hat{f}_2 \leq \hat{f}_3$, aleshores $D_{\hat{f}_1} \subseteq D_{\hat{f}_2} \subseteq D_{\hat{f}_3}$ (en particular, $D_{\hat{f}_1} \subseteq D_{\hat{f}_3}$) i $\hat{f}_{2|D_{\hat{f}_1}} = \hat{f}_1$ i $\hat{f}_{3|D_{\hat{f}_2}} = \hat{f}_2$ impliquen, fent composició de restriccions, que $\hat{f}_{3|D_{\hat{f}_1}} = \hat{f}_1$, d'on tenim $\hat{f}_1 \leq \hat{f}_3$.

Anem a veure que S és un conjunt inductiu, és a dir, que tot subconjunt totalment ordenat de S té una fita superior a S. Sigui $C = \{\hat{f}_\alpha\}$ una cadena arbitrària de S.

Volem demostrar que existeix $\hat{f} \in S$ tal que $\hat{f}$ és una fita superior de T. Defineixo $\hat{f}$ aquella funció que té domini $\cup_\alpha D_{\hat{f}_\alpha}$ i tal que $\hat{f}(x)_{|D_{\hat{f}_\alpha}} = \hat{f}_\alpha(x)$ per a tot α, de manera que està definida a tot el conjunt (l'hem definit a un recobriment del seu domini).

Per veure que $\hat{f} \in S$, en primer lloc hem de comprovar que el domini $\cup_\alpha D_{\hat{f}_\alpha}$ és un subespai vectorial de X. Si $x \in \cup_\alpha D_{\hat{f}_\alpha}$, aleshores existeix algun α tal que $x \in D_{\hat{f}_\alpha}$ i, com que $D_{\hat{f}_\alpha}$ és un subespai vectorial, aleshores per a tot escalar λ es compleix $\lambda x \in D_{\hat{f}_\alpha}$. Ara suposem que $x, y \in \cup_\alpha D_{\hat{f}_\alpha}$. Aleshores existeixen α_1 i α_2 de manera que $x \in D_{\hat{f}_{\alpha_1}}$ i $y \in D_{\hat{f}_{\alpha_2}}$, i com que C és una cadena (conjunt totalment ordenat amb la relació d'inclusió) aleshores $D_{\hat{f}_{\alpha_1}} \subseteq D_{\hat{f}_{\alpha_2}}$ o $D_{\hat{f}_{\alpha_2}} \subseteq D_{\hat{f}_{\alpha_1}}$. Sense pèrdua de generalitat, podem suposar que $D_{\hat{f}_{\alpha_1}} \subseteq D_{\hat{f}_{\alpha_2}}$. Aleshores, com que $x, y \in D_{\hat{f}_{\alpha_2}}$ i $D_{\hat{f}_{\alpha_2}}$ és un espai vectorial, tenim que $x+y \in D_{\hat{f}_{\alpha_2}} \subseteq \cup_\alpha D_{\hat{f}_\alpha}$, tal com volíem veure. Per tant, $\cup_\alpha D_{\hat{f}_\alpha}$ és un subespai vectorial de X.

Per veure que $\hat{f}$ està ben definida, suposem que $x \in D_{\hat{f}_\alpha}$ i $x \in D_{\hat{f}_\beta}$. Per definició de $\hat{f}$, tenim que $\hat{f}(x) = \hat{f}_\alpha(x)$ i $\hat{f}(x) = \hat{f}_\beta(x)$. Com que C és una cadena, tenim que $\hat{f}_\alpha$ és una extensió de $\hat{f}_\beta$ o viceversa, i podem suposar

sense pèrdua de generalitat que $\hat{f}_\alpha$ és una extensió de $\hat{f}_\beta$. Però en qualsevol cas es té que $\hat{f}_\alpha = \hat{f}_\beta$ sobre el domini més petit, de manera que $\hat{f}$ està ben definida. Observem que $\hat{f}$ és una aplicació lineal que estén f i que compleix $\hat{f}(x) \leq p(x)$ quan $x \in D_{\hat{f}}$ (per definició de S). A més a més, per a qualsevol $\hat{f}_\alpha \in C$, es té que $\hat{f}_\alpha \leq \hat{f}$. Per tant, $\hat{f}$ és una fita superior (a S) de la cadena C i, per tant, S és un conjunt inductiu.

Aplicant el lema de Zorn, existeix $F \in S$ element maximal de S. Donat que $F \in S$, F és un funcional lineal que estén f i amb la propietat $F(x) \leq p(x)$ per a tot $x \in D_F$, essent D_F un espai vectorial.

Per acabar la demostració, només ens fa falta veure que $D_F = X$. La inclusió $D_F \subseteq X$ és trivial perquè X és l'espai vectorial ambient. Per veure l'altra inclusió, suposem per tal d'arribar a contradicció que no és així i que, per tant, $X \setminus D_F \neq \emptyset$. Sigui $x_0 \in X \setminus D_F$. Aplicant el lema 37, tenim que F pot estendre's a un funcional F' que estén f i satisfà $F'(x) \leq p(x)$ per a tot $x \in < D_F \cup \{x_0\} >$. Per tant, F' pertany a S i estén F amb $D_F \subsetneq D_{F'}$, fet que contradiu la maximalitat de F. Per tant, x_0 no pot existir i el domini de F és tot X. $\square$.

Observació. El teorema de Hahn-Banach és un resultat intermedi entre l'axioma de l'elecció i l'existència de subconjunts no mesurables de $\mathbb{R}$. En primer lloc, fem una llista de proposicions per poder establir les relacions entre aquests resultats:

(1) Axioma de l'elecció.
(2) Teorema de Krull: En un anell commutatiu, tot ideal propi es pot estendre a un ideal primer maximal.
(3) Teorema dels Ideals Booleans Primers: En un anell booleà, existeix com a mínim un ideal maximal.
(4) Teorema de Hahn-Banach.
(5) Per a tota àlgebra booleana, existeix una mesura additiva m sobre els reals que satisfà $m(0) = 0$ i $m(1) = 1$.
(6) Existeix un subconjunt no mesurable a $\mathbb{R}$.

Aleshores, entre aquests resultats, es tenen les següents implicacions:

$$(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Leftrightarrow (5) \Rightarrow (6),$$

mentre que es poden trobar contraexemples que demostren que

$$(1) \not\Leftarrow (2) \not\Leftarrow (3) \not\Leftarrow (4)$$

i

$$(5) \not\Leftarrow (6).$$

Observem, doncs, que el teorema de Hahn-Banach és una versió més feble de l'axioma de l'elecció que, en particular, es pot demostrar a partir del Teorema dels Ideals Booleans Primers que, en certes ocasions i per a alguns matemàtics, s'escull com a versió més feble de l'axioma.

El teorema originalment va ser demostrat per Banach usant el principi de la bona ordenació com a versió equivalent de l'Axioma de l'elecció en comptes del lema de Zorn, que posteriorment va ser demostrat a partir del Teorema dels Ideals Booleans Primers (3) i demostrat equivalent al resultat (5) per Los i Ryll-Nardzewski l'any 1951.

3. Existència de conjunts no mesurables

Per demostrar l'existència de conjunts no mesurables, primer necessitem definir el conjunt de diferències d'un conjunt mesurable i demostrar el teorema de Steinhaus.

Definició. Sigui $E \subseteq \mathbb{R}$ un conjunt mesurable. Definim el seu *conjunt de diferències* per

$$D(E) = \{x - y \mid x, y \in E\}.$$

Proposició 39. *(Teorema de Steinhaus) Si $m(E) > 0$, existeix un $\delta > 0$ tal que $(-\delta, \delta) \subseteq D(E)$.*

Demostració: Com que les mesures són regulars, existeix un tancat $F \subseteq E$ amb $0 < m(F) < +\infty$, i com que $D(F) \subseteq D(E)$, només farà falta comprovar el teorema per F.

Com que suposem que F és tancat, si anomenem $U_n := \left\{x \in E \mid d(x,F) < \frac{1}{n}\right\}$, tindrem que

$$F = \bigcap_{n\in\mathbb{N}} U_n$$

i, en conseqüència,

$$m(F) = \lim_{n\to+\infty} m(U_n),$$

de manera que podrem escollir n_0 tal que $m(U_{n_0}) < \frac{3}{2}m(F)$.

Aleshores prenem $\delta < \frac{1}{n_0}$. Afirmem que si $|x| < \delta$, aleshores $F \cap (F+x) \neq \emptyset$: si suposem que no, tindrem que:

$$m(U_{n_0} - [F \cap (F+x)]) \leq m(U_{n_0} - F) + m(U_{n_0} - (F+x))$$

atès que

$$U_{n_0} - [F \cap (F+x)] = (U_{n_0} - F) \cup (U_{n_0} - (F+x))$$

però per l'elecció de δ, $F + x \subseteq U_{n_0}$, de manera que obtenim que:

$$m(U_{n_0} - [F \cap (F+x)]) \leq m(U_{n_0}) - m(F) + m(U_{n_0}) - m(F+x) =$$

$$= 2m(U_{n_0}) - 2m(F) < m(F)$$

per la invariància per translacions de la mesura de Lebesgue i per $m(U_{n_0}) < \frac{3}{2}m(F)$.

Aquest fet és una contradicció, atès que com que $F \subseteq U_{n_0}$, hem de tenir que $|F| \leq |U_{n_0}|$. Aquesta contradicció prové de suposar que F i $F+x$ eren disjunts i, per tant, existirà $y \in F$ tal que $x + y \in F$, i per tant $x = (y+x) - y \in D(F) \subseteq D(E)$. □

Anem a demostrar que existeixen conjunts no mesurables a partir de l'Axioma de l'elecció. Una de la família d'exemples més senzills de conjunts no mesurables són els conjunts de Vitali, que són els que veurem en aquest apartat a continuació. Al següent apartat, apareixeran altres conjunts no mesurables que ens permetran arribar a la paradoxa de Banach-Tarski.

Proposició 40. *(Existència de conjunts no mesurables o teorema de Vitali). Existeixen conjunts no mesurables (Lebesgue). Encara més: Qualsevol conjunt mesurable $E \subseteq \mathbb{R}$ amb mesura positiva conté un conjunt no mesurable.*

Demostració: En el conjunt E (mesurable amb mesura positiva) definim la següent relació d'equivalència:

$$x \simeq y \Leftrightarrow x - y \in \mathbb{Q}.$$

Considerem aquesta relació restringida al conjunt E. Aquesta relació ens determina una partició del conjunt E en classes d'equivalència, és a dir, la partició $\mathcal{P} := \{[x] \mid x \in E\}$. Utilitzant l'axioma de l'elecció, construïm un conjunt V escollint un (únic) element de cada classe d'equivalència de E, atès que existeixen funcions d'elecció que ens permeten fer-ho. Anomenem a qualsevol conjunt V çonstruït"d'aquesta manera *conjunt de Vitali*. Anem a veure que V és un conjunt no mesurable tal com cercàvem:

En primer lloc, enumerem el conjunt $\mathbb{Q}$ dels racionals com $\mathbb{Q} = (q_k)_{k\in\mathbb{N}}$, cosa que podem fer perquè $\mathbb{Q}$ és un conjunt numerable, i defineixo els conjunts traslladats següents:

$$V_k = V + q_k.$$

Observem que si $x \in E$, aleshores $x \simeq v$ per a algun $v \in V$, atès que V conté un element de cada classe d'equivalència. Aleshores, $x - v \in \mathbb{Q}$, és a dir $x - v = q_k$ per a algun k i conseqüentment $x \in V_k$. Per tant, deduïm que

$$E \subseteq \bigcup_{k\in\mathbb{N}} V_k.$$

Si V fos mesurable, cada V_k seria mesurable i $m(V_k) = m(V)$ (per la invariància per translacions de la mesura de Lebesgue). Per tant, per la σ-subadditivitat, tenim:

$$m(E) \leq \sum_{k=1}^{\infty} m(V_k) = \sum_{k=1}^{\infty} m(V).$$

Si $m(V) = 0$, tindríem que $m(E) = 0$, en contra de la hipòtesi inicial.

Per tant, ens queda analitzar què passaria si $m(V) > 0$. En aquest cas, $D(V)$ hauria de contenir un interval de la forma $(-\delta, \delta)$ per a algun $\delta > 0$ pel teorema 39. Però per definició de V, tenim que

$$D(V) \cap \mathbb{Q} = \{0\}.$$

Aquest fet és una contradicció, atès que per la densitat dels nombres racionals a l'interval $(-\delta, \delta)$ hi han d'haver infinits nombres racionals.

Hem arribat a una contradicció que provenia de suposar que V era mesurable, Per tant, V no pot ser mesurable. □

Observació. El mateix argument que la demostració anterior ens permet demostrar que qualsevol subconjunt mesurable d'un conjunt de Vitali té mesura 0.

4. Paradoxa de Banach-Tarski

Una de les conseqüències més sorprenents de l'axioma de l'elecció és la paradoxa de Banach-Tarski, que en la seva versió clàssica més coneguda ens diu, de manera informal, que podem tallar una bola unitat de l'espai euclidià en un nombre finit de trossos i reordenar-los i enganxar-los de manera que podem obtenir dues boles unitat de la mateixa mida (volum, mesura Lebesgue) que l'original.

En aquest apartat demostrarem tal com ho va fer originalment Felix Hausdorff usant menys eines de teoria de grups que les que es fan servir habitualment a les demostracions més recents. A l'Apèndix he incorporat un capítol amb demostració alternativa (amb més o menys totes les demostracions dels resultats previs) més moderna que utilitza el fet que a $\mathcal{SO}(3)$ existeix un subgrup lliure isomorf a $\mathbb{F}_2$ i utilitzant teoremes de grups paradoxals. En canvi, aquí construïrem una descomposició de l'esfera en un nombre finit de conjunts construïts a partir d'un subgrup de $\mathcal{SO}(3)$ que no és lliure però també ens permet arribar a la paradoxa de Banach-Tarski, un *grup modular* $\Gamma := \langle u, v \rangle$ generat per dos elements u, v tals que $u^2 = v^3 = 1$, de manera que $\Gamma \cong C_2 * C_3$, on $*$ denota el producte de la categoria de grups.

A continuació, anem a donar algunes definicions, enunciar i demostrar alguns lemes i teoremes previs que ens permetran demostrar aquesta versió de la paradoxa:

Lema 41. *Sigui G el producte lliure de grups $\{1, \phi\}$ i $\{1, \psi, \psi^2\}$, és a dir, el grup de tots els productes formals per ϕ, ψ i ψ^2, amb l'especificació que $\phi^2 = 1$ i $\psi^3 = 1$. Considerem dos eixos de rotació (diferents) a_ϕ i a_ψ que atravessen el centre de la bola unitat $U \subseteq \mathbb{R}^3$, i interpreto G com el grup de totes les rotacions generades per la rotació ϕ de π radians sobre l'eix a_ϕ i la rotació ψ de $\frac{2\pi}{3}$ radians sobre l'eix a_ψ. Aleshores podem determinar els eixos a_ϕ i a_ψ de manera que elements diferents de G representen rotacions diferents generades per ϕ i ψ, és a dir, l'acció del grup G sobre l'esfera U és fidel.*

Demostració: N'hi ha prou buscant un angle θ entre els eixos a_ϕ i a_ψ de manera que cap element diferent de la identitat de G representi la rotació identitat (atès que es tindria la igualtat si i només i dos elements diferents de G diferents de la identitat fossin iguals, al multiplicar pels inversos d'alguna(/es) de les components de la paraula $\alpha = I$ a banda i banda de la igualtat). Prenem a_ψ com l'eix Z i a_ϕ de manera que pertanyi al pla XZ formant un angle θ amb l'eix a_ψ.

Aleshores les rotacions ψ i ϕ vénen representades per les matrius següents:

$$\psi = \begin{bmatrix} \lambda & \mu & 0 \\ -\mu & \lambda & 0 \\ 0 & 0 & 1 \end{bmatrix};$$

$$\phi = \begin{bmatrix} -\cos(\theta) & 0 & \sin(\theta) \\ 0 & -1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{bmatrix}$$

on $\lambda = \cos\frac{2\pi}{3} = -\frac{1}{2}$ i $\mu = \sin\frac{2\pi}{3} = \frac{\sqrt{3}}{2}$.

Escollim θ de manera que $\cos(\theta)$ sigui irracional. Volem veure que si $\alpha \in G$ diferent de la identitat, aleshores la seva acció corresponent no és la identitat.

Si $\alpha = \sigma_m\sigma_{m-1}\dots\sigma_2\sigma_1$ on cada σ_i és o bé de la forma $\psi\phi$ o bé de la forma $\psi^2\phi$, aleshores l'acció de cadascuna de les σ_i ve representada per una de les dues matrius següents:

$$\sigma = \begin{bmatrix} -\lambda\cos(\theta) & \mp\mu & \lambda\sin(\theta) \\ \pm\cos(\theta) & -\lambda & \mp\mu\sin(\theta) \\ \sin(\theta) & 0 & \cos(\theta) \end{bmatrix}.$$

Sigui $K = (0,0,1)$ l'eix z. Tenim que

$$\alpha \cdot K = \sigma_m\sigma_{m-1}\dots\sigma_1 \cdot K = (\sin(\theta)P_m(\cos(\theta)), \sqrt{3}\sin(\theta)Q_m(\cos(\theta)), R_m(\cos(\theta))),$$

on P_m, Q_m i R_m són polinomis amb coeficients racionals. De fet, tenim que $P_1(x) = -\frac{1}{2}$, $Q_1(x) = \mp\frac{1}{2}$ i $R_1(x) = x$ i

$$P_{m+1}(x) = -\lambda x P_m(x) \mp \frac{3}{2}Q_m(x) + \lambda R_m(x);$$
$$Q_{m+1}(x) = \pm\frac{1}{2}xP_m(x) - \lambda Q_m(x) \pm \frac{1}{2}R_m(x);$$
$$R_{m+1}(x) = (1-x^2)P_m(x) + xR_m(x).$$

Com que $\cos(\theta)$ és irracional, no és l'arrel de cap polinomi amb coeficients racionals. Per tant, $\alpha \cdot K \neq K$, ja que altrament tindríem que $R_m(\cos(\theta)) - 1 = 0$ a l'igualtat les terceres components dels vectors i aquesta equació hauria de tenir una solució, que no existeix per la hipòtesi d'haver-se escollit $\cos(\theta)$ irracional.

Els casos en els quals α és de les formes $\phi\sigma_m\dots\sigma_1, \sigma_m\dots\sigma_1\psi^{\pm}$ i $\phi\sigma_m\dots\sigma_1\psi^{\pm}$ es dedueixen d'aquest darrer cas de forma anàloga, obtenint el que volíem demostrar. Per tant, l'acció del grup G sobre l'esfera és fidel. □

LEMA 42. *Sigui G el grup determinat per les hipòtesis del lema 41 i amb els eixos escollits de manera que elements diferents de G representin rotacions diferents generades per ϕ i ψ. Aleshores el grup G es pot descompondre en tres conjunts disjunts*

$$A \cup B \cup C$$

de manera que $\phi(A) = B \cup C$, $\psi(A) = B$ i $\psi^2(A) = C$.

Demostració: Construïm els conjunts A, B, C recursivament a partir de les llargades dels elements de G. Siguin $1 \in A$, $\phi, \psi \in B$ i $\psi^2 \in C$ i, aleshores, procedim com a continuació per a qualsevol $\alpha \in G$:

Sabem que $\alpha \neq 1$ acaba o bé amb $\psi^{\pm 1}$ o bé amb ϕ. En el primer cas, estarem a la primera fila, mentre que en el segon cas estarem a la segona i tercera fila:

$\alpha \in A$	$\alpha \in B$	$\alpha \in C$
$\alpha\phi \in B$	$\alpha\phi \in A$	$\alpha\phi \in A$
$\alpha\psi \in B$	$\alpha\psi \in C$	$\alpha\psi \in A$
$\alpha\psi^{-1} \in C$	$\alpha\psi^{-1} \in A$	$\alpha\psi^{-1} \in B$

(No calen fer més comprovacions, atès que si α acaba amb ϕ ja sabem que a l'aplicar-hi ϕ ens quedarà α i ja sabrem a quin conjunt pertany, mentre que la resposta a les expressions amb α acabant amb $\psi^{\pm 1}$ a la segona i tercera fila també les podrem trobar dins de la taula, atès que sabem que $\psi^3 = 1$.)

Per tant, en aquest punt tenim construïts els conjunts A, B i C definits de manera recursiva a l'haver contemplat tots els casos.

D'aquesta manera, la condició $\phi(A) = B \cup C$, $\psi(A) = B$ i $\psi^2(A) = C$ de l'enunciat del lema estarà satisfeta, fent totes les comprovacions i tota la casuística possible a partir de la taula. Per exemple, si $\alpha \in A$ i acaba amb $\psi^{\pm}$, aleshores $\phi(\alpha) = \alpha\phi \in B$, si acaba amb ϕ aleshores o bé és $\alpha = \phi \in B$ o bé té una penúltima lletra que per força ha de ser $\psi^{\pm}$ per tractar-se d'una expressió reduïda i, en el cas que sigui ψ^+, tindrem que $\phi(\alpha) \in B$ per la taula mentre que en el cas que sigui ψ^{-1} tindrem que $\phi(\alpha) \in C$ per la taula, i l'últim cas que queda és $\alpha = 1$ de manera que $\phi(\alpha) = \phi \in B$, de manera que $\phi(A) = B \cup C$ (i els altres casos es comproven de forma anàloga). Efectivament, les condicions que hem demostrat al lema 41 juntament amb la manera com hem construït els elements dels conjunts ens garanteixen que els conjunts A, B i C són disjunts amb $G = A \cup B \cup C$, d'on obtenim el que volíem demostrar. □

Definicions. Donats $A, B \subseteq \mathbb{R}^3$ subconjunts de l'espai euclidià tridimensional, direm que A i B són *congruents* si existeix una isometria $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ (element del grup euclidià $E(3)$) tal que $f(A) = B$. En aquest cas, denotarem $A \simeq B$.

Direm que X i Y són *equidescomponibles* si existeix una descomposició finita de X en conjunts disjunts $X_1, \ldots, X_n$ de manera que

$$X = X_1 \cup \ldots \cup X_n$$

i una descomposició de Y en el mateix nombre de conjunts disjunts

$$Y = Y_1 \cup \ldots \cup Y_n$$

de manera que $X_i \simeq Y_i$ per a tot $i = 1, \ldots, n$. En tal cas denotarem $X \approx Y$.

Ara ja estem preparats per demostrar una teorema previ a la paradoxa de Banach-Tarski, que se sol anomenar la paradoxa de Hausdorff:

Proposició 43. *(Paradoxa de Hausdorff) Existeix una descomposició disjunta de l'esfera $\mathbb{S}^2$ en quatre conjunts A, B, C, Q tals que $A, B, C, B \cup C$ són congruents i Q és enumerable.*

Demostració: Sigui Q el conjunt de tots els punts fixos de l'esfera S de totes les rotacions $\alpha \in G$. Cada rotació $\alpha \in G$ té dos punts fixos; per tant, Q és numerable, atès que tenim una quantitat numerable de rotacions. El conjunt $S \setminus Q$ és unió disjunta de totes les òrbites P_x del grup G:

$$P_x = \{x\alpha \mid \alpha \in G\}.$$

Aleshores, per l'axioma de l'elecció, sabem que existeix un conjunt M que conté exactament un element de P_x per a tot $x \in S \setminus Q$. Fent una partició de G com al lema 42 amb $G = \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ i $\phi(\mathcal{A}) = \mathcal{B} \cup \mathcal{C}$, $\psi(\mathcal{A}) = \mathcal{B}$ i $\psi^2(\mathcal{A}) = \mathcal{C}$, aleshores puc construir $A := M \cdot \mathcal{A}$, $B := M \cdot \mathcal{B}$ i $C := M \cdot \mathcal{C}$, de manera que per les condicions del lema sabem que $A, B, C \subseteq S$ són disjunts, congruents els uns amb els altres i amb $B \cup C$ per ser ϕ, ψ, ψ^2 isometries i, per construcció,

$$S = A \cup B \cup C \cup Q,$$

tal com volíem demostrar. □

Anem a veure alguns lemes previs sobre equidescomponibilitat i ja podrem demostrar la paradoxa de Banach-Tarski.

Lema 44. *Sigui $\approx$ la relació d'equidescomponibilitat. Aleshores:*

(1) $\approx$ és una relació d'equivalència.
(2) Si X i Y són unió disjuntes de dos conjunts, és a dir, $X = X_1 \sqcup X_2$ i $Y = Y_1 \sqcup Y_2$ i $X_i \approx Y_i$ per a $i = 1, 2$, aleshores $X \approx Y$.

(3) Si $X_1 \subseteq Y \subseteq X$ i si $X \approx X_1$, aleshores $X \approx Y$.

Demostració:

(1) Senzill de demostrar (comprovar que és reflexiva, simètrica i transitiva).
(2) També és senzill de demostrar utilitzant les reunions dels trossos en els quals els conjunts descomponen X_1, X_2, Y_1, Y_2 descomponen i considerant les seves reunions (nombre finit de conjunts).
(3) Sigui $X = X^1 \cup \ldots \cup X^n$ i $X_1 = X_1^1 \cup \ldots X_1^n$ tals que $X^i \simeq X_1^i$ per a tot $i = 1, \ldots, n$. Escollim les isometries que ens donin aquestes congruències $f_i \colon X^i \to X_1^i$ per a tot $i = 1, \ldots, n$, i sigui f l'aplicació bijectiva corresponent de X a X_1 que coincideix amb f_i en cada X_i. Ara, siguin

$$X_0 = X,\ X_1 = f''X_0, \ldots,\ X_n = f''X_{n-1}, \ldots;$$
$$Y_0 = Y,\ Y_1 = f''Y_0,\ \ldots,\ X_n = f''Y_{n-1},\ \ldots,$$

on $f''X := \{f(x) \mid x \in X\}$ és la imatge de X per f. Si fem

$$Z = \cup_{n=0}^{\infty}(X_n - Y_n),$$

aleshores $f''Z$ i $X - Z$ són disjunts per construcció, $Z \approx f''Z$ i

$$X = Z \cup (X - Z),\ Y = f''Z \cup (X - Z)$$

també per construcció, i per tant $X \approx Y$ utilitzant l'apartat (2). □

Ara ja estem preparats per demostrar la versió clàssica de la paradoxa de Banach-Tarski:

Proposició 45. *(Paradoxa de Banach-Tarski) La bola unitat $U \subseteq \mathbb{R}^3$ és equidescomponible com a unió disjunta de dues boles unitat, és a dir, $U = X \cup Y$ amb $U \approx X$ i $U \approx Y$.*

Demostració: Sigui U una bola tancada i sigui

$$S = A \cup B \cup C \cup Q$$

la descomposició de la seva superfície trobada a la proposició 43 (paradoxa de Hausdorff). Aleshores tenim que

$$U = \bar{A} \cup \bar{B} \cup \bar{C} \cup \bar{Q} \cup \{c\},$$

on c és el el centre de l'esfera i per a tot $X \subseteq S$, $\bar{X}$ és el conjunt de tots els $x \in U$, diferents de c, tals que la seva projecció a la superfície és a X. Clarament,

$$\bar{A} \approx \bar{B} \approx \bar{C} \approx \bar{B} \cup \bar{C},$$

atès que si podem trobar aquestes isometries per punts de l'esfera també les podem estendre a isometries entre segments que uneixen aquests punts amb el centre de l'esfera (tots ells isomètrics a intervals semioberts i semitancats de $\mathbb{R}$). Siguin $X := \bar{A} \cup \bar{Q} \cup \{c\}$ i $Y := U - X$. Aleshores, per l'observació anterior, el lema 44 (primera part, que ens diu que és una relació d'equivalència) i el fet que $\bar{A} \cup \bar{B} \approx \bar{A} \cup \bar{B} \cup \bar{C}$ sabem que

$$\bar{A} \approx \bar{A} \cup \bar{B} \cup \bar{C},$$

i per tant deduïm que $X \approx U$.

Ara, atès que Q és numerable mentre que $\mathcal{SO}_3$ no ho és, és fàcil buscar una rotació α que no sigui de G de manera que $\alpha(Q)$ i Q siguin disjunts i, per tant, utilitzant que

$$\bar{C} \approx \bar{A} \cup \bar{B} \cup \bar{C}$$

(ho podem deduir per transitivitat amb $\bar{A} \approx \bar{A} \cup \bar{B} \cup \bar{C}$), existeix $S \subseteq C$ tal que $\bar{S} \approx \bar{Q}$. Sigui $p \in \bar{C} \setminus \bar{S}$ (no buit, atès que $\bar{C}$ no és numerable mentre que $\bar{S}$ sí que ho és). Òbviament,

$$\bar{A} \cup \bar{Q} \cup \{c\} \approx \bar{B} \cup \bar{S} \cup \{p\}.$$

Com que

$$\bar{B} \cup \bar{S} \cup \{p\} \subseteq Y \subseteq U,$$

podem utilitzar la definició de X i que $X = \bar{A} \cup \bar{Q} \cup \{c\} \approx \bar{B} \cup \bar{S} \cup \{p\}$ juntament amb el lema 44.3 per obtenir $Y \approx U$. Amb això acabem la demostració del teorema. □

5. Existència de clausura algebraica

En aquest apartat demostrarem que tot cos admet una (única llevat d'isomorfisme) clausura algebraica.

Recordem la definició següent:

Definició. Sigui F un cos. Una *clausura algebraica de* F és un cos algebraicament tancat C que és algebraic sobre F; és a dir, tot polinomi no constant té una arrel i tot element de C és una arrel d'un polinomi de $F[x]$.

PROPOSICIÓ 46. *Tot cos F té una clausura algebraica.*

Demostració: Sigui

$$S := \{(k, a_0, \ldots, a_n, 0, \ldots) \in \mathbb{N} \times F \times F \times \ldots \mid a_i \in F,\ 1 \leq k \leq n\}.$$

Qualsevol extensió algebraica E de F té com a molt tants elements com el conjunt S, atès que tota arrel $\alpha \in E$ és una arrel d'algun polinomi $f(x) = a_0 + \ldots + a_n x^n \in F[x]$, que té com a molt n arrels diferents a E. Per tal d'aconseguir encara més elements, puc agafar el conjunt potència $\mathcal{P}(S)$ de S i recordar que el seu cardinal satisfà $|\mathcal{P}(S)| > |S|$. Com que la funció $f\colon F \to \mathcal{P}(S)$ donada per $f(a) = \{(1, a, -1, 0, 0, \ldots)\}$ és injectiva, puc considerar el conjunt $\Omega := (\mathcal{P}(S) \setminus f[F]) \cup F$ per tal que $F \subseteq \Omega$ i $|\Omega| > |S|$.

Donat un cos $(E, +, \cdot)$ amb $E \subseteq \Omega$, la suma és una funció $f\colon E \times E \to E$, que és un subconjunt de $(E \times E) \times E \subseteq (\Omega \times \Omega) \times \Omega = \Omega^3$. El mateix es pot dir de la multiplicació. Per tant, podem definir el conjunt $\mathcal{E}$ de totes les extensions algebraiques E de F amb $E \subseteq \Omega$ com

$$\mathcal{E} = \left\{(E, +, \cdot) \in \mathcal{P}(\Omega) \times \mathcal{P}(\Omega^3) \times \mathcal{P}(\Omega^3) \mid (E, +, \cdot) \text{ és una extensió algebraica de } F\right\}.$$

Observem que el conjunt $\mathcal{E}$ és no buit, atès que conté el cos F.

Podem definir la següent relació d'ordre parcial a $\mathcal{E}$: per a tot $E_1, E_2 \in \mathcal{E}$, $E_1 \leq E_2$ si i només si E_2 és una extensió algebraica de E_1. Aleshores tota cadena $\mathcal{C} \subseteq \mathcal{E}$ té una fita superior K a $\mathcal{E}$, atès que si $\mathcal{C}$ és una cadena no buida de $\mathcal{E}$, aleshores podem agafar $K := \bigcup_{E \in \mathcal{C}} E$ juntament amb les operacions binàries tal com les definim a continuació: Si $\alpha, \beta \in K$, aleshores $\alpha \in E_1$ i $\beta \in E_2$ per a certs $E_1, E_2 \in \mathcal{C}$. Com que $\mathcal{C}$ és una cadena, aleshores o bé E_1 és un subcos de E_2 o bé E_2 és un subcos de E_1 (sense pèrdua de generalitat, podem suposar que E_1 és un subcos de E_2). En qualsevol dels casos, existeix un $E \in \mathcal{C}$ tal que $\alpha, \beta \in E \subseteq K$ i podem definir $\alpha + \beta$ i $\alpha \cdot \beta$ com el resultat de l'operació corresponent al conjunt $(E, +, \cdot)$. Aquestes definicions són independents de la tria del conjunt E, atès que si E' és un altre element de la cadena $\mathcal{C}$ amb $\alpha, \beta \in E'$, aleshores o bé E és un subcos de E' o viceversa. Per tant $\alpha + \beta$ és el mateix element a E que a E' i, per tant, és un element que està ben definit a K, i anàlogament això passa amb l'element $\alpha \cdot \beta$. Com que tots aquests càlculs de K es fan en algun cos $E \in \mathcal{C}$, deduïm que aquestes operacions doten el conjunt K d'una estructura de cos. De fet, K és una extensió algebraica de F, atès que tot element $\alpha \in K$ pertany a alguna extensió algebraica $E \in \mathcal{C}$ de F i és, per tant, algebraica sobre F. Per tant, deduïm que $(K, +, \cdot) \in \mathcal{E}$ i que $E \leq K$ per a tot $\mathcal{C}$, és a dir, K és una fita superior de la cadena $\mathcal{C}$ en $\mathcal{E}$.

Ara podem aplicar el lema de Zorn per garantir l'existència d'un element maximal $\bar{F}$ de $\mathcal{E}$. Per definició de $\mathcal{E}$, $\bar{F}$ és una extensió algebraica de F i $\bar{F} \subseteq \Omega$. Tant sols queda per veure que $\bar{F}$ és algebraicament tancat. En efecte, sigui $f(x) \in \bar{F}[x]$ un polinomi no constant i suposem, per reducció a l'absurd, que $f(x)$ no tingués cap arrel a $\bar{F}$. Aleshores sabem de l'assignatura d'Estructures Algebraiques que existeix una extensió finita E de $\bar{F}$ i un element $\alpha \in E$ tal que $f(\alpha) = 0$. Com que E és una extensió finita de F, aleshores E és una extensió algebraica de F.

Com que $|\bar{F}| \leq |E| \leq |S| < |\Omega|$, tenim que $|\Omega| = |\Omega \setminus \bar{F}|$ pel lema 13 aplicat als conjunts $A = \Omega$, $B = \bar{F}$ i $C = \Omega \setminus \bar{F}$. Per tant, tenim que $|E \setminus \bar{F}| \leq |E| < |\Omega| = |\Omega \setminus \bar{F}|$, de manera que existeix una funció injectiva $g\colon E \to \Omega$ que estén la funció inclusió $i\colon \bar{F} \to \Omega$. Reescrivint les taules d'addició i multiplicació del cos E a la seva imatge $g[E] \subseteq \Omega$, mentre es mantenen les operacions de F tal com les tenim definides, podem assumir que $E \subseteq \Omega$ i $E \in \mathcal{E}$, en primer lloc. Com que $f(x)$ no té cap arrel a $\bar{F}$, tenim que $\alpha \notin \bar{F}$ de manera que $\bar{F} \subsetneq E$, contradient la maximalitat de $\bar{F}$. □

Observació. Es pot demostrar que aquesta clausura algebraica és única llevat d'isomorfismes de cossos aplicant el lema de Zorn a la família de tots els isomorfismes parcials donats entre dues clausures algebraiques de F.

Capítol 5
Versions febles de l'Axioma de l'elecció i algunes aplicacions

En aquest capítol enunciarem i demostrarem algunes de les versions febles de l'Axioma de l'elecció i demostrarem algun teorema que resulta com a aplicació d'aquestes versions més febles.

1. Axioma de les eleccions dependents. Aplicació al teorema de la base de Hilbert

Anem a enunciar i demostrar l'Axioma de les eleccions dependents a partir de l'Axioma de l'elecció:

PROPOSICIÓ 47. *(Axioma de les eleccions dependents) Sigui $X \neq \emptyset$ un conjunt i $R \subseteq X \times X$ una relació satisfent la següent propietat:*

$$\forall x \in X \ \exists y \in X \mid (x, y) \in R.$$

Aleshores existeix una successió

$$(x_k)_{k\in\mathbb{N}} \mid (x_k, x_{k+1}) \in R.$$

Demostració: Per l'enunciat de la proposició, tenim $X \neq \emptyset$ un conjunt i $R \subseteq X \times X$ una relació satisfent la següent propietat:

$$\forall x \in X \ \exists y \in X \mid (x, y) \in R.$$

Per a tot element $x \in X$, defineixo el conjunt $R(x) := \{y \in S \mid xRy\}$. Per hipòtesi, $R(x)$ és un conjunt no buit per a tot $x \in X$. Sigui $\mathcal{H} := \{R(x) \mid x \in X\}$ la família formada per tots aquests conjunts. Aleshores, l'Axioma de l'elecció ens permet afirmar que existeix una funció $f\colon \mathcal{H} \to X$ tal que $f(H) \in H$ per a tot $H \in \mathcal{H}$. Si defineixo $i\colon X \to \mathcal{H}$ com la funció $i(x) := R(x)$, aleshores puc definir $g\colon X \to X$ per $g(x) := (f \circ i)(x)$, de manera que $g(x) \in R(x)$ per a tot $x \in X$. Per tant, tenim que $xRg(x)$. Aleshores, per a tot $x \in X$, puc definir la successió $(x_n)_{n\geq 1} \subseteq X$ tal que $x_n := g^n(x)$, on g^n denota la composició de g amb ella mateixa n vegades. Per construcció, aquesta successió satisfà que $x_n R x_{n+1}$ per a tot $n \in \mathbb{N}$, tal com volíem demostrar. □

Aquesta versió té com a aplicacions, per exemple, els següents teoremes importants d'Àlgebra Commutativa: la caracterització que ens permet definir els anells noetherians i el teorema de la base de Hilbert. Anem-los a veure:

PROPOSICIÓ 48. *Sigui A un anell. Aleshores les següents condicions són equivalents:*

(1) Tot ideal $I \subseteq A$ és finitament generat, és a dir, per a tot ideal $I \subseteq A$, existeixen $f_1, \ldots, f_k \in I$ tals que $I = (f_1, \ldots, f_k)$.
(2) Tot cadena ascendent

$$I_1 \subsetneq \ldots \subsetneq I_m \subseteq \ldots$$

d'ideals de A és estacionària, és a dir, existeix $N > 0$ natural tal que $I_N = I_{N+1} = \ldots$ (que anomenem condició de cadena ascendent).

(3) Tot subconjunt no buit d'ideals de A té un element maximal.

Demostració:

- (1) $\Rightarrow$ (2) Donat $I_1 \subseteq \ldots \subseteq I_m \subseteq \ldots$, sigui $I = \cup_{m\geq 1} I_m$. Aleshores clarament I segueix sent un ideal (vist a Estructures Algebraiques a la demostració del teorema de Krull). Si $I = (f_1, \ldots, f_k)$, aleshores cada f_i és un element d'algun I_{m_i} per a algun m_i, de manera que prenent $m := \max\{m_i\}$ obtenim $I = I_m$ i, per tant, la cadena estaciona a I_m.
- (2) $\Rightarrow$ (3) Suposem per reducció a l'absurd que la condició (3) és falsa i sigui

$$\Sigma := \{I \neq \emptyset \mid I \subseteq A,\ I \text{ no és maximal}\}.$$

  Aleshores, si considero $X = \Sigma$ i $R \subseteq X \times X$ la relació $\subsetneq$, es verifiquen les hipòtesis de l'axioma de les eleccions dependents. En efecte, per definició de Σ, tenim que si $I_i \in \Sigma$, com que I_i no és maximal, existeix $I_k \in \Sigma$ tal que $I_i \subsetneq I_k$. Per tant, utilitzant l'axioma de les eleccions dependents 47 i la negació de l'enunciat (3), tenim que existeix una cadena ascendent $(I_k)_{k\in\mathbb{N}}$ tal que

$$\forall k \in \mathbb{N}, I_k \in \Sigma \text{ i } I_k \subsetneq I_{k+1}.$$

  Per tant, hem trobat una cadena ascendent no estacionària, contradient la hipòtesi (2).
- (3) $\Rightarrow$ (1) Sigui I un ideal i sigui $\Sigma := \{J \subseteq I \mid J \text{ és ideal finitament generat}\}$. Aleshores, (3) ens diu que Σ té un element maximal, diguem-ne J_0. Però aleshores $J_0 = I$, atès que si no fos així tindríem que qualsevol $f \in I \setminus J_0$ ens permet construir l'ideal $J_0 + Af$ que segueix sent finitament generat, però és estrictament més gran que J_0. Per tant, qualsevol ideal $I \subseteq A$ és finitament generat, tal com volíem demostrar. □

Definició. Si A és un anell i es compleix alguna de les condicions de la proposició anterior 48, direm que A és un *anell noetherià*.

Anem a enunciar una proposició sense demostrar (extreta de

https://homepages.warwick.ac.uk/staff/Miles.Reid/MA4A5/UAG.pdf, Proposició 3.2, pàgina 58):

PROPOSICIÓ 49. *(1) Sigui A un anell noetherià i $I \subseteq A$ un ideal. Aleshores l'anell quocient $B := A/I$ és noetherià.*

(2) Sigui A un anell noetherià que sigui domini d'integritat, amb $A \subseteq K$ el seu cos de fraccions. Sigui $0 \notin S \subseteq A$ un subconjunt de A i defineixo

$$B := A[S^{-1}] = \left\{\frac{a}{b} \in K \mid a \in A,\ b = 1 \text{ o un producte d'elements de } S\right\}.$$

Aleshores B és noetherià.

Anem a enunciar i demostrar el teorema de la base de Hilbert:

PROPOSICIÓ 50. *Sigui A un anell. Aleshores, si A és noetherià, $A[X]$ també és noetherià.*

Demostració: Sigui $J \subseteq A[X]$ un ideal qualsevol de $A[X]$. Per demostrar que $A[X]$ és noetherià, només fa falta veure que J és finitament generat. En efecte, si defineixo

$$J_n := \left\{a \in A \mid \exists f = aX^n + b_{n-1}X^{n-1} + \ldots + b_0 \in J\right\}.$$

Aleshores és senzill comprovar que J_n és un ideal de A (comprovant les propietats d'ideal) i $J_n \subseteq J_{n+1}$ (multiplicant l'equació de la condició dels elements de J_n per X i comprovant tots els elements de J_n són de J_{n+1} posant com a condició que aquest producte per X sigui de l'ideal). Per tant, utilitzant la condició de cadena ascendent enunciada a la proposició 48 que caracteritza els anells noetherians, tenim que existeix $N > 0$ tal que

$$J_N = J_{N+1} = \ldots.$$

Ara construïm un sistema de generadors de J de la manera següent: per a tot $i \leq N$, $i \geq 0$, siguin $a_{i_1}, \ldots, a_{i_{m_i}}$ generadors de J_i i, com a la definició de J_i, per a cadascun dels a_{i_k}, sigui $f_{i_k} = a_{i_k}X^i + \ldots \in J$ un element de grau i amb terme dominant a_{i_k} (que sempre es pot aconseguir per definició dels elements de J_i).

Anem a veure que el conjunt

$$\{f_{i_k} \mid i = 0, \ldots, N, k = 1, \ldots, m_i\}$$

acabat de construir genera l'ideal J. En efecte, sigui $g \in J$ i suposem que $\deg(g) = m$. Aleshores el terme dominant de g és bX^m amb $b \in J_m$ i, pel que sabem sobre J_m, podré escriure b com la combinació lineal $b = \sum c_{m'k}a_{m'k}$, de manera que $m' = m$ si $m \geq N$ mentre que $m' = N$ altrament. Aleshores defineixo $g_1 := g - X^{m-m'} \cdot \sum c_{m'k}f_{m'k}$. Per construcció, el terme de grau m s'anul·la, de manera que $\mathrm{gr}(g_1) \leq \mathrm{gr}(g) - 1$. Per tant, per inducció -el cas base és trivial, atès que l'ideal $J_0 \subseteq A$ és finitament generat per $f_{0_k} = a_{0_k}$ -, puc escriure g (aïllant a l'equació anterior) com a combinació lineal de f_{i_k}, de manera que en efecte generen J. □

Corol·lari 51. *Si K és un cos, aleshores una K-àlgebra finitament generada és un anell noetherià.*

Demostració: Una K àlgebra finitament generada és un anell de la forma $A = K[a_1, \ldots, a_n]$, de manera que A està generat com a anell pel cos K i $a_1, \ldots, a_n$. Clarament, pel primer teorema d'isomorfisme, tenim que $A \cong K[X_1, \ldots, X_n]/I$, essent I l'ideal generat pels polinomis que anul·len $a_1, \ldots, a_n$ (nucli del morfisme d'anells avaluació en $X_1 = a_1, \ldots, X_n = a_n$). Tot cos és noetherià (finitament generat per l'element neutre) i, per tant, aplicant el teorema de la base de Hilbert 1 n vegades per inducció, tenim que $K[X_1, \ldots, X_n]$ també és noetherià, de manera que, al passar al quocient, podem aplicar la proposició 49 per demostrar que A també és noetherià. □

Capítol 6
Apèndix

1. Conseqüències en anàlisi funcional del teorema de Hahn-Banach: versions geomètriques

Anem a veure uns quants resultats conseqüència del teorema de Hahn-Banach. en particular les seves versions geomètriques.

Corol·lari 52. *Siguin E un espai vectorial, F un subespai vectorial de E i $p\colon E \to [0,\infty)$ una seminorma. Sigui $f\colon F \to \mathbb{R}$ lineal tal que $|f(x)| \leq p(x)$ per a tot $x \in F$. Aleshores, existeix $g\colon E \to \mathbb{R}$ lineal i tal que $g_{|F} = f$ i $|g(x)| \leq p(x)$.*

Demostració: Com que p és una seminorma, en particular és un funcional convex i també un funcional sublineal, i tenim $f(x) \leq |f(x)| \leq p(x)$ per a tot $x \in F$. Aplicant el teorema de Hahn-Banach, tenim que existeix $g\colon E \to \mathbb{R}$ lineal tal que $g_{|F} = f$ i $g(x) \leq p(x)$ per a tot $x \in E$. A més a més, $-g(x) = g(-x) \leq p(-x) = |-1|p(x) = p(x)$, on a la segona igualtat hem utilitzat que p és una seminorma. Per tant, $g(x) \leq p(x)$ i $g(-x) \leq p(x)$ per a tot $x \in E$, d'on en deduïm que $|g(x)| \leq p(x)$ per a tot $x \in E$. □

Corol·lari 53. *Sigui E espai normat, $E \neq \{0\}$ i F un subespai vectorial de E. Sigui $f\colon F \to \mathbb{R}$ lineal i contínua. Aleshores, existeix $g\colon E \to \mathbb{R}$ lineal, contínua i tal que $g_{|F} = f$ i $||g|| = ||f||$.*

Demostració: Sigui $p(x) := ||f||||x||$. Com que f és lineal i contínua, aleshores $|f(x)| \leq ||f||||x|| = p(x)$ per a tot $x \in F$ (on $||f|| \in \mathbb{R}, ||f|| \geq 0$). Definim $p\colon E \to \mathbb{R}$, $p(x) = ||f||||x||$ per a tot $x \in E$. L'aplicació p és una seminorma, ja que $p(x) \geq 0$, $p(x+y) = ||f||||x+y|| \leq ||f||(||x|| + ||y||) = p(x) + p(y)$ i $p(\lambda x) = ||f||||\lambda x|| = \lambda||f||||x|| = |\lambda|p(x)$. Per tant, es compleixen les hipòtesis del corol·lari 52 i tenim que existeix $g\colon E \to \mathbb{R}$ lineal tal que $g_{|F} = f$ i $|g(x)| \leq p(x) := ||f||||x||$ per a tot $x \in E$. Per tant, $g \in E \to \mathbb{R}$ és lineal, contínua i $||g|| \leq ||f||$, ja que $|g(x)| \leq p(x) = ||f||||x||$ implica que és contínua (per definició) i que $|g(x)|/||x|| \leq ||f||$ i prenent el suprem del terme de l'esquerra de la desigualtat per a les $x \in E$, en deduïm $||g|| \leq ||f||$. Per a veure que es té la igualtat de normes, fem el següent càlcul:

$$||f|| = \sup_{||x||\leq 1, x\in F}|f(x)| = \sup_{||x||\leq 1, x\in F}|g(x)| \leq \sup_{||x||\leq 1, x\in E}|g(x)| = ||g||.$$

Per tant, $||f|| = ||g||$, tal com volíem veure. □

Corol·lari 54. *Sigui E un espai normat, $E \neq \{0\}$ i $x_0 \in E$. Aleshores existeix $g\colon E \to \mathbb{R}$ lineal i contínua tal que $g(x_0) = ||x_0||$, $||g|| = 1$.*

Demostració:

- Si $x_0 \neq 0$, definim $F := [x_0]$ com a subespai vectorial de E i, escrivint $y = \lambda x_0$ per a tot $y \in F$ (escrit en la base $[x_0]$), definim l'aplicació $f\colon F \to \mathbb{R}$ com $f(y) = f(\lambda x_0) := \lambda||x_0||$, lineal per construcció. Tenim que $|f(\lambda x)| = |\lambda||x||| = |\lambda|||x|| = ||\lambda x||$ i, per tant, f és contínua amb $||f|| = 1$. Aleshores,

aplicant el corol·lari 53, existeix $g\colon E \to \mathbb{R}$ lineal i contínua de manera que $g_{|F} = 1$ i $||g|| = ||f|| = 1$, tal com volíem veure.

- Si $x_0 = 0$, aleshores agafem $y_0 \in E \setminus \{0\} \neq \emptyset$. Ara apliquem el mateix argument que al punt anterior canviant x_0 per y_0, i en deduïm que existeix $g\colon E \to \mathbb{R}$ lineal i contínua tal que $g(y_0) = ||y_0||$ i $||g|| = 1$. Aquesta mateixa funció, pel fet de ser lineal, satisfà que $g(x_0) = g(0) = 0 = ||0|| = ||x_0||$ i, per tant, també satisfà el que volíem. □

COROL·LARI 55. *Sigui E un espai normat, $E \neq \{0\}$. Sigui F un subespai vectorial de E i sigui $x_0 \in E \setminus \overline{F}$. Aleshores existeix $g\colon E \to \mathbb{R}$ lineal i contínua tal que $g_{|F} = 0$ i $g(x_0) = 1$.*

Demostració: Prenem el subespai vectorial $G := F \oplus [x_0]$ de E i, si escrivim els $y \in G$ com $y = x + \lambda x_0$ amb $x \in F$ (que es pot fer de manera única per ser una suma directa), podem prendre $f\colon G \to \mathbb{R}$ de manera que $f(x + \lambda x_0) := \lambda$.

Es pot comprovar que f és lineal, ja que si $y_1, y_2 \in G$ i $\alpha_1, \alpha_2 \in \mathbb{R}$, aleshores $y_1 = x_1 + \lambda_1 x_0$ i $y_2 = x_2 + \lambda_2 x_0$, d'on $\alpha_1 y_1 + \alpha_2 y_2 = \alpha_1(x_1 + \lambda_1 x_0) + \alpha_2(x_2 + \lambda_2 x_0) = \alpha_1 x_1 + \alpha_2 x_2 + (\alpha_1\lambda_1 + \alpha_2\lambda_2)x_0$ i, per tant, $f(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1\lambda_1 + \alpha_2\lambda_2 = \alpha_1 f(y_1) + \alpha_2 f(y_2)$.

Anem a veure que f és contínua. En efecte, si considerem $y \in G \setminus F$ de manera que $y = x + \lambda x_0$ amb $\lambda \neq 0$, aleshores

$$||x + \lambda x_0|| = ||\lambda(x/\lambda + x_0)|| = |\lambda|||x/\lambda + x_0|| \geq |\lambda| d(x_0, F) > 0,$$

on hem utilitzat que $\lambda \neq 0$ i la definició de distància induïda per una norma d'un punt a un subespai vectorial. Ara fem el càlcul següent:

$$\Big|f\Big(\frac{x + \lambda x_0}{||x + \lambda x_0||}\Big)\Big| = \frac{1}{||x + \lambda x_0||}|f(x + \lambda x_0)| = \frac{|\lambda|}{||x + \lambda x_0||} \leq \frac{1}{d(x_0, F)} < \infty$$

per a tot $x \in F$ i per a tot $\lambda \neq 0 \in \mathbb{R}$, on hem utilitzat la linealitat de f i la desigualtat anterior. Aleshores, $\sup_{||u||=1, u \in F \oplus ([x_0] \setminus \{0\})} ||f(u)|| \leq \frac{1}{d(x_0, F)}$. Si $y \in F$ (i, per tant, $\lambda = 0$), aleshores $f(y) = 0$ i $||f(y)|| = 0$, d'on usant que $||f|| \geq 0$ en deduïm que $||f|| = \sup_{||u||=1, u \in F \oplus [x_0]} ||f(u)|| = \sup_{||u||=1, u \in F \oplus ([x_0] \setminus \{0\})} ||f(u)||$ i, per tant, aplicant la desigualtat anterior amb els $y \in F \oplus [x_0] \setminus \{0\}$, tenim que f és afitat. Per tant, com que f és lineal i afitat, aleshores f és continu, tal com volíem veure.

Aleshores usant el corol·lari 53, tenim que existeix $g\colon E \to \mathbb{R}$ lineal i contínua de manera que $g_{|G} = f$ i $||g|| = ||f||$. En particular, això passa per a tot $x \in F \subseteq F \oplus [x_0] = G$, d'on tenim que $g(x) = f(x) = f(x + 0x_0) = 0$, és a dir, $g_{|F} = 0$. Per al vector $x_0 \in [x_0] \subseteq F \oplus [x_0] = G$, tenim que $g(x_0) = f(x_0) = f(0 + 1x_0) = 1$, és a dir, $g(x_0) = 1$. Per tant, g compleix les condicions que volíem demostrar. □.

Ara anem a enunciar i demostrar el teorema per al cas complex ($K = \mathbb{C}$), fent alguna petita modificació en les hipòtesis (tenint en compte que la condició $f(x) \leq p(x)$ no té sentit si $f(x) \in \mathbb{C}$) i a demostrar-ne alguna de les seves conseqüències. Tenint això en compte, comencem donant les definició següents:

Definició. Sigui X un espai vectorial sobre $K = \mathbb{C}$. Aleshores direm que $p\colon X \to \mathbb{R}$ és un *funcional convex (complex)* si verifica:

(1) (a) $p(x + y) \leq p(x) + p(y)$ per a tot $x, y \in X$ (subadditivitat).
(2) (b) $p(\alpha x) = \alpha p(x)$ per a tot $\alpha \in \mathbb{R}$ tal que $\alpha \geq 0$ (positivament homogènia respecte $\mathbb{R} \subseteq \mathbb{C}$).
(3) (c) $p(x) \geq 0$ per a tot $x \in X$ (positiva).

Definició. Sigui $p(x)$ un funcional convex. Aleshores direm que $p(x)$ és *simètric* si per a tot escalar $\alpha \in \mathbb{C}$ es satisfà que $p(\alpha x) = |\alpha| p(x)$.

Definició. Sigui X un espai vectorial sobre $K = \mathbb{C}$ i $f\colon X \to \mathbb{C}$ una funció. Aleshores direm que $f(x)$ és un *funcional lineal* si

(1) $f(x + y) = f(x) + f(y)$ per a tot $x, y \in X$.
(2) $f(\alpha x) = \alpha f(x)$ per a tot $\alpha \in \mathbb{C}$, $x \in X$.

Definició. Sigui X un espai vectorial sobre $K = \mathbb{C}$. Direm que g és un *funcional lineal real* sobre l'espai complex X si $\alpha \in \mathbb{R}$ implica que $g(\alpha x) = \alpha g(x)$ per a tot $x \in X$.

Observem que afirmar que un funcional convex és simètric implica enfortir la condició (b) de la definició de funcional convex complex, que se satisfà automàticament per als valors $\alpha \in \mathbb{R} \subseteq \mathbb{C}$ amb $\alpha \geq 0$ per complir-se $|\alpha| = \alpha$.

Anem a enunciar i demostrar el teorema de Hahn-Banach per al cas complex:

PROPOSICIÓ 56 (Teorema de Hahn-Banach, cas complex). *Sigui X un espai vectorial complex, M un subespai vectorial de X, p un funcional convex simètric definit sobre X i f un funcional lineal definit sobre M amb la propietat que $|f(x)| \leq p(x)$ per a tot $x \in M$. Aleshores existeix un funcional lineal F definit en tot X que estén f i que satisfà la condició $|F(x)| \leq p(x)$ per a tot $x \in X$.*

Demostració: Per fer aquesta demostració, treballarem amb l'espai vectorial complex X com si es tractés d'un espai vectorial real escrivint la funció f com $f(x) = f_1(x) + if_2(x)$, essent $f_1(x)$ la part real de $f(x)$ i $f_2(x)$ la part imaginària de $f(x)$ per definició, de manera que $f_1(x)$ i $f_2(x)$ són funcionals reals si considerem la seva restricció a $\mathbb{R}$.

Anem a veure que a més a més $f_1(x)$ i $f_2(x)$ són funcionals lineals. En efecte, sigui $\alpha \in \mathbb{R}$ i considerem $\alpha f(x) = \alpha f_1(x) + i\alpha f_2(x)$. Donat que f és un funcional lineal, aquesta expressió ha de ser igual a $f(\alpha x) = f_1(\alpha x) + f_2(\alpha x)$. Igualant les parts reals i les parts imaginàries, tenim que $f_1(\alpha x) = \alpha f_1(x)$ i $f_2(\alpha x) = \alpha f_2(x)$. De manera anàloga, es pot comprovar que es conserven les sumes, així que f_1 i f_2 són funcionals lineals reals tal com volíem veure.

Aplicant la linealitat de f, podem establir les següents igualtats:

$$i(f_1(x) + if_2(x)) = if(x) = f(ix) = f_1(ix) + if_2(ix).$$

Per tant, igualant les parts reals i imaginàries, obtenim que $f_1(ix) = -f_2(x)$.

De les hipòtesis del teorema en podem deduir fàcilment la desigualtat $f_1(x) \leq p(x)$ per a tot $x \in M$. Per tant, aplicant el teorema de Hahn-Banach real al funcional lineal $f_1(x)$, sabem que existeix un funcional lineal F_1 amb valors reals, definit a tot l'espai, que estén f_1 i que satisfà $F_1(x) \leq p(x)$ per a tot $x \in X$.

Definim ara $F(x) := F_1(x) - iF_1(ix)$ i anem a veure que F estén f. Per tal de fer-ho, sigui $x \in M$ i, utilitzant que F_1 estén a f_1, tenim que $F_1(x) = f_1(x)$ i $F_1(ix) = f_1(ix) = -f_2(x)$, on a la segona igualtat hem utilitzat que $f_1(ix) = -f_2(x)$. Així doncs, $F(x) = f_1(x) - i(-f_2(x)) = f_1(x) + if_2(x) = f(x)$, és a dir, $F_{|M} = f$. Donat que F és un funcional lineal real per les conseqüències del teorema de Hahn-Banach utilitzat anteriorment, per tal de demostrar que també és un funcional complex només fa falta demostrar que $F(ix) = iF(x)$. Anem-ho a calcular:

$$F(ix) = F_1(ix) - iF_1(-x) = F_1(ix) + iF_1(x).$$

Per altra banda, tenim

$$iF(x) = iF_1(x) + F_1(ix),$$

d'on obtenim el que buscàvem igualant les dues expressions, i F és un funcional lineal complex.

Finalment, hem de veure que $|F(x)| \leq p(x)$. La desigualtat es satisfà si $F(x) = 0$ per ser $p(x) \geq 0 = F(x)$, així que podem suposar $F(x) \neq 0$. En aquest cas, podem escriure el nombre complex $F(x)$ en forma polar com $F(x) = re^{i\theta}$ amb $r \in \mathbb{R}$ i $\theta \in [0, 2\pi)$. Per tant, $F(e^{-i\theta}x) = e^{-i\theta}F(x) = r = |F(x)|$. Així doncs, $F(e^{-i\theta}x)$ és un nombre real, fet que implica que la seva part imaginària, $-F_1(ie^{-i\theta}x)$, s'anul·la, cosa que implica que $F(e^{-i\theta}x) = F_1(e^{-i\theta}x)$. Però com que $F_1(x) \leq |p(x)|$ per a tot $x \in X$, veiem que $|F(x)| = r = F_1(e^{-i\theta}x) \leq p(e^{-i\theta}x) = |e^{-i\theta}|p(x) = p(x)$.

Així doncs, el funcional F definit d'aquesta manera satisfà les hipòtesis del teorema. □

Anem a veure algunes conseqüències més del teorema.

En primer lloc, recordem la definició d'espai dual d'un espai vectorial E i el dotem de norma:

Definició. Sigui E un espai vectorial sobre un cos K. Aleshores l'*espai dual* (de E) es defineix com $E' := \mathcal{L}(E, K)$, és a dir, l'espai de tots els funcionals lineals continus sobre E.

Observació. Sigui E un espai vectorial i E' el seu dual. Si defineixo

$$||f||_{E'} = \sup_{||x||\leq 1, x\in E}|f(x)| = \sup_{||x||\leq 1, x\in E} f(x),$$

aleshores $(E', ||\cdot||_{E'})$ és un espai normat amb la norma $||\cdot||_{E'}$.

Sempre que no hi hagi confusió, escriurem $||f||$ en comptes de $||f||_{E'}$.

El teorema següent ens relaciona la separabilitat d'un espai vectorial X amb la separabilitat del seu espai dual associat:

PROPOSICIÓ 57. *Sigui X un espai vectorial sobre $K = \mathbb{R}$ o $K = \mathbb{C}$ i X' el seu espai dual associat. Llavors, si X és separable, aleshores X' també és separable.*

Demostració: Sigui $S := \{f \in X' \mid ||f|| = 1\}$. Com que qualsevol subconjunt d'un espai mètric separable és separable, aleshores S ha de ser separable. En particular, podem agafar una successió de funcions $(f_n)_{n\geq 1} \subseteq S$ de manera que sigui un subconjunt numerable dens de S. Donat que per a cada $f_n \in S$ es té que $||f_n|| = 1$ i que $||f_n|| = \sup_{||x||=1}|f_n(x)|$, es dedueix que per a tot $n \geq 1$ existeix un vector $x_n \in X$ tal que $||x_n|| = 1$ tal que $f_n(x_n) > \frac{1}{2}$, ja que si no existís tal x_n satisfent aquesta propietat per a cap $n \geq 1$ es contradiria el fet que $||f_n|| = 1$. Sigui ara $M = \overline{[(x_n)_{n\geq 1}]}$ la clausura del subespai generat per aquesta successió de vectors $(x_n)_{n\geq 1} \subseteq X$ i suposem que $M \neq X$. Per tant, existeix algun vector $x_0 \in X \setminus M$. Aleshores, la distància $d := d(x_0, M) > 0$ i es pot aplicar el corol·lari 55, de manera que existeix un funcional lineal afitat (continu) F tal que $||F|| = 1$, $F(x_0) \neq 0$ i $F_{|M} = 0$. Però com que $||F|| = 1$, aleshores $F \in S$ i $F(x_n) = 0$ per a tot $n \geq 1$ per ser $(x_n)_{n\geq 1} \subseteq M$.

Anem a escriure $f_n(x_n)$ de la manera següent:

$$f_n(x_n) = f_n(x_n) - F(x_n) + F(x_n).$$

Aleshores, per la desigualtat triangular, tenim que

$$|f_n(x_n)| \leq |f_n(x_n) - F(x_n)| + |F(x_n)|.$$

Donat que $F(x_n) = 0$ i

$$f_n(x_n) - F(x_n) = (f_n - F)(x_n),$$

podem escriure aquesta desigualtat de la següent manera:

$$|f_n(x_n)| \leq |(f_n - F)(x_n)|.$$

D'aquí se'n dedueix que

$$\frac{1}{2} < |f_n(x_n)| \leq |(f_n - F)(x_n)| \leq ||f_n - F||||x_n||$$

i, donat que $||x_n|| = 1$, arribem a la desigualtat

$$\frac{1}{2} < ||f_n - F||$$

per a tot $n \geq 1$. Si $(f_n)_{n\geq 1}$ fos un subconjunt dens de S i $F \in S$, aleshores per a tot $\epsilon > 0$ existeix algun $n_\epsilon := n(\epsilon)$ tal que $1/2 < ||f_{n_\epsilon} - F|| < \epsilon$ (en particular, per $\epsilon = \frac{1}{4}$ arribem a contradicció). Per tant, arribem a una contradicció que prové de suposar que $M \neq X$ i, per tant, el conjunt de totes les combinacions lineals (finites) de $(x_n)_{n\geq 1}$ és dens en X.

Si X és un espai vectorial sobre $K = \mathbb{R}$, ja n'hem deduït la separabilitat agafant les combinacions lineals finites de $(x_n)_{n\geq 1}$. Per tal de comprovar que aquest fet implica que X és separable en el cas que $K = \mathbb{C}$, fem el següent argument de numerabilitat:

(1) Sigui $\mathbb{Q}[i] := \{a + bi \mid a, b \in \mathbb{Q}\}$ el conjunt dels nombres racionals gaussians.
(2) Donat que el producte cartesià d'un nombre finit de conjunts numerables és numerable, aleshores el conjunt dels gaussians racionals és numerable, fet que ens permet enumerar-los com una successió $(\alpha_n)_{n\geq 1}$.

(3) Els nombres racionals són densos a $\mathbb{R}$, fet que implica que els nombres racionals gaussians són densos a $\mathbb{C}$.
(4) El conjunt $(\alpha_m x_n)_{n\geq 1}$, amb m fixat, és numerable, i com que la unió numerable de conjunts numerables és numerable, aleshores el conjunt $\cup_{m\geq 1}(\alpha_m x_n)_{n\geq 1}$ és numerable.
(5) Ara bé, donat que el conjunt de totes les combinacions lineals finites d'aquest conjunt és un conjunt numerable dens, i donat que $M = X$, se'n dedueix la separabilitat de X. □

A continuació, anem a demostrar alguns corol·laris més en el cas real per als espais duals:

COROL·LARI 58. *Sigui E un espai vectorial normat sobre $\mathbb{R}$ i $F \subseteq E$ un subespai vectorial. Si $f\colon F \to \mathbb{R}$ és un funcional lineal continu, aleshores existeix $G \in E'$ de manera que $G_{|F} = f$ i tal que*

$$||G||_{E'} = sup_{x\in F, ||x||\leq 1}|f(x)| = ||f||_{G'}.$$

Demostració: Aquest corol·lari és immediat utilitzant $p(x) = ||f||_{F'}||x||$ i aplicant el teorema de Hahn-Banach real, ja que es satisfà la hipòtesi $f(x) \leq ||f||_{F'}||x|| = p(x)$ per a tot $x \in F$ per definició. De fet, es tracta del corol·lari 53 reescrit amb la definició dels espais duals. □

COROL·LARI 59. *Sigui $E \neq \{0\}$ un espai vectorial normat sobre $\mathbb{R}$ i sigui $x_0 \in E$. Aleshores existeix $f_0 \in E'$ tal que $||f_0|| = ||x_0||$ i $\langle f_0, x_0\rangle = ||x_0||^2$.*

Demostració: Utilitzem el corol·lari 58 amb $F = [x_0]$ i l'aplicació $f\colon F \to \mathbb{R}$ definida com $f(\lambda x_0) = \lambda||x_0||^2$ per a tot $\lambda \in \mathbb{R}$ (a tot F per ser F el subespai generat per x_0), de manera que $||f||_{F'} = ||x_0||$. Si $\lambda = 1$, aleshores es compleix la condició de l'enunciat per a f, ja que $f(x_0) = ||x_0||$ i, per ser f_0 una extensió de f, es té $\langle f_0, x_0\rangle = \langle f, x_0\rangle = f(x_0) = ||x_0||$. La condició $||f_0||_{E'} = ||f||_{F'}$ també se'n dedueix del resultat del corol·lari 58. □

Observació. L'element f_0 donat al corol·lari 59 en general no és únic. Tanmateix, si E' és estrictament convex o si $E = L^p(\Omega)$ amb $1 < p < \infty$, aleshores f_0 és únic. En particular, podem definir una aplicació F que en comptes d'enviar vectors de E a vectors de E' envia vectors de E a subconjunts de E' de la següent manera:

$$F(x_0) := \left\{f_0 \in E' \colon\ ||f_0|| = ||x_0|| \text{ i } \langle f_0, x_0\rangle = ||x_0||^2\right\},$$

anomenada *aplicació dual*.

COROL·LARI 60. *Sigui E un espai vectorial real normat. Aleshores per a tot $x \in E$ es té que*

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

Demostració: Suposem que $x \neq 0$ (si $x = 0$, tots els termes de la igualtat són sempre 0 i la igualtat està clara). En aquest cas, està clar pel teorema de Cauchy-Schwarz que

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

D'altra banda, sabem pel corol·lari 59 que existeix algun $f_0 \in E'$ tal que $||f_0|| = ||x||$ i $\langle f_0, x\rangle = ||x||^2$. Si defineixo $f_1 := \frac{f_0}{||x||}$ (ho puc fer perquè $x \neq 0$), aleshores es té que $||f_1|| = \frac{||f_0||}{||x_0||} = 1$, de manera que $\langle f_1, x\rangle = \frac{\langle f_0, x\rangle}{||x||} = ||x||$. Per tant, s'assoleix el màxim en $f_1 \in E'$, $||f_1|| \leq 1$, de manera que se'n dedueix la igualtat de l'enunciat. □

A continuació, anem a enunciar i demostrar versions geomètriques del teorema de Hahn-Banach.

Per començar, introduïm algunes definicions preliminars sobre hiperplans i demostrem alguns teoremes i lemes que necessitarem per a demostrar els teoremes.

Definició. Sigui E un espai vectorial normat real. Un *hiperplà afí* és un subconjunt H de E de la forma

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

on f és un funcional lineal (no necessàriament continu) que no s'anul·la a tot arreu i $\alpha \in \mathbb{R}$ és una constant. En aquests casos, escrivim $H := [f = \alpha]$ i diem que $f = \alpha$ és l'equació de l'hiperplà H.

PROPOSICIÓ 61. *Sigui E un espai vectorial normal real. L'hiperplà $H = [f = \alpha] \subseteq E$ és tancat si i només si f és contínua.*

Demostració: Per tal de demostrar la implicació cap a la dreta, sabem de topologia que $\{\alpha\}$ és un tancat de $\mathbb{R}$ i l'antiimatge d'un tancat per una aplicació contínua és un tancat, és a dir, $H := f^{-1}(\{\alpha\})$ és un tancat.

Per tal de demostrar la implicació cap a l'esquerra, suposem que H és tancat. Aleshores, el complementari de E per H és $E \setminus H \neq \emptyset$ (ja que f no s'anul·la a tot arreu) un conjunt obert (per ser complementari d'un tancat). Sigui $x_0 \in E \setminus H$ de manera que $f(x_0) \neq \alpha$, de manera que podem suposar sense pèrdua de generalitat que $f(x_0) < \alpha$.

Sigui $r > 0$ un valor fixat tal que $B(x_0, r) \subseteq E \setminus H$, on

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

és la bola oberta de radi r centrada en x_0.

Podem comprovar que $f(x) < \alpha$ per a tot $x \in B(x_0, r)$: En efecte, suposem per tal d'arribar a contradicció que existeix algun $x_1 \in B(x_0, r)$ tal que $f(x_1) > \alpha$. Aleshores, com que les boles són un conjunt convex, sabem que el segment

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

que uneix x_0 i x_1 està contingut a $B(x_0, r)$ i, per tant, $f(x_t) \neq \alpha$ per a tot $t \in [0,1]$. Aleshores, aplicant el teorema del valor intermedi sobre aquest segment amb f contínua, deduïm que existeix algun $t_0 \in (0,1)$ tal que $f(x_{t_0}) = \alpha$. Per tant, $B(x_0, r) \cap H \neq \emptyset$, arribant a contradicció provinent de suposar que existeixi tal x_0. Per tant, deduïm que

$$f(x_0 + rz) < \alpha$$

per a tot $z \in B(0,1)$. Aleshores, en deduïm que f és contínua i que

$$||f|| \leq \frac{1}{r}(\alpha - f(x_0))$$

(al prendre suprems sobre $z \in B(0,1)$ a l'expressió anterior). □

Definicions. Sigui E un espai vectorial normat real, $A \subseteq E$ i $B \subseteq E$ dos subconjunts de E. Diem que l'hiperplà $H = [f = \alpha]$ *separa* A i B si $f(x) \leq \alpha$ per a tot $x \in A$ i $f(x) \geq \alpha$ per a tot $x \in B$.

Diem que H *separa esctrictament* A i B si existeix $\epsilon > 0$ tal que $f(x) \leq \alpha - \epsilon$ per a tot $x \in A$ i $f(x) \geq \alpha + \epsilon$ per a tot $x \in B$.

Geomètricament, la separació significa que A pertany a un dels dos semiespais determinat per H, i B pertany a l'altre.

Finalment, diem que un subconjunt $A \subseteq E$ és *convex* si $tx + (1-t)y \in A$ per a tot $x, y \in A$ i per a tot $t \in [0,1]$.

LEMA 62. *Sigui E un espai vectorial normat real. Sigui $C \subseteq E$ un obert convex amb $0 \in C$. Per a tot $x \in E$, defineixo $p(x) := inf\{\alpha > 0 \mid \alpha^{-1}x \in C\}$ (p s'anomena el gauge de C o el funcional de Minkowski de C). Aleshores p satisfà:*

(1) $p(\lambda x) = \lambda p(x)$ per a tot $x \in E$ i per a tot $\lambda > 0$.
(2) Existeix una constant M tal que $0 \leq p(x) \leq M||x||$ per a tot $x \in E$.
(3) $C = \{x \in E \mid p(x) < 1\}$.
(4) $p(x + y) \leq p(x) + p(y)$ per a tot $x, y \in E$.

Demostració: Anem a demostrar cadascuna de les propietats de p:

(1) $p(\lambda x) = \inf\{\alpha > 0 \mid \alpha^{-1}\lambda x \in C\} = \lambda\inf\{\alpha > 0 \mid \alpha^{-1}x \in C\} = \lambda p(x)$ per a tot $x \in E$ i per a tot $\lambda > 0$.
(2) Sigui $r > 0$ tal que $B(0,r) \subseteq C$, essent $B(0,r)$ la bola oberta de centre 0 i radi r. Per tant, tenim que $p(x) \leq \frac{1}{r}||x||$ per a tot $x \in E$, d'on tenim que la constant $M := \frac{1}{r}$ ens serveix. En particular, $p(x)$ està ben definit i es té $p(x) \geq r > 0$. Per tant, $0 \leq p(x) \leq M||x||$ per a tot $x \in E$.
(3) Suposem, en primer lloc, que $x \in C$. Com que C és obert, aleshores se segueix que $(1+\epsilon)x \in C$ per a $\epsilon > 0$ prou petit, de manera que $p(x) \leq \frac{1}{1+\epsilon} < 1$. Recíprocament, si $p(x) < 1$, aleshores existeix $\alpha \in (0,1)$ tal que $\alpha^{-1}x \in C$, i per tant $x = \alpha(\alpha^{-1}x) + (1-\alpha)0 \in C$ per la convexitat de C. Al tenir les dues inclusions, en deduïm la igualtat $C = \{x \in E \mid p(x) < 1\}$.
(4) Siguin $x, y \in E$ i sigui $\epsilon > 0$. Aleshores usant (1) i (3) tenim que $\frac{x}{p(x)+\epsilon} \in C$ i $\frac{y}{p(y)+\epsilon} \in C$. Per tant, $\frac{tx}{p(x)+\epsilon} + \frac{(1-t)y}{p(y)+\epsilon} \in C$ per a tot $t \in [0,1]$ (per la convexitat de C). Escollint el valor $t = \frac{p(x)+\epsilon}{p(x)+p(y)+2\epsilon}$, trobem que $\frac{x+y}{p(x)+p(y)+2\epsilon} \in C$. Utilitzant (1) i (3) una altra vegada, deduïm que $p(x+y) < p(x)+p(y)+2\epsilon$ per a tot $\epsilon > 0$ (ja que $f\big(\frac{x+y}{p(x)+p(y)+2\epsilon}\big) < 1$), d'on deduïm $p(x+y) \leq p(x)+p(y)$ per a tot $x, y \in E$, tal com volíem veure.

Anem ara a demostrar el lema previ a les versions geomètriques del teorema que utilitza la versió del teorema en forma analítica (i, per tant, és en aquest punt de la demostració on utilitzem el lema de Zorn):

LEMA 63. *Sigui E un espai vectorial normat real. Sigui $C \subseteq E$ un obert convex no buit i $x_0 \in E$ amb $x_0 \notin C$. Aleshores existeix $f \in E'$ tal que $f(x) \leq f(x_0)$ per a tot $x \in C$. En particular, l'hiperplà $\{f = f(x_0)\}$ separa $\{x_0\}$ i C.*

Demostració: Si considero $D := C + a$ per a certa $a \in X$, podem considerar que el conjunt C traslladat per a (és a dir, D) conté el 0, i D també satisfà les hipòtesis de ser $D \subseteq E$ un obert convex no buit tal que $y_0 := a + x_0 \notin D$. En aquest cas que $0 \in D$, considerem el subespai lineal $[y_0]$ i el funcional lineal $g\colon [y_0] \to \mathbb{R}$ definit per $g(ty_0) = t$ per a tot $t \in \mathbb{R}$. Aleshores, està clar que $g(x) \leq p(x)$ per a tot $x \in [y_0]$, estudiant per separat els casos $t > 0$ i $t \leq 0$, que ambdós són anàlegs. Aleshores aplicant la versió analítica del teorema de Hahn-Banach (38) i els apartats (1) i (2) del lema 62 per comprovar que $p(x)$ és un funcional sublineal tenim que existeix un funcional lineal $f\colon E \to \mathbb{R}$ tal que $f_{|[y_0]} = g$ i satisfà $f(x) \leq p(x)$ per a tot $x \in E$. En particular, tenim que $f(y_0) = 1$ i que f és contínua per l'apartat (2) del lema 63. Per tant, per l'apartat (3) del lema 63, també deduïm que $f(x) < 1$ per a tot $x \in D$. Si traslladem aquest funcional al conjunt C, tornem a obtenir les condicions del teorema amb x_0 en comptes de y_0, i deduïm que l'hiperplà $H := [f = f(x_0)]$ separa el conjunt $\{x_0\}$ i C, satisfent la condició $f(x) < 1 = f(x_0)$ per a tot $x \in C$ i essent $f \in E'$. □

PROPOSICIÓ 64 (Teorema de Hahn-Banach, primera versió geomètrica). *Sigui E un espai vectorial normat real, $A \subseteq E$ i $B \subseteq E$ dos subconjunts no buits convexos tals que $A \cap B = \emptyset$. Suposem que un dels dos és obert. Aleshores existeix un hiperplà tancat que separa A i B.*

Demostració: En primer lloc suposem sense pèrdua de generalitat que el conjunt A és el conjunt que és obert. Aleshores, els conjunts traslladats $x + A$ són oberts per a tot $x \in E$.

Sigui $C := A - B = \{a - b \mid a \in A, b \in B\}$. Aleshores C és convex, ja que si $c_1 = a_1 - b_1 \in C$ i $c_2 = a_2 - b_2 \in C$ amb $a_1, a_2 \in A$ i $b_1, b_2 \in B$, aleshores per a tot $t \in [0,1]$ es té que

$$(1-t)c_1 + tc_2 = (1-t)(a_1 - b_1) + t(a_2 - b_2) = ((1-t)a_1 + ta_2) - ((1-t)b_1 + tb_2) =$$
$$= a_3 - b_3 \in A - B = C$$

amb $a_3 := (1-t)a_1 + ta_2 \in A$ i $b_3 := (1-t)b_1 + tb_2 \in B$ per ser A i B convexos. També tenim que C és obert, ja que $C = \bigcup_{y \in B}(A - y)$ essent els conjunts $A - y$ oberts per a tot $y \in B$. A més a més, també sabem que $0 \notin C$, ja que $A \cap B = \emptyset$. Per tant, utilitzant el lema 63 aplicant les hipòtesis amb $x_0 = 0$, tenim que existeix algun $f \in E'$ tal que $f(z) < 0 = f(0)$ per a tot $z \in C$ (aplicant la linealitat de f per veure que $f(0) = 0$), és a dir, $f(a) < f(b)$ per a tot $a \in A$ i per a tot $a \in B$ (al fer $c = a - b$ i $f(c) = f(a) - f(b)$ aplicant la linealitat de f).

Només cal agafar una constant $\alpha \in \mathbb{R}$ que satisfaci

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

per obtenir un hiperplà $H := [f = \alpha]$ (tancat, per ser f contínua i aplicant la proposició 61) que separi A i B, tal com volíem demostrar. □

PROPOSICIÓ 65 (Teorema de Hahn-Banach, segona versió geomètrica). *Sigui E un espai vectorial normat real. Siguin $A \subseteq E$ i $B \subseteq E$ dos conjunts convexos no buits tals que $A \cap B = \emptyset$. Suposem que A és tancat i B és compacte. Aleshores existeix un hiperplà tancat que separa estrictament A i B.*

Demostració: Defineixo els conjunts $A_\epsilon := A + B_(0, \epsilon)$ i $B_\epsilon := B + B(0, \epsilon)$, essent $B(0, \epsilon)$ una bola oberta de radi ϵ. Tenint en compte que A i B són disjunts, volem veure que podem trobar un valor de ϵ prou petit de manera que A_ϵ i B_ϵ siguin disjunts. Anem-ho a veure per reducció a l'absurd. Suposem que per a tot $\epsilon > 0$, existeix $z(\epsilon) \in A_\epsilon \cap B_\epsilon$. Aleshores podem escollir una successió $(\epsilon_n)_{n\geq 1}$ i $z(\epsilon_n) := z_n$ ens defineix una successió $(z_n)_{n\geq 1}$ de manera que $z_n \in A_{\epsilon_n} \cap B_{\epsilon_n}$ per a tot $n \geq 1$. Per definició de A_ϵ i B_ϵ, existeixen $x_n \in A$ i $y_n \in B$ de manera que $||x_n - z_n|| < \epsilon_n$ i $||y_n - z_n|| < \epsilon_n$ i, per tant, per la desigualtat triangular tenim que $||x_n - y_n|| < 2\epsilon_n$. Com que B és compacte, podem agafar una subsuccessió $(y_{n_k}) \subseteq (y_n)_{n\geq 1}$ de manera que (y_{n_k}) convergeixi a $y \in B$. Aleshores aplicant novament la desigualtat triangular obtenim que $(x_n)_{n\geq 1}$ també convergeix a $y \in B$. Com que A és tancat, aleshores necessàriament $y \in A$ per ser de la seva clausura, d'on tenim que $y \in A \cap B$ arribant a contradicció amb la nostra hipòtesi que A i B són disjunts. Per tant, existeix algun ϵ satisfent $A_\epsilon \cap B_\epsilon = \emptyset$.

Aleshores, com que clarament A_ϵ i B_ϵ són oberts convexos no buits i disjunts, podem aplicar la versió anterior 64 del teorema de Hanh-Banach per a aquests conjunts i tenim que existeix $f \in E'$ no nul·la i $\alpha \in \mathbb{R}$ tal que $f(x) \leq \alpha \leq f(y)$ per a tot $x \in A_\epsilon$ i per a tot $y \in B_\epsilon$. Per tant, per a tot $x \in A$, per a tot $y \in B$ i per a tot $z \in B(0, 1)$, tenim que $f(x + \epsilon z) \leq \alpha \leq f(y + \epsilon z)$. Per tant, podem escollir $z_0 \in B(0, 1)$ de manera adequada per tal que es satisfacin les desigualtats $f(x) \leq \alpha - \epsilon||f||_{E'}$ i $f(y) \geq \alpha - \epsilon||f||_{E'}$ per a tot $x \in A$ i per a tot $y \in B$, obtenint doncs que l'hiperplà $H := [f = \alpha]$ (tancat, per ser f contínua i aplicant la proposició 61) separa A i B estrictament, tal com volíem demostrar. □

De manera anàloga al teorema 64, puc definir el conjunt $C := A - B$ i veure que és convex amb $0 \notin C$, utlitzant les mateixes hipòtesis que ens dóna l'enunciat d'aquest teorema.

Volem comprovar que C és tancat. En efecte, sigui

Anem a demostrar el següent corol·lari del teorema de Hahn-Banach usant les seves versions geomètriques. Observem que en realitat és un corol·lari que es desprèn immediatament del corol·lari 55, però ara anem a demostrar-ho d'aquesta altra manera usant la segona versió geomètrica:

COROL·LARI 66. *Sigui E un espai vectorial normat real. Sigui $F \subseteq E$ un subespai lineal tal que $\overline{F} \neq E$. Aleshores existeix alguna $f \in E'$, f no idènticament nul·la, tal que $\langle f, x\rangle = 0$ per a tot $x \in F$.*

Demostració: Sigui $x_0 \in E \setminus \overline{F}$. Aleshores aplicant el teorema 65 amb $A = \overline{F}$ i $B = \{x_0\}$, trobem un hiperplà tancat $H := [f = \alpha]$ que separa estrictament $\overline{F}$ i $\{x_0\}$. Per tant, tenim que

$$\langle f, x\rangle < \alpha < \langle f, x_0\rangle$$

per a tot $x \in F$. Per tant, se'n dedueix que $\langle f, x\rangle = 0$ per a tot $x \in F$ donat que $\lambda\langle f, x\rangle < \alpha$ per a tot $\lambda \in \mathbb{R}$ (recordant que $[x_0] \subseteq E \setminus \overline{F}$ i que la desigualtat s'ha de satisfer per a tot $x_1 = \lambda x_0 \in [x_0]$ i podem treure el factor escalar λ per la linealitat de f). □

2. Demostració alternativa de la paradoxa de Banach-Tarski (més moderna)

En aquest apartat de l'Apèndix donarem un esquema de demostració alternativa de la paradoxa de Banach-Tarski utilitzant conceptes de teoria de grups més moderns que la demostració original. Enunciarem i deixarem sense demostrar la majoria dels resultats.

Definició. Sigui G un grup i X un G-conjunt. Dos subconjunts A i B de X s'anomenen *G-congruents* si existeix un element g de G tal que $gA = B$.

Observació. La relació de G-congruència és una relació d'equivalència a $\mathcal{P}(X)$. En aquest cas, denotarem la relació de ser G-congruent per $A \sim_G B$ per a $A, B \in \mathcal{P}(X)$.

Definició. Un *mapa G-congruent* entre $A, B \in \mathcal{P}(\mathcal{X})$ és una bijecció definida per l'acció de qualsevol element $g \in G$.

Definició. Direm que una partició finita d'un conjunt X és una *descomposició* de X.

Definició. Sigui G un grup i X un G-conjunt. Dos subconjunts $A, B \subseteq X$ s'anomenen *G-equidescomponibles* si existeixen descomposicions $\{A_i \mid 1 \leq i \leq n\}$ i $\{B_i \mid \mid \leq i \leq n\}$ de A i B de manera que per a tot $i \in \{1, \ldots, n\}$, A_i i B_i són G-congruents.

Observació. La relació de G-equidescomponibilitat és una relació d'equivalència a $\mathcal{P}(X)$. En aquest cas, denotarem la relació de ser G-congruent per $A \approx_G B$ per a $A, B \in \mathcal{P}(X)$.

Definició. Defineixo un *mapa G-trencaclosques* entre A i B com una aplicació bijectiva $f\colon A \to B$ amb la propietat que existeixen descomposicions $\{A_i \mid 1 \leq i \leq n\}$ de A i $\{B_i \mid 1 \leq i \leq n\}$ de B tals que la restricció de f a A_i és un mapa G-congruent entre A_i i B_i.

Observació. Descomposicions de A i B satisfent la definició de G-equidescomponibilitat defineixen i estan definits per un mapa G-trencaclosques entre A i B.

Definició. Direm que A és *G-subdescomponible* respecte B si existeix un conjunt $B' \subseteq B$ de manera que A i B' són G-equidescomponibles. En tal cas, denotarem $A \precsim_G B$.

PROPOSICIÓ 67. *Sigui S un conjunt, R una relació a S i E una relació d'equivalència a S. Suposem que les següents afirmacions es verifiquen:*

(1) Per a tot parell de conjunts $a, b \in S$ tals que $(a, b) \in R$, existeix una aplicació injectiva $f\colon a \to b$ tal que per a qualsevol subconjunt c de a, es té que $(c, f(c)) \in E$.
(2) Per a qualsevol quàdrople de conjunts $a_1, a_2, a_3, a_4 \in S$, si $a_1 \cap a_2 = \emptyset$, $b_1 \cap b_2 = \emptyset$, $(a_1, b_1) \in E$ i $(a_2, b_2) \in E$, aleshores $(a_1 \cup a_2, b_1 \cup b_2) \in E$.

Aleshores per a tot $a, b \in S$, $(a, b) \in R$ i $(b, a) \in R$ implica que $(a, b) \in E$.

PROPOSICIÓ 68. *(Teorema de Banach-Schröder-Bernstein) Si G és un grup i X és un G-conjunt, aleshores la G-subdescomponibilitat és una relació d'ordre de $\mathcal{P}(X)$ respecte la G-equidescomponibilitat.*

Demostració: La reflexivitat i la transitivitat de G són fàcils de demostrar. Pel que fa a la propietat antisimètrica, hem de veure que es satisfan les propietats (1) i (2) de la proposició 67 amb la G-descomponibilitat i la G-equidescomponibilitat i quedarà automàticament comprovada. En efecte, si A i B són subconjunts de X tals que $A \sim_G B'$ per a algun subconjunt B' de B, aleshores un mapa G-trencaclosques entre A i B' satisfà la condició (1). La condició (2) és òbvia per la definició de G-equidescomponibilitat. □

Definició. Sigui G un grup i X un G-conjunt. Direm que un subconjunt C de X és *G-paradoxal* si existeixen dos subconjunts disjunts $A, B \subseteq C$ tals que A, B i C són G-descomponibles dos a dos. Direm que el grup G és *paradoxal* si és G-paradoxal quan el veiem com un conjunt en el qual G actua per translació (per l'esquerra).

COROL·LARI 69. *Un subconjunt C de X és tal que existeix un subconjunt $A \subseteq C$ tal que $A \sim_G C$ i $C \setminus A \sim_G C$ (es pot duplicar) si i només si és G-paradoxal.*

Demostració: Aplicant el teorema 68 és immediat. □

Observació. Observem que si un G-conjunt és H-paradoxal per a algun subgrup H de G, aleshores també és G-paradoxal.

En aquests termes, podem reescriure la paradoxa de Banach-Tarski com *la bola unitat de $\mathbb{R}^3$ és G-paradoxal, on $G = \mathrm{Isom}(\mathbb{R}^3)$ és el grup d'isometries de $\mathbb{R}^3$*. De fet, només necessitem utilitzar isometries directes, però aquest fet no juga cap paper important en el que es discuteix a continuació.

Anem a estudiar breument els grups lliures i les accions lliures. Denotem per F_n el grup lliure format per n generadors.

PROPOSICIÓ 70. *El grup F_2 és paradoxal.*

Demostració: Sabem que tot element de F_n el podem expressar de manera única com un producte reduït (una *paraula*) dels seus generadors i els seus inversos (fent correspondre la paraula buida a l'element unitat). Siguin a i b generadors de F_2. Denotem per A_+ (respectivament, A_-, B_+ i B_-) el subconjunt de F_2 consistent en totes les paraules que comencen amb a per l'esquerra (respectivament, a^{-1}, b i b^{-1}). Aleshores $A_- \cup A_+$ i $B_- \cup B_+$ són subconjunts disjunts de F_2 i tenen descomposicions $\{A_-, A_+\}$ i $\{B_-, B_+\}$. Per tant, $A_- \cup A_+ \approx_{F_2} aA_- \cup A_+ = F_2$ i $B_- \cup B_+ \approx_{F_2} bB_- \cup B_+ = F_2$, tal com volíem demostrar. □

En aquesta proposició fem servir l'axioma de l'elecció.

PROPOSICIÓ 71. *Sigui G un grup que actua lliurement sobre un conjunt X. Si G és un grup paradoxal, aleshores X és G-paradoxal.*

Demostració: Utilitzant l'Axioma de l'elecció, existeix un conjunt T que conté exactament un element de cada òrbita de G en X. Si H_1, H_2 són subconjunts disjunts de G amb $G = H_1 \cup H_2$ tals que G, H_1 i H_2 són G-equidescomponibles dos a dos, aleshores clarament tenim que $G(T), H_1(T)$ i $H_2(T)$ són G-equidescomponibles dos a dos. Com que $H_1(T)$ i $H_2(T)$ són disjunts (atès que l'acció de G és lliure) i $G(T) = X$, aleshores X és G-paradoxal. □

PROPOSICIÓ 72. *Sigui G un grup i X un G-conjunt. Aleshores G actua lliurement sobre $X \setminus D$ on D és el conjunt de tots els punts fixos per un element no trivial de G.*

Demostració: En primer lloc, anem a veure que $X \setminus D$ és estable sota l'acció de G. Si $y \in D$ i $g \in G$, i si $h \in G$ és tal que $hy = y$, aleshores hg^{-1} deixa fix gx i, per tant, $gx \in D$. Per complementarietat, tenim que $G(X \setminus D) = X \setminus D$. Ara hem de demostrar que per a qualsevol parell d'elements diferents $g, h \in G$ i per a qualsevol punt $x \in X \setminus D$ tenim que $gx \neq hx$. Suposem que $gy = hy$ per a algun $y \in X$. Aleshores $y = g^{-1}h(y)$, i com que hem suposat que $g^{-1}h$ és no trivial, aleshores $y \in D$. □

Ara hem de veure que l'esfera unitat $\mathbb{S}^2$ de $\mathbb{R}^3$ és SO(3)-paradoxal, essent SO(3) el grup de rotacions de l'espai euclidià $\mathbb{R}^3$.

El primer pas consisteix en demostrar el lema següent:

LEMA 73. *Existeixen rotacions A i B a través d'eixos de $\mathbb{R}^3$ que passen per l'origen generant un subgrup de $\mathcal{SO}(3)$ isomorf a $\mathbb{F}_2$, el grup lliure de dos generadors.*

Demostració: Siguin

$$A^{\pm} = \begin{bmatrix} \frac{1}{3} & \mp\frac{2\sqrt{2}}{3} & 0 \\ \pm\frac{2\sqrt{2}}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix}, B^{\pm} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & \mp\frac{2\sqrt{2}}{3} \\ 0 & \pm\frac{2\sqrt{2}}{3} & \frac{1}{3} \end{bmatrix}$$

les nostres rotacions. Observem que $A^{\pm}$ corresponen a matrius de rotació d'angles $\arccos(\frac{1}{3})$ al voltant de l'eix z i $B^{\pm}$ a matrius de rotació d'angles $\arccos(\frac{1}{3})$ al voltant de l'eix x. Ara, sigui ω una paraula reduïda en $A^{\pm}$ i $B^{\pm}$ que no sigui la paraula buida I (matriu identitat). Com que ω no actua com a element identitat sobre $\mathbb{R}^3$, necessàriament tenim que $\langle A, B\rangle \simeq \mathbb{F}_2$, on $\langle A, B\rangle$ és el subgrup de $\mathcal{SO}(3)$ generat per A i B. Fem notar que, sense pèrdua de generalitat, podem assumir que ω acaba en $A^{\pm}$, atès que si ω fos la paraula buida, aleshores la conjugació per $A^{\pm}$ no l'alteraria. Afirmem que

$$w \cdot \begin{bmatrix}1\\0\\0\end{bmatrix} = \frac{1}{3^k}\begin{bmatrix}a\\b\sqrt{2}\\c\end{bmatrix}$$

amb $a, b, c \in \mathbb{Z}$, $3 \nmid b$ i k la llargada de la paraula ω, de manera que en tal cas ω no pot actuar com a identitat a $\mathbb{R}^3$. Anem-ho a veure per inducció sobre k. Per hipòtesi, doncs, el cas base és $w = A^{\pm}$, de manera que

$$w \cdot \begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}\frac{1}{3} & \mp\frac{2\sqrt{2}}{3} & 0\\ \pm\frac{2\sqrt{2}}{3} & \frac{1}{3} & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix} = \frac{1}{3^k}\begin{bmatrix}\frac{1}{3}\\ \pm\frac{2\sqrt{2}}{3}\\0\end{bmatrix} = \frac{1}{3}\begin{bmatrix}1\\ \pm 2\sqrt{2}\\0\end{bmatrix}.$$

Per tant, efectivament, el cas $k = 1$ funciona. Per veure el cas general, sigui $w = A^{\pm}w'$ o $w = B^{\pm}w'$ de manera que

$$w' \cdot \begin{bmatrix}1\\0\\0\end{bmatrix} = \frac{1}{3^{k-1}}\begin{bmatrix}a'\\b'\sqrt{2}\\c'\end{bmatrix}$$

essent $a', b', c' \in \mathbb{Z}$ i $3 \nmid b'$. Aleshores un càlcul senzill ens mostra que, si estem en el primer cas,

$$w \cdot \begin{bmatrix}1\\0\\0\end{bmatrix} = A^{\pm}\frac{1}{3^{k-1}}\begin{bmatrix}a'\\b'\sqrt{2}\\c'\end{bmatrix} = \frac{1}{3^{k-1}}\begin{bmatrix}\frac{1}{3} & \mp\frac{2\sqrt{2}}{3} & 0\\ \pm\frac{2\sqrt{2}}{3} & \frac{1}{3} & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}a'\\b'\sqrt{2}\\c'\end{bmatrix} = \frac{1}{3^{k-1}}\begin{bmatrix}\frac{1}{3}a' \mp \frac{1}{3}4b'\\ \pm a'\frac{2\sqrt{2}}{3} + \frac{1}{3}b'\sqrt{2}\\ c'\end{bmatrix} =$$

$$= \frac{1}{3^k}\begin{bmatrix}a' \mp 4b'\\ \sqrt{2}(b' \pm 2a')\\ 3c'\end{bmatrix} = \frac{1}{3^k}\begin{bmatrix}a\\b\sqrt{2}\\c\end{bmatrix},$$

i de manera similar podem fer un càlcul d'aquest estil amb $w = B^{\pm}w'$. En aquests dos casos, tenim:

$$\begin{cases} a = a' \mp 4b', b = b' \pm 2a', c = 3c' & \text{si } \omega = A^{\pm}w' \\ a = 3a', b = b' \mp 2c', c = c' \pm 4b' & \text{si } w = B^{\pm}w' \end{cases}$$

Evidentment, $a, b, c \in \mathbb{Z}$. Hem de demostrar que $3 \nmid b$ i, d'aquesta manera, haurem obtingut la fórmula general. En efecte:

(1) Cas 1: $\omega = A^{\pm}B^{\pm}v$ (on possiblement $v = I$). Aleshores tenim

$$w \cdot \begin{bmatrix}1\\0\\0\end{bmatrix} = A^{\pm}B^{\pm}v\begin{bmatrix}1\\0\\0\end{bmatrix}$$

de manera que de les equacions anteriors tenim que $b = b' \pm 2a' = b' \pm 6a''$. Com que, per hipòtesi, $3 \nmid b'$, se'n dedueix que $3 \nmid b$.

(2) Cas 2: $\omega = B^{\pm}A^{\pm}v$. Es demostra igual que el cas 1.

(3) Cas 3: $\omega = A^{\pm}A^{\pm}v$. Per hipòtesi, tenim

$$v \cdot \begin{bmatrix}1\\0\\0\end{bmatrix} = \frac{1}{3^{k-2}}\begin{bmatrix}a''\\b''\sqrt{2}\\c''\end{bmatrix}$$

amb $a'', b'', c'' \in \mathbb{Z}$ i $3 \nmid b''$. De les equacions anteriors, deduïm que $b = b' \pm 2a' = b' \pm 2(a'' \mp 4b'') = b' + b'' \pm 2a'' - 9b'' = 2b' - 9b''$ de manera que $3 \nmid b$ perquè $3 \nmid b'$ per hipòtesi d'inducció.

(4) Cas 4: $\omega = B^{\pm}B^{\pm}v$. Aquest cas es tracta de manera anàloga que el Cas 3.

Per tant, com que per a qualsevol llargada k de ω obtenim que l'aplicació de la paraula ω a l'eix $(1,0,0)$ ens dóna elements diferents del vector $(1,0,0)$ de $\mathcal{SO}(3)$ per a qualsevol $\omega \neq I$ arbitrari però resultat d'una simple comprovació de casuística-, atès que hem comprovat que hem tret els factors múltiples de tres que apareixen a les components vectorials, i comprovant els casos que ens falten per a paraules de la mateixa longitud per comprovar que obtenim elements diferents com a l'article de referència Tom Weston, *The Banach-Tarski Paradox* -càlculs llargs i tediosos-, podem assegurar que $\langle A, B\rangle \simeq \mathbb{F}_2$ tal com volíem veure. □

Aleshores, usant aquest lema, podem demostrar el teorema següent:

PROPOSICIÓ 74. *Si S és una esfera de $\mathbb{R}^3$ i D és un subconjunt enumerable de S, aleshores S i $S \setminus D$ són $\mathbb{SO}(3)$-equidescomponibles.*

Demostració: La idea de la demostració és demostrar, utilitzant l'enumerabilitat de D, que existeix una rotació de l'esfera que es pot aplicar al conjunt D tantes vegades com es vulgui sense que cap punt de D retorni al conjunt D. Ara, si apliquem aquesta rotació una vegada més al conjunt de tots els punts que es poden obtenir d'aquesta manera juntament amb D, tornarem a caure en el mateix conjunt, però D haurà desaparegut.

En efecte, sigui O el centre de l'esfera S. Com que D és enumerable, existeix $d \in S$ tal que ni d ni el seu oposat pertanyen a D. Denotem per L la recta que conté O i d. Siguin x i y dos punts de D i considerem rotacions ρ d'eix L de manera que $\rho^n(x) = y$ per a algun enter positiu n. Òbviament, si x i y no estan continguts al pla normal de L, aleshores no existeix tal rotació. En canvi, com que x no pertany a l'eix L, existeixen exactament n rotacions que ho satisfan (dit d'una altra manera, si la rotació d'angle α envia x a y, aleshores les n rotacions són d'angles $(\alpha + 2k\pi/n$ per a $k \in \{0, \dots, n-1\})$. Aleshores el conjunt R de totes aquestes rotacions per a qualsevol parell de punts $x, y \in D$ i per a tot enter positiu n és enumerable. Com que el conjunt de totes les rotacions d'eix L no és numerable, existeix una rotació r d'eix L que no pertany a R, és a dir, de manera que per a tot $x, y \in D$ i per a tot enter positiu n es tingui que $r^n(x) \neq y$. Això significa que els conjunts $r^n(D)$ per a tot enter no negatiu n són disjunts dos a dos. Sigui U la unió disjunta dels conjunts següents: $u = \cup_{n\geq 0} r^n(D)$. Aleshores $r(U) = \cup_{n\geq 1} r^n(D) = U \setminus D$, i se'n segueix per definició que $U \sim_{\mathrm{SO}(3)} U \setminus D$. Aquest fet implica que $S = U \cup (S \setminus U) \sim_{\mathrm{SO}(3)} (U \setminus D) \cup (S \setminus U) = S \setminus D$ atès que les unions són disjuntes, i així concloem la demostració. □

A partir d'ara, assumim que $G = \mathrm{Isom}(\mathbb{R}^3)$ per a totes les definicions de G-congruència, G-equidescomponibilitat i G-paradoxalitat de subconjunts de $\mathbb{R}^3$ que apareguin.

Observació. Siguin A i B dos subconjunts disjunts de l'esfera unitat de manera que A, B i $\mathbb{S}^2$ siguin SO(3)-equidescomponibles dos a dos. Aleshores obtindrem la $\mathbb{SO}(3)$-paradoxalitat de l'esfera unitat tancada (sense el seu centre) amb $\mathbb{B}^3 \setminus \{0\}$ si considerem les unions A' i B' de tots els segments que surten del 0 (no inclòs) als punts de A i B respectivament. En altres paraules, els conjunts $A' = \cup_{a\in A}(0, a]$ i $B' = \cup_{b\in B}(0, b]$ i $\mathbb{B}^3 \setminus \{0\}$ són SO(3)-equidescomponibles dos a dos.

PROPOSICIÓ 75. *Sigui B una bola tancada de centre O. Aleshores B i $B \setminus O$ són equidescomponibles.*

Demostració: Suposem sense pèrdua de generalitat que B és la bola unitat tancada. Aleshores considero el conjunt

$$D = \{(\cos(n), \sin(n), 0) \mid n \in \mathbb{N}\}.$$

Sigui ρ una rotació d'un radian al voltant de l'eix vertical. Com que cap múltiple de 2π és enter, és obvi que $\rho(D) = D \setminus \{(1,0,0)\}$, de manera que $B = (B \setminus D) \cup D \sim_G (B \setminus D) \cup (D \setminus \{(1,0,0)\}) = B \setminus \{(1,0,0)\}$. Ara, utilitzant la descomposició $\{B \setminus \{(1,0,0), 0\}, \{(1,0,0)\}\}$, deduïm que $B \setminus \{(1,0,0)\} \sim_G B \setminus \{0\}$, i per transitivitat $B \setminus \{0\}$, que és el resultat desitjat. □

Per tant, la bola unitat $\mathbb{B}^3$ és paradoxal. En particular, existeix una descomposició $\{A, B\}$ de $\mathbb{B}^3$ tal que A, B i $\mathbb{B}^3$ són equidescomponibles dos a dos. Utilitzant una translació, el conjunt A també és equidescomponible a una altra bola tancada unitat de radi 1 disjunta amb $\mathbb{B}^3$, d'on obtenim una *duplicació* de la bola. Aquest resultat el podem generalitzar de la manera següent:

PROPOSICIÓ 76. *Una unió finita de boles tancades de radi* 1 *i la bola unitat tancada són equidescomponibles.*

Demostració: Utilitzant l'observació anterior a la proposició 75 i la proposició 75. □

També podem generalitzar aquest resultat, usant la proposició 67 i la proposició 76, als teoremes següents:

LEMA 77. *Qualsevol bola unitat de* $\mathbb{R}^3$ *és equidescomponible a la bola unitat tancada.*

PROPOSICIÓ 78. *Sigui A un conjunt fitat de* $\mathbb{R}^3$ *amb interior no buit. Aleshores A i l'esfera unitat* $\mathbb{B}^3$ *són equidescomponibles.*

Per transitivitat, en deduïm la paradoxa de Banach-Tarski:

COROL·LARI 79. *Qualsevol parell de conjunts fitats de* $\mathbb{R}^3$ *amb interiors no buits són equidescomponibles.*

Bibliografia

[1] Kevin Barnum, *The axiom of choice and its implications*, http://math.uchicago.edu/ may/REU2014/REUPapers/Barnum.pdf

[2] Andreas Blass, *Existence of basis implies the Axiom of Choice*, http://www.math.lsa.umich.edu/ ablass/bases-AC.pdf.

[3] Donald Brower, *Notes on the Banach-Tarski Paradox* http://www3.nd.edu/ dbrower/papers/banach-tarski.pdf, 2006.

[4] Josep Maria Brunat, *(Ma)temàtiques Clàssiques*.

[5] Keith Devlin, *The joy of sets. Fundamentals of Contemporary Set Theory.* Springer-Verlag

[6] Cristian Gerardo Allen. *The Axiom of Choice.*
digarchive.library.vcu.edu/bitstream/10156/2937/1/thesis_AC.pdf.
A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University. Richmond, Virginia, Maig 2010.

[7] Hanspeter Fisher, *Algebraic clousure*, http://www.cs.bsu.edu/ fischer/math412/Closure.pdf

[8] Edwin Hewitt and Karl Stromberg. *Real and Abstract Analysis*. Springer-Verlag, New York, 1965.

[9] James M. Henle. *An Outline of Set Theory*. Dover Publications Inc. Mineola, New York, 1986.

[10] Marc Hoyons, *A short proof of the Banach–Tarski paradox*, 2006.

[11] Paul Howard, Jean E. Rubin. *Consequences of the Axiom of Choice*, American Mathematical Society. Mathematical Surveys and Monographs, Volume 59.

[12] Thomas J. Jech. *Set Theory. The Third Millennium Edition, Revised and Expanded* 3r edition, Springer

[13] Thomas J. Jech. *The Axiom of Choice*. Dover Publications, Inc. Mineola, New York, 1973.

[14] J. L. Kelley. *General topology*. Number 21 in International Student Editions. Van Nostrand Reinhold, New York, 1970.

[15] Anthony W. Knapp. *Basic algebra*. Cornestones. Birkhäuser, Boston, 2006.

[16] Kelley, J. L. *The Tychonoff product theorem implies the axiom of choice.* Fund. Math. 37, (1950). 75–76 (MR)

[17] Serge Lang. *Algebra*. Addison Wesley, Reading, third edition, 1993.

[18] Fredrik Meyer, Nadia S. Larsen, *The Banach-Tarski Paradox*, http://folk.uio.no/fredrme/BanachTarski.pdf, University of Oslo.

[19] Pere Pascual, Agustí Roig. *Topologia*. Apunts de topologia.
www.ma1.upc.edu/docencia/assignatures/fme/topologia-fme/apunts/.
Novembre 2014.

[20] Josep Pla i Carrera, *Lliçons de Lògica Matemàtica (Primera i Segona part)*.

[21] H. Rubin, J.E. Rubin, *Equivalents of the Axiom of Choice, II. Studies in logic and the Foundations of Mathematics, Volume 116.*

[22] Eric Schechter. A home page for the Axiom of Choice.
http://www.math.vanderbilt.edu/∼ schectex/ccc/choice.html,
Novembre 2014.

[23] Raymond M. Smullyan, Melvin Fitting. *Set theory and the Continuum Problem*. Dover Publications, Inc. Mineola, New York, 2010.

[24] Robert R. Stoll. *Set theory and Logic*. Dover Publications, Inc. New York, 1979.

[25] Tom Weston, *The Banach-Tarski Paradox*, http://people.math.umass.edu/ weston/oldpapers/banach.pdf.