Universidad Nacional Autónoma de México

facultad de ciencias

Paradojas de Hausdorff-Banach-Tarski

T E S I S

que para obtener el título de:

Matemático

p r e s e n t a:

Ayax Calderón Camacho

Director de Tesis:
Dr. Pierre Michel Bayard

Ciudad Universitaria, CDMX, 2021

Agradecimientos

A mis padres, Lilia y Javier, por su apoyo incondicional en todo lo que he decidido emprender, por alentarme siempre que lo he necesitado y por nunca dejar de creer en mí.

A mi hermano Irving, porque desde que tengo memoria y hasta la fecha ha sido mi ejemplo a seguir, porque en distintas etapas de mi vida me ha presentado una nueva y bonita cara de las matemáticas, por animarme a estudiarlas y por ayudarme a seguir adelante cuando he dudado de mí.

A mi cuñada Claudia, por toda la ayuda que me brindó cuando empecé la carrera y por ser como una hermana para mí.

A Irving y Claudia, a quiénes ya agradecí por separado, pero ahora quiero agradecer en conjunto, por cuidar de mí cuando recién empezaba mi vida universitaria, por todos sus consejos, por todos los momentos divertidos que pasamos viviendo juntos. Gracias porque, a pesar de que estos últimos años han estado muy lejos y nos hemos visto muy poco, nos hemos seguido frecuentando y compartiendo momentos como si la distancia no existiera.

A Marcela, Edwin, Julio, César, Diego y los demás amigos que hice en el camino, y que espero que me sigan acompañando al final de éste. Por todas esas tardes en la biblioteca, por todos los retos que superamos a lo largo de este tiempo, por todas las risas y, sobre todo, por todos los buenos recuerdos que han dejado en mí.

También quiero agradecer a mi asesor, Pierre Bayard, por sus cursos de teoría de la medida y análisis funcional, los cuales sirvieron como cimientos para la realización de esta tesis, por haberme sugerido este tema y por su muy paciente asesoría.

Finalmente, agradezco a mis sinodales, Alberto Saldaña, Natalia Jonard, Judith Campos y Carmen Martínez Adame por tomarse el tiempo de leer con cautela este trabajo y ayudarme a enriquecerlo con sus valiosos comentarios.

Índice general

Introducción

En matemáticas hay un teorema que de manera coloquial se puede interpretar de la siguiente manera:

Teorema. *Una naranja puede cortarse en una cantidad finita de pedazos, y esos pedazos pueden ser reensamblados para obtener dos naranjas idénticas a la original.*

Otra forma de interpretar el mismo teorema, que en lo personal me parece aún más sorprendente, es la siguiente:

Teorema. *Dado un chícharo cualquiera, existe una descomposición de este en una cantidad finita de piezas de manera que al reensamblarlas se obtenga una pelota del tamaño del Sol.*

Este resultado que parece desafiar el sentido común el problema central a desarrollar en esta tesis y lo enunciamos en dos partes.

Teorema (Hausdorff, 1914)**.** *La esfera unitaria de* $\mathbb{R}^3$ *es* $SO(3)$*-paradójica.*

Teorema (Banach-Tarski, 1924)**.** *Dos subconjuntos acotados de* $\mathbb{R}^3$ *con interior no vacío son* $Is^+(\mathbb{R}^3)$*-equidescomponibles, donde* $Is^+(\mathbb{R}^3)$ *es el grupo de las isometrías directas de* $\mathbb{R}^3$.

En esencia, la prueba de estos teoremas se apoya en dos pilares fundamentales:

- El axioma de elección.
- El hecho de que los grupos libres de rango 2 no son promediables.

Resulta que el teorema de Banach-Tarski no sólo es válido en dimensión 3, sino en dimensión $n \geq 3$.

El objetivo de esta tesis es desarrollar con todo detalle el camino presentado por Denis Choimet y Hervé Queffélec en el Capítulo 6 de su libro [1], donde culminan con la prueba del problema principal de este trabajo.

En el Capítulo 1 introducimos las herramientas de álgebra y análisis funcional que serán necesarias para el desarrollo del Capítulo 2, empezando por la definición de acción de un grupo y el concepto de representación. Luego hablamos del *teorema de Hahn-Banach.* En seguida damos un breve repaso acerca de la topología débil-$\star$ generada por una familia de semi-normas sobre un espacio vectorial topológico y finalizamos la sección con un teorema de compacidad en esta topología.

En el Capítulo 2 se desarrolla el concepto de *promedio* definido sobre un conjunto E, el cual simplemente es una medida μ finitamente aditiva y tal que $\mu(E) = 1$. Luego estudiamos el hecho de que los promedios se pueden identificar con funcionales lineales positivas con la finalidad de hacer uso de las herramientas de análisis funcional, específicamente haremos uso de la dualidad entre funcionales lineales y promedios.

En seguida definimos el concepto de *G-invariancia* para los promedios, el cual nos dice que si G es un grupo que actúa sobre E, entonces un promedio es G-invariante si

$$\mu(A) = \mu(gA) \quad \forall A \in \mathscr{P}(E), \forall g \in G.$$

Análogamente se define el concepto para funcionales lineales y se establece la equivalencia de la G-invariancia entre un promedio y su correspondiente funcional lineal.

En el capítulo 3 definimos el concepto de *equidescomponiblidad* entre subconjuntos ajenos de un conjunto y la noción de ser *G-paradójico* con la finalidad de presentar las construcciones paradójicas de Hausdorff y Banach-Tarski, esto haciendo uso de toda la teoría desarrollada en los capítulos anteriores.

Una de las utilidades de estas construcciones paradójicas es responder a la pregunta de la existencia de una extensión de la medida de Lebesgue en dimensión $d \geq 3$ que sea finitamente aditiva e invariante bajo isometrías. Existen numerosas consecuencias de estas construcciones en diversas áreas de las matemáticas, tales como la teoría de la medida, teoría de grupos, geometría, teoría de conjuntos y lógica. Sin embargo, el marco de esta tesis no es lo suficientemente amplio para explorarlas a todas.

Asimismo, al final del texto se incluyen dos apéndices. En el primero de ellos se presenta el *teorema de Steinhaus*, mismo que se usa para probar que cualquier conjunto de medida positiva contiene un subconjunto no medible en el sentido de Lebesgue.

Para concluir, en el último apéndice se presenta una tercera versión del teorema de Hahn-Banach, esta dada ahora por los matemáticos Agnew y Morse.

Motivación del problema: La idea de Vitali

A principios del siglo XX la escuela francesa encabezada por Émile Borel buscaba una medida[1] λ que cumpliera las siguientes propiedades.

1. $\lambda : \mathscr{P}(\mathbb{R}^d) \to [0, +\infty]$, es decir, que esté definida sobre todos los subconjuntos de $\mathbb{R}^d$.

2. $\lambda\left(\bigcup_{k\in\mathbb{N}} A_k\right) = \sum_{k\in\mathbb{N}} \lambda(A_k)$, donde $(A_k)_{k\in\mathbb{N}} \subseteq \mathscr{P}(\mathbb{R}^d)$ es una sucesión de subconjuntos de $\mathbb{R}^d$ ajenos por pares. A esta propiedad se le conoce como σ-aditividad.

3. $\lambda([0,1]^d) = 1$.

4. $\lambda(g(A)) = \lambda(A)$ para cualesquiera $A \in \mathscr{P}(\mathbb{R}^d)$ y $g \in Is(\mathbb{R}^d)$, donde $Is(\mathbb{R}^d)$ es el grupo de isometrías de $\mathbb{R}^d$.

El objetivo de esto era generalizar la noción de longitud, área y volumen que se tiene en dimensiones 1, 2 y 3, respectivamente.

En 1901, mientras escribía su tesis doctoral, Lebesgue definió la medida que hoy lleva su nombre y casi da al clavo con lo que se buscaba. Sin embargo, la medida de Lebesgue no satisface la propiedad 1, en el sentido de que esta medida no está definida sobre todos los subconjuntos de $\mathbb{R}^d$, sino sobre la σ-álgebra de Lebesgue $\mathscr{L}$.

En 1905, Vitali demostró que si se supone el axioma de elección y la existencia de una medida que satisfaga las cuatro propiedades anteriores, se llega a una contradicción, destruyendo así las esperenzas de todos aquellos que deseaban encontrar una teoría oculta detrás de esta pregunta. Como ejemplo de este hecho tenemos el siguiente:

Teorema. 1. *No existe una medida de probabilidad definida sobre $\mathscr{P}(\mathbb{S}^1)$, (donde $\mathbb{S}^1$ es la circunferencia unitaria en $\mathbb{C}$) que sea invariante bajo rotaciones.*

2. *No existe una medida positiva ν definida sobre $\mathscr{P}(\mathbb{R})$ que sea invariante bajo traslaciones y tal que $\nu([0,1]) = 1$.*

[1] Las definiciones correspondientes a teoría de la medida se encuentran en la sección 1.2.

Demostración. 1. Diremos que $p \in \mathbb{S}^1$ es un punto racional si p es de la forma $e^{2\pi i x}$ con $x \in \mathbb{Q}$. Al conjunto de dichos puntos los llamaremos G. Resulta que G es un grupo multiplicativo, pues $\mathbb{Q}$ es un subgrupo de $\mathbb{R}$.
Con G podemos asociar una relación de equivalencia definida por

$$x \sim y \Leftrightarrow x^{-1}y \in G.$$

Veamos que en efecto $\sim$ es una relación de equivalencia.
Reflexividad. $x^{-1}x = 1 = e^{2\pi i(0)}$, entonces $x \sim x$.
Simetría. Supongamos que $x \sim y$ entonces $x^{-1}y \in G$, es decir, existe $k \in \mathbb{Q}$ tal que $e^{2\pi i k} = x^{-1}y$. Luego $y^{-1}x = e^{2\pi i(-k)}$ con $-k \in \mathbb{Q}$. Así $y \sim x$.
Transitividad. Supongamos que $x \sim y$, $y \sim z$, entonces $x^{-1}y \in G$, $y^{-1}x \in G$. Luego $x^{-1}yy^{-1}z = x^{-1}z \in G$. Así, $x \sim z$.
Se concluye que $\sim$ es relación de equivalencia.

Note que la clase de equivalencia de $x \in \mathbb{S}^1$ es xG, pues si $y \in [x]_\sim$, entonces $x^{-1}y \in G$ y así, $y \in xG$. Conversamente, si $y \in xG$, entonces existe $z \in G$ tal que $y = xz$. Luego, $x^{-1}y = z \in G$ y así $y \in [x]_\sim$. Concluimos que $xG = [x]_\sim$.

El axioma de elección nos garantiza la existencia de $M \subseteq \mathbb{S}^1$ tal que M intersecta a cada clase de equivalencia en exactamente un punto.
Luego

$$\mathbb{S}^1 = \bigcup_{g \in G} gM,$$

donde los gM son ajenos por pares.

Ahora supongamos que existe una medida de probabilidad μ sobre $\mathscr{P}(\mathbb{S}^1)$ que es invariante bajo rotaciones. Como G es numerable, entonces

$$1 = \mu(\mathbb{S}^1) = \mu\left(\bigcup_{g \in G} gM\right) = \textstyle\sum_{g \in G} \mu(gM) = \sum_{g \in G} \mu(M) \in \{0, +\infty\}^{2},$$

lo cual es una contradicción. Se sigue el resultado deseado.

2. Supongamos, para generar una contradicción, que tal medida ν existe. Tenemos la biyección natural

$$\begin{aligned} u : [0,1) &\to \mathbb{S}^1 \\ x &\mapsto e^{2\pi i x} \end{aligned}$$

Para cada subconjunto A de $\mathbb{S}^1$ definimos $\mu(A) = \nu(u^{-1}(A))$. Veamos que μ es una medida sobre $\mathscr{P}(\mathbb{S}^1)$.

[2] $\displaystyle\sum_{g \in G} \mu(M) = 0$ cuando $\mu(M) = 0$ y $\displaystyle\sum_{g \in G} \mu(M) = +\infty$ cuando $\mu(M) > 0$.

a) $\mu(\varnothing) = \nu(u^{-1}(\varnothing)) = \nu(\varnothing) = 0$.

b) Sea $(A_n)_{n\in\mathbb{N}} \subseteq \mathscr{P}(\mathbb{S}^1)$ una sucesión de conjuntos ajenos por pares. Entonces

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \nu\left(u^{-1}\left(\bigcup_{n\in\mathbb{N}} A_n\right)\right) = \nu\left(\bigcup_{n\in\mathbb{N}} u^{-1}(A_n)\right)$$
$$= \sum_{n\in\mathbb{N}} \nu(u^{-1}(A_n)) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

Se sigue que μ es una medida sobre $\mathscr{P}(\mathbb{S}^1)$.

Más aún, usando la invariancia bajo traslaciones tenemos para todo $n \geq 1$

$$\nu\left(\left\{\frac{1}{n}\right\}\right) = \nu(\{1\})$$

y como $\nu([0,1])$ es finito, entonces $\nu(\{1\}) = 0$.
Así,

$$\mu(\mathbb{S}^1) = \nu(u^{-1}(\mathbb{S}^1)) = \nu([0,1)) = \nu([0,1]) = 1.$$

Finalmente verifiquemos que μ es invariante bajo rotaciones.

Si tomamos $\alpha \in [0,1), A \subseteq \mathbb{S}^1$ y $B = u^{-1}(A)$ tenemos

$$\begin{aligned}
\mu(e^{2\pi i\alpha}A) &= \mu(\{e^{2\pi i(\alpha+x)} : x \in B\} \\
&= \mu(\{e^{2\pi i(\alpha+x)} : x \in B \cap [0, 1-\alpha)\}) \\
&\quad + \mu(\{e^{2\pi i(\alpha+x)} : x \in B \cap [1-\alpha, 1)\}) \\
&= \mu(\{e^{2\pi ix} : x \in (B+\alpha) \cap [\alpha, 1)\}) \\
&\quad + \mu(\{e^{2\pi ix} : x \in (B+\alpha-1) \cap [0, \alpha)\}) \\
&= \nu((B+\alpha) \cap [\alpha,1)) + \nu((B+\alpha-1) \cap [0,\alpha)) \\
&= \nu(B) \\
&= \nu(u^{-1}(A)) \\
&= \mu(A).
\end{aligned}$$

Pero esto resulta absurdo, ya que contradice el inciso 1 de esta proposición.

$\square$

Capítulo 1

Preliminares

1.1. Una pizca de álgebra

Definición 1.1.1. Sean X un conjunto y G un grupo. Diremos que G *actúa sobre* X si existe una función

$$\begin{aligned} G \times X &\to X \\ (g, x) &\mapsto gx \end{aligned}$$

tal que

1. $(gh)x = g(hx) \quad \forall g, h \in G,\ \forall x \in X.$
2. $ex = x \quad \forall x \in X$, donde e es el elemento neutro de G.

Definición 1.1.2. Sea G un grupo y V un espacio lineal. Una *acción lineal* de G en V, también llamada *representación* de G, es una acción de G sobre V tal que para cada $g \in G$, la función

$$\begin{aligned} V &\to V \\ x &\mapsto gx \end{aligned}$$

es lineal, o equivalentemente, define un homomorfismo de G al grupo de automorfismos lineales de V.

1.2. Una cucharada de medida

Sean X un conjunto y $\mathscr{P}(X)$ la potencia de X.

Definición 1.2.1. Diremos que $\mathscr{A} \subseteq \mathscr{P}(X)$ es una *σ-álgebra* de X si se satisfacen las siguientes condiciones:

1. $\varnothing \in \mathscr{A}$.
2. Si $A \in \mathscr{A}$, entonces $X \backslash A \in \mathscr{A}$.

3. Si $(A_n)_{n\in\mathbb{N}} \subseteq \mathscr{A}$, entonces $\displaystyle\bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}$.

Adicionalmente, diremos que el par $(X, \mathscr{A})$ es un *espacio medible* y a los elementos de $\mathscr{A}$ los llamamos *subconjuntos medibles*.

Definición 1.2.2. Sea $(X, \mathscr{A})$ un espacio medible. Una *medida* sobre $(X, \mathscr{A})$ es una función

$$\mu : \mathscr{A} \to [0, +\infty]$$

que satisface las siguientes propiedades:

1. $\mu(\varnothing) = 0$.
2. μ es σ-aditiva, es decir, si $(A_n)_{n\in\mathbb{N}} \subseteq \mathscr{A}$ es una sucesión de conjuntos ajenos por pares, entonces
$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n).$$
A la tercia $(X, \mathscr{A}, \mu)$ la llamamos *espacio de medida*.

Definición 1.2.3. Una *medida de probabilidad* en un espacio de medida $(X, \mathscr{A})$ es una función

$$\mu : \mathscr{A} \to [0, 1]$$

que satisface $\mu(X) = 1$.

1.3. Un puñado de análisis funcional

1.3.1. Teorema de Hahn-Banach

Sea E un espacio lineal sobre $\mathbb{R}$. Recordemos que una *funcional* es una función definida sobre E, o sobre un subespacio de E, *con valores en* $\mathbb{R}$.

Teorema 1.3.1. *[Helly, Hahn-Banach forma analítica] Sea $p : E \to \mathbb{R}$ una función que satisface*[1]

1. $p(\lambda x) = \lambda p(x) \qquad \forall x \in E \quad y \quad \forall \lambda > 0,$
2. $p(x + y) \leq p(x) + p(y) \qquad \forall x, y \in E.$

Sea $G \subseteq E$ un subespacio lineal y $g : G \to \mathbb{R}$ una funcional lineal tal que

$$g(x) \leq p(x) \quad \forall x \in G.$$

Bajo estas hipótesis, existe una funcional lineal f definida sobre E que extiende a g, es decir, $g(x) = f(x) \quad \forall x \in G$ y tal que

$$f(x) \leq p(x) \quad \forall x \in E.$$

[1] Una funcional que satisface las propiedades 1 y 2 a veces es llamada *funcional de Minkowski.*

La prueba de este teorema es muy técnica, por lo cual no se incluye en esta tesis, pero puede consultarse en [6].

En el capítulo 2 veremos una versión del teorema de Hahn-Banach, donde la funcional lineal que se extiende es invariante con respecto a la acción de un grupo G. Asimismo, en el Apéndice B se puede encontrar una tercera versión de este teorema (¡existen más!), en esta ocasión la funcional que se extiende es invariante con respecto a una familia de operadores lineales y continuos.

1.3.2. Topología débil-$\star$

Sea E un espacio de Banach sobre $\mathbb{R}$.

Notación. *Denotamos por $E^\star$ al espacio dual de E, es decir, al espacio de todas las funcionales lineales continuas sobre E; la norma dual de E está definida por*

$$\|f\|_{E^\star} = \sup \{|f(x)| : \|x\| \leq 1,\ x \in E\}.$$

Observación 1.3.2. $E^\star$ es un espacio de Banach aún cuando E no lo sea.

Definimos, para todo $x \in E$ y todo $\varphi \in E^\star$

$$p_x(\varphi) = |\varphi(x)|.$$

$p_x(\varphi)$ es una semi-norma sobre $E^\star$.

Definición 1.3.3. $\sigma(E^\star, E)$ es la topología sobre $E^\star$ definida por la familia de semi-normas $(p_x(\varphi))_{x\in E}$ y es la topología más gruesa tal que $\forall x \in E$

$$\begin{aligned} E^\star &\to \mathbb{R} \\ \varphi &\mapsto \varphi(x) \end{aligned}$$

es continua. A $\sigma(E^\star, E)$ se le conoce como la *topología débil-$\star$*.

Teorema 1.3.4. *[Banach–Alaoglu] La bola unitaria cerrada*

$$B_{E^\star} = \{f \in E^\star : \|f\|_{E^\star} \leq 1\}$$

es compacta en la topología débil-$\star$ $\sigma(E^\star, E)$.

Demostración. Recordemos que por definición, la topología producto sobre $\prod_{i\in I} X_i$ es la topología con menos abiertos tal que las proyecciones

$$p_j : \prod_{i\in I} X_i \to X_j$$

son continuas para toda $j \in I$.

Consideremos el producto cartesiano $\mathbb{R}^E = \{f : E \to \mathbb{R}\}$ el cual consiste de todas las funciones $E \to \mathbb{R}$. Por definición se tiene que

$$\mathbb{R}^E = \prod_{x \in E} \mathbb{R}_x,$$

donde $\mathbb{R}_x = \mathbb{R}\ \forall x$. Con esta notación, cada función $f \in \mathbb{R}^E$ se puede expresar de la forma $f = (f(x))_{x \in E}$.

Por definición, la topología producto sobre $\mathbb{R}^E$ es la topología con menos abiertos tal que la función

$$\begin{aligned} \mathbb{R}^E &\to \mathbb{R} \\ f &\mapsto f(x) \end{aligned}$$

es continua para todo $x \in E$.

La topología inducida sobre $E^\star \subseteq \mathbb{R}^E$ es la topología débil-$\star$ $\sigma(E^\star, E)$. Tenemos que $B_{E^\star} = F_1 \cap F_2$, donde

$$F_1 = \{f \in \mathbb{R}^E : |f(x)| \leq \|x\| \quad \forall x \in E\}$$

y

$$F_2 = \{f \in \mathbb{R}^E : f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)\ \ \forall (x, y, \lambda, \mu) \in E \times E \times \mathbb{R} \times \mathbb{R}\}.$$

Mostremos que F_1 y F_2 son cerrados en $\mathbb{R}^E$:

Para $x \in E$ fijo, se tiene que la aplicación $f \overset{j_x}{\longmapsto} f(x)$ es continua en $\mathbb{R}^E$ y además

$$F_1 = \bigcap_{x \in E} j_x^{-1}(\overline{D}(0, \|x\|)),$$

(donde $\overline{D}(0, \|x\|)$ es el disco con centro en 0 y radio $\|x\|$) lo que muestra que F_1 es cerrado en $\mathbb{R}^E$.

De manera similar se tiene que si tomamos $x, y \in E$ y $\lambda, \mu \in \mathbb{R}$ fijos, entonces la aplicación

$$e_{x,y,\lambda,\mu} : f \mapsto f(\lambda x + \mu y) - \lambda f(x) - \mu f(y)$$

es continua en $\mathbb{R}^E$ y además

$$F_2 = \bigcap_{x,y,\lambda,\mu} e_{x,y,\lambda,\mu}^{-1}(0),$$

lo cual muestra que F_2 es cerrado en $\mathbb{R}^E$. Por lo tanto, $B_{E^\star} = F_1 \cap F_2$ es cerrado en $\mathbb{R}^E$.

Consideremos ahora para $r \geq 0$, el disco cerrado

$$\overline{D}_r = \{\lambda \in \mathbb{R} : |\lambda| \leq r\}.$$

Tenemos por definición que

$$B_{E^\star} \subseteq \prod_{x \in E} \overline{D}_{\|x\|},$$

ya que $f \in B_{E^\star}$ implica que $|f(x)| \leq \|x\| \quad \forall x \in E$. Además el conjunto $\prod_{x \in E} \overline{D}_{\|x\|}$ es compacto, por el teorema de Tychonoff.

Finalmente, el subconjunto $B_{E^\star}$ es cerrado en el compacto $\prod_{x \in E} \overline{D}_{\|x\|}$ y por lo tanto es compacto. $\square$

Observación 1.3.5. El teorema de Banach-Alaouglu también depende del axioma de elección, pues en la prueba se utiliza el teorema de Tychonoff, que es equivalente al axioma de elección.

Capítulo 2

Promedios

El teorema de Vitali muestra que a una función $\mu : \mathscr{P}(\mathbb{R}) \to [0, +\infty]$ tal que $\mu([0,1]) = 1$, no podemos pedirle simultáneamente invariancia bajo isometrías y σ-aditividad. De esta manera resulta natural debilitar un poco esta última condición, lo cual nos lleva a las siguientes definiciones:

Definición 2.0.1. Sea G un grupo mutiplicativo que actúa sobre un conjunto no vacío E.

- Una *medida finitamente aditiva* es una función $\mu : \mathscr{P}(E) \to [0, +\infty]$ que satisface

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i)$$

  para cualesquiera $A_1, \ldots, A_n \subseteq E$ ajenos por pares.

- Decimos que la medida anterior es *G-invariante* si adicionalmente satisface

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

  para todo $g \in G$ y $A \subseteq E$.

- Un *promedio* es una medida finitamente aditiva μ normalizada por la condición

$$\mu(E) = 1.$$

2.1. Interpretación en términos de funcionales lineales

Queremos establecer algunos teoremas de existencia sobre promedios invariantes. Para lograr esto, primero necesitamos interpretar los promedios de E en términos de funcionales lineales positivas, lo cual nos permitirá hacer uso de las herramientas desarrolladas por Banach y sus sucesores: el análisis funcional y específicamente, la dualidad entre funcionales lineales y promedios.

Sea $(\ell^\infty(E), \|\cdot\|_\infty)$ el álgebra de Banach de las funciones acotadas de E a $\mathbb{R}$. Sea φ una funcional lineal en $\ell^\infty(E)$, positiva (en el sentido de que $\varphi(f) \geq 0$ si $f \geq 0$) y normalizada por $\varphi(\mathbb{1}) = 1$, donde $\mathbb{1}$ es la función constante 1. Dicha funcional lineal es automáticamente continua:

En efecto, si tomamos $f \in \ell^\infty(E)$, por definición de $\|\cdot\|_\infty$, se tiene que

$$|f(x)| \leq \|f\|_\infty \quad \forall x \in E,$$

entonces

$$0 \leq \|f\|_\infty \cdot \mathbb{1} \pm f.$$

Como φ es positiva, se sigue que

$$0 \leq \varphi(\|f\|_\infty \cdot \mathbb{1} \pm f) = \|f\|_\infty \varphi(\mathbb{1}) \pm \varphi(f) = \|f\|_\infty \pm \varphi(f)$$

y así

$$|\varphi(f)| \leq \|f\|_\infty .$$

Más aún, la función

$$\begin{aligned} \mu_\varphi : \mathscr{P}(E) &\to [0,1] \\ A &\mapsto \varphi(\mathbb{1}_A), \end{aligned}$$

donde $\mathbb{1}_A$ denota la función indicadora de A,

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{si} \quad x \in A \\ 0 & \text{si} \quad x \notin A, \end{cases}$$

es un promedio sobre E. En efecto: Sean $A_1, \ldots, A_n$ ajenos por pares y $U := \bigcup_{i=1}^{n} A_i$. Como los $(A_i)_{i=1}^n$ son ajenos por pares, se tiene que $\mathbb{1}_U = \sum_{i=1}^{n} \mathbb{1}_{A_i}$, pues tener $x \in \bigcup_{i=1}^{n} A_i$ garantiza la existencia de un único $1 \leq i \leq n$ tal que $x \in A_i$.
Entonces

$$\mu_\varphi\left(\bigcup_{i=1}^{n} A_i\right) = \varphi(\mathbb{1}_U) = \sum_{i=1}^{n} \varphi(\mathbb{1}_{A_i}) = \sum_{i=1}^{n} \mu_\varphi(A_i).$$

Finalmente notemos que $\mu_\varphi(E) = \varphi(\mathbb{1}_E) = \varphi(\mathbb{1}) = 1$. Se sigue que μ_φ es un promedio sobre E.

Definición 2.1.1. Decimos que f es una función simple si se puede escribir como

$$f = \sum_{i=1}^{r} a_i \mathbb{1}_{A_i},$$

donde $(A_i)_{i=1}^r$ es una colección de subconjuntos de E ajenos por pares, tales que $\bigcup_{i=1}^{r} A_i = E$ y los $a_1, \ldots a_r$ son números reales o complejos.
Equivalentemente, f es una función simple si toma solamente una cantidad finita de valores.

Proposición 2.1.2. El espacio de las funciones simples es cerrado bajo sumas, productos y productos por un escalar.

Demostración.

- La función constante cero solamente toma un valor y por lo tanto es simple.
- Sea f una función simple y $\lambda \in \mathbb{R}$ un escalar, entonces

$$f = \sum_{i=1}^{r} a_i \mathbb{1}_{A_i}$$

y luego

$$(\lambda f) = \sum_{i=1}^{r} (\lambda a_i) \mathbb{1}_{A_i}.$$

Se sigue que (λf) es una función simple.

- Sean $f, g : E \to \mathbb{R}$ funciones simples dadas por

$$f = \sum_{i=1}^{r} a_i \mathbb{1}_{A_i} \quad \text{y} \quad g = \sum_{j=1}^{s} b_j \mathbb{1}_{B_j}.$$

Como f y g tienen a E como dominio y son simples, entonces $(A_i)_{i=1}^{r}$ y $(B_j)_{j=1}^{s}$ son colecciones de E ajenos por pares y tales que $\bigcup_{i=1}^{r} A_i = E = \bigcup_{j=1}^{s}$. Definimos

$$C_{ij} = A_i \cap B_j.$$

La colección de los C_{ij} no necesariamente es una partición de E, ya que algunas C_{ij} podrían ser vacías[1]. Sin embargo, veremos que los C_{ij} son ajenos por pares y que su unión es E.
Notemos que $A_i \subseteq E = \bigcup_{j=1}^{s} B_j$ y $A_i = A_i \cap E = A_i \cap \left(\bigcup_{j=1}^{s} B_j \right) = \bigcup_{j=1}^{s} C_{ij}$. Similarmente, $B_j = \bigcup_{i=1}^{r} C_{ij}$.

Entonces,

$$E = \bigcup_{i=1}^{r} A_i = \bigcup_{i=1}^{r} \bigcup_{j=1}^{s} C_{ij}.$$

Finalmente si $(i, j) \neq (k, l)$, entonces supongamos sin pérdida de generalidad que $i \neq k$ de donde

$$C_{ij} \cap C_{kl} = \varnothing,$$

[1] Esto no representa un problema, ya que si C_{ij} es vacío, entonces la función $\mathbb{1}_{C_{ij}}$ es idénticamente cero.

pues $C_{ij} \cap C_{kl} = A_i \cap B_j \cap A_k \cap B_l \subseteq A_i \cap A_k = \varnothing$.

Como los (C_{ij}) son ajenos por pares, entonces

$$\mathbb{1}_{A_i} = \sum_{j=1}^{s} \mathbb{1}_{C_{ij}} \quad \text{y} \quad \mathbb{1}_{B_j} = \sum_{i=1}^{r} \mathbb{1}_{C_{ij}}.$$

Así,

$$f = \sum_{i=1}^{r} \sum_{j=1}^{s} a_i \mathbb{1}_{C_{ij}} \quad \text{y} \quad g = \sum_{j=1}^{s} \sum_{i=1}^{r} b_j \mathbb{1}_{C_{ij}}.$$

Finalmente,

$$f + g = \sum_{i=1}^{r} \sum_{j=1}^{s} (a_i + b_j) \mathbb{1}_{C_{ij}}$$

y

$$fg = \sum_{i=1}^{r} \sum_{j=1}^{s} (a_i b_j) \mathbb{1}_{C_{ij}}$$

son funciones simples.

□

Lema 2.1.3. *El conjunto de las funciones simples de $\ell^\infty(E)$ es denso en $\ell^\infty(E)$.*

Demostración. Sea $f \in \ell^\infty(E)$. Multiplicando por un escalar, de ser necesario, podemos suponer que $\|f\|_\infty \leq 1$. Para $n \geq 1$, tomamos

$$f_n = \sum_{k=-n}^{n-1} \frac{k}{n} \mathbb{1}_{f^{-1}\left(\left[\frac{k}{n}, \frac{k+1}{n}\right)\right)} + \mathbb{1}_{f^{-1}(\{1\})}.$$

Sea $\varepsilon > 0$, entonces existe $N \in \mathbb{N}$ tal que $0 < \frac{1}{N} < \varepsilon$.
Luego $\|f - f_N\|_\infty \leq \frac{1}{N} < \varepsilon$. Así, el conjunto de funciones simples es denso en $\ell^\infty(E)$. □

Teorema 2.1.4. *La función $\varphi \mapsto \mu_\varphi$ es una biyección del conjunto de funcionales lineales normalizadas y positivas sobre $\ell^\infty(E)$ al conjunto de promedios sobre E.*

Demostración. Empezaremos probando que la función $\varphi \mapsto \mu_\varphi$ es suprayectiva. Sea μ un promedio sobre E y $f : E \to \mathbb{R}$ una función simple.
Denotemos por $\{x_1, \ldots, x_r\}$ los valores distintos que toma f y definimos

$$\varphi_\mu(f) = \sum_{i=1}^{r} x_i \mu(f^{-1}(\{x_i\})).$$

Veamos que φ_μ es una funcional lineal en el espacio de las funciones simples de E en $\mathbb{R}$.
Sean $f, g \in \ell^\infty(E)$ funciones simples; $\{x_1, \ldots, x_r\}$ los valores distintos que toma f y $\{y_1, \ldots, y_s\}$

los valores distintos que toma g.
Sean $A_i = f^{-1}(\{x_i\})$ y $B_j = g^{-1}(\{y_j\})$, entonces

$$\varphi_\mu(f) = \sum_{i=1}^{r} x_i\mu(A_i) \quad \text{y} \quad \varphi_\mu(g) = \sum_{j=1}^{s} y_j\mu(B_j).$$

Reciclando la notación y las ideas de la prueba de la linealidad del espacio de funciones simples tenemos que
$A_i = \bigcup_{j=1}^{s} C_{ij}$, donde los C_{ij} son ajenos por pares, entonces $\mu(A_i) = \sum_{j=1}^{s} \mu(C_{ij})$. Similarmente se llega a que $\mu(B_j) = \sum_{i=1}^{r} \mu(C_{ij})$, pues $B_j = \bigcup_{i=1}^{r} C_{ij}$.
Entonces,

$$\varphi_\mu(f) = \sum_{i=1}^{r}\sum_{j=1}^{s} x_i\mu(C_{ij}) \quad \text{y} \quad \varphi_\mu(g) = \sum_{j=1}^{s}\sum_{i=1}^{r} y_j\mu(C_{ij}).$$

Luego,

$$\varphi_\mu(f) + \varphi_\mu(g) = \sum_{i=1}^{r}\sum_{j=1}^{s}(x_i + y_j)\mu(C_{ij}), \quad \text{donde } C_{ij} = (f+g)^{-1}(\{x_i + y_j\}).$$

Así,

$$\varphi_\mu(f) + \varphi_\mu(g) = \sum_{i=1}^{r}\sum_{j=1}^{s}(x_i + y_j)\mu((f+g)^{-1}(\{x_i + y_j\})) = \varphi_\mu(f+g).$$

Por lo tanto, φ_μ es una funcional lineal sobre el espacio de las funciones simples en $\ell^\infty(E)$.

Notemos además que

$$\varphi_\mu(1) = 1 \cdot \mu(1^{-1}(\{1\})) = \mu(E) = 1.$$

Finalmente, si $f \in \ell^\infty(E)$ es simple y positiva, entonces el conjunto $\{x_1, \ldots, x_r\}$ de todos los valores distintos que toma f, está contenido en $\mathbb{R}_{\geq 0}$, por lo que $x_i\mu(f^{-1}) \geq 0$, y así $\varphi_\mu(f) = \sum_{i=1}^{r} x_i\mu(f^{-1}(\{x_i\})) \geq 0$.
Como φ_μ es funcional lineal positiva sobre el espacio lineal de las funciones simples de E a $\mathbb{R}$, entonces es continua. Por el lema anterior, podemos extender φ_μ a una funcional lineal continua sobre $\ell^\infty(E)$ (que aún denotaremos como φ_μ), positiva y que satisface $\varphi_\mu(1) = 1$.

Si $f \in \ell^\infty(E)$, el número real $\varphi_\mu(f)$ se puede denotar por[2]

$$\int_E f d\mu = \int_E f(x) d\mu(x).$$

[2]Estamos trabajando con una teoría *finita* de integración, donde la diferencia con respecto a la teoría usual es la ausencia de los teoremas límite (convergencia dominada, etc.), los cuales requieren medidas σ-aditivas.

Mostremos que las dos correspondencias $M : \varphi \mapsto \mu_\varphi$ y $\Phi : \mu \mapsto \varphi_\mu$ que acabamos de describir, son inversas una de la otra.

Sea $A \subseteq E$, entonces

$$(\Phi \circ M)(\varphi)(\mathbb{1}_A) = \Phi(M(\varphi(\mathbb{1}_A))) = \Phi(\mu_\varphi)(\mathbb{1}_A) = \varphi_{\mu_\varphi}(\mathbb{1}_A) = \int_E \mathbb{1}_A d\mu_\varphi = \mu_\varphi(A) = \varphi(\mathbb{1}_A).$$

Por otro lado

$$(M \circ \Phi)(\mu)(A) = M(\Phi(\mu(A))) = M(\varphi_\mu(A)) = \mu_{\varphi_\mu}(A) = \varphi_\mu(\mathbb{1}_A) = \int_E \mathbb{1}_A d\mu = \mu(A).$$

Así, $M \circ \Phi = id$.

Se sigue que M y Φ son inversas una de la otra. □

Ahora veremos cómo la eventual G-invariancia de un promedio μ afecta a la funcional lineal φ_μ.

Para esto necesitamos la siguiente notación:

Sean $f \in \ell^\infty(E)$ y $g \in G$ que actúa en E, diremos que ${}_gf$ es el *trasladado* de f, definida por:

$$\begin{aligned} {}_gf : E &\to \mathbb{R} \\ x &\mapsto f(g^{-1}x). \end{aligned}$$

Decimos que una funcional lineal φ sobre $\ell^\infty(E)$ es G-*invariante* si

$$\varphi({}_gf) = \varphi(f) \text{ para } f \in \ell^\infty(E) \text{ y } g \in G.$$

Por la linealidad y la continuidad de φ, y el lema anterior, esto es equivalente a

$$\varphi({}_g(\mathbb{1}_A)) = \varphi(\mathbb{1}_A) \text{ para } A \subseteq E \text{ y } g \in G$$

o

$$\varphi((\mathbb{1}_{gA})) = \varphi(\mathbb{1}_A) \text{ para } A \subseteq E \text{ y } g \in G.$$

Por lo que finalmente

$$\mu_\varphi(gA) = \mu_\varphi(A) \text{ para } A \subseteq E \text{ y } g \in G.$$

Obtenemos la siguiente proposición:

Proposición 2.1.5. Un promedio μ es G-invariante si y sólo si φ_μ es una funcional lineal G-invariante.

Demostración. $\Rightarrow]$ Sea μ un promedio G-invariante. Tenemos que

$$\varphi_\mu({}_gf) = \sum_i x_i\mu({}_gf^{-1}(\{x_i\})) = \sum_i x_i\mu(gA_i) = \sum_i x_i\mu(A_i) = \sum_i x_i\mu(f^{-1}(\{x_i\})) = \varphi_\mu(f).$$

$\Leftarrow]$ Sean φ_μ una funcional lineal G-invariante, $g \in G$ y $A \subseteq E$. Tenemos que

$$\mu(gA) = \int_E \mathbb{1}_{gA} d\mu = \varphi_\mu(\mathbb{1}_{gA}) = \varphi_\mu({}_g(\mathbb{1}_A)) = \varphi_\mu(\mathbb{1}_A) = \int_E \mathbb{1}_A d\mu = \mu(A).$$

Por lo tanto μ es un promedio G-invariante.
Se concluye el resultado deseado. □

Observación 2.1.6. En general, la G-invariancia de un promedio μ sobre E también se puede escribir como

$$\int_E f(gx)d\mu(x) = \int_E f(x)d\mu(x)$$

para $g \in G$, $f \in \ell^\infty(E)$.

2.2. Grupos promediables

Primero construiremos promedios invariantes en el grupo G y luego los transportamos al conjunto E sobre el cual G actúa.

Definición 2.2.1. Decimos que un grupo G es *promediable* si, actuando sobre sí mismo mediante traslaciones izquierdas, admite un promedio G-invariante.

Ejemplo 2.2.2. *Un grupo finito es promediable: basta con tomar para todo $A \subseteq G$*

$$\mu(A) = \frac{|A|}{|G|}.$$

Proposición 2.2.3. Sea G un grupo que actúa sobre un conjunto E, y $f : G \to E$ una función tal que

$$f(gh) = gf(h) \quad \text{para todo } (g,h) \in G^2.$$

Entonces, si μ es un promedio G-invariante sobre G, el *promedio inducido por f* definido por

$$\mu_f(X) = \mu(f^{-1}(X)) \quad \text{para todo } X \subseteq E$$

es un promedio G-invariante sobre E.

Demostración. Primero veamos que μ_f es un promedio sobre E.
Sean $A_1, \ldots, A_n \subseteq E$ ajenos por pares, entonces

$$\mu_f\left(\bigcup_{i=1}^n A_i\right) = \mu\left(f^{-1}\left(\bigcup_{i=1}^n A_i\right)\right) = \mu\left(\bigcup_{i=1}^n f^{-1}(A_i)\right) = \sum_{i=1}^n \mu(f^{-1}(A_i)) = \sum_{i=1}^n \mu_f(A_i).$$

Además μ_f está normalizada, pues

$$\mu_f(E) = \mu(f^{-1}(E)) = \mu(G) = 1.$$

Ahora verificamos la G-invariancia de μ_f: Sean $g \in G$ y $X \subseteq E$, tenemos que

$$f^{-1}(gX) = gf^{-1}(X).$$

$\subseteq$] Sea $h \in f^{-1}(gX)$, entonces $f(h) \in gX$ y luego $g^{-1}f(h) \in X$, de donde $f(g^{-1}h) \in X$, luego $g^{-1}h \in f^{-1}(X)$ y finalmente $h \in gf^{-1}(X)$.
Así, $f^{-1}(gX) \subseteq gf^{-1}(X)$.
$\supseteq$]Conversamente, si tomamos $k \in gf^{-1}(X)$ entonces existe $h \in f^{-1}(X)$ tal que $k = gh$.
Como $h \in f^{-1}(X)$, entonces $f(h) \in X$. Luego $f(k) = f(gh) = gf(h) \in gX$, de donde $k \in f^{-1}(gX)$. Así, $gf^{-1}(X) \subseteq f^{-1}(gX)$.
Se sigue la igualdad deseada.

Finalmente notemos que

$$\mu_f(gX) = \mu(f^{-1}(gX)) = \mu(gf^{-1}(X)) = \mu(f^{-1}(X)) = \mu_f(X).$$

Se concluye que μ_f es un promedio G-invariante sobre E. □

Una aplicación del resultado anterior es la siguiente:

Sea H un subgrupo normal de G. El grupo G actúa sobre el cociente G/H mediante traslaciones izquierdas ($g\overline{h} := \overline{gh}$, donde $\overline{h} \in G/H$), y la proyección canónica $\rho : G \to G/H$ satisface $\rho(gh) = g\rho(h)$ para $(g, h) \in G^2$.

Como consecuencia de lo anterior tenemos el siguiente resultado:

Proposición 2.2.4. Cualquier cociente de un grupo promediable es promediable.

Proposición 2.2.5. Cualquier subgrupo de un grupo promediable es promediable.

Demostración. Sea G un grupo promediable y H un subgrupo de G. Definamos la siguiente relación de equivalencia sobre G:

$$x \sim y \Leftrightarrow yx^{-1} \in H.$$

Notemos que la clase de equivalencia de x es Hx. Gracias al axioma de elección podemos tomar un subconjunto M de G tal que M intersecta a cada clase de equivalencia en exactamente un punto.

Sea $A \subseteq H$, definimos

$$\begin{aligned} \nu : \mathscr{P}(H) &\to [0, +\infty] \\ A &\mapsto \mu(AM), \end{aligned}$$

donde $AM = \{am : (a,m) \in A \times M\}$.

Veamos que ν es un promedio H-invariante sobre H:
Sean $A_1, \ldots, A_n \subseteq H$ ajenos por pares. Notemos que los $A_1M, \ldots, A_nM$ también son ajenos por pares por la siguiente razón:

Si existe $x \in A_iM \cap A_jM$, entonces existen $a_i \in A_i, a_j \in A_j$, $m_i, m_j \in M$ únicos tales que $a_im_i = x = a_jm_j$. En particular se tiene que $x = a_jm_j \in A_iM$, de donde $a_j \in A_i$ y por lo tanto $a_j \in A_i \cap A_j$, lo cual es absurdo.

Además,

$$\nu\left(\bigcup_{i=1}^{n} A_i\right) = \mu\left(\left(\bigcup_{i=1}^{n} A_i\right)M\right) = \mu\left(\bigcup_{i=1}^{n} A_iM\right) = \sum_{i=1}^{n} \mu(A_iM) = \sum_{i=1}^{n} \nu(A_i).$$

También se tiene que ν es normal con respecto a H, pues

$$\nu(H) = \mu(HM) = \mu\left(\bigcup_{x\in M} Hx\right) = \mu(G) = 1.$$

La H-invariancia es consecuencia de lo siguiente: Si $h \in H$ y $A \subseteq H$, entonces

$$\nu(hA) = \mu((hA)M) = \mu(h(AM)) = \mu(AM) = \nu(A),$$

pues $h \in G$ y μ es G-invariante. Concluimos que ν es H-invariante sobre H y por consiguiente H es promediable. □

Proposición 2.2.6. Sea G un grupo y H un subgrupo normal de G. Si H y G/H son promediables, entonces G también es promediable.

Demostración. Por hipótesis tenemos que existen μ y $\overline{\mu}$ promedios sobre H y G/H, H y G/H-invariantes, respectivamente. Nuestro objetivo es, a partir de los anteriores, construir un promedio G-invariante sobre G. Definimos para $A \subseteq G$

$$\mu_H(A) = \mu(A \cap H).$$

Notemos que μ_H es un promedio sobre G, pues dados $A_1, \ldots, A_n \subseteq G$ ajenos por pares, se tiene que

$$\mu_H\left(\bigcup_{i=1}^{n} A_i\right) = \mu\left(\left(\bigcup_{i=1}^{n} A_i\right) \cap H\right) = \mu\left(\bigcup_{i=1}^{n}(A_i \cap H)\right) = \sum_{i=1}^{n} \mu(A_i \cap H) = \sum_{i=1}^{n} \mu_H(A_i)$$

y es normal, pues

$$\mu_H(G) = \mu(G \cap H) = \mu(H) = 1.$$

Sin embargo, μ_H no tiene razón aparente para ser G-invariante, así que la modificaremos un poco para lograr nuestro objetivo.

Primero notemos que si tomamos $g \in G$, entonces $\mu_H(g^{-1}A)$ depende únicamente de la clase derecha de g módulo H. Sea $g_1 = gh$ con $h \in H$, entonces

$$\mu_H(g_1^{-1}A) = \mu_H((gh)^{-1}A) = \mu_H(h^{-1}g^{-1}A) = \mu(h^{-1}g^{-1}A \cap H) = \mu(g^{-1}A \cap H) = \mu_H(g^{-1}A).$$

Lo anterior nos permite definir para $\overline{g} \in G/H$

$$\mu_H^*(\overline{g}^{-1}A) = \mu_H(g^{-1}A).$$

Para obtener una medida G-invariante sobre G, promediaremos $\mu_H^*(\overline{g}^{-1}A)$ con $\overline{g} \in G/H$. Veamos que

$$\nu(A) = \int_{G/H} \mu_H^*(x^{-1}A)d\overline{\mu}(x)$$

es el promedio G-invariante que buscamos.

Veamos que ν es finitamente aditiva. Sean $A_1, \dots, A_n \subseteq G$ ajenos por pares. Entonces

$$\nu\left(\bigcup_{i=1}^{n} A_i\right) = \int_{G/H} \mu_H^*\left(x^{-1}\bigcup_{i=1}^{n} A_i\right) d\overline{\mu}(x) = \int_{G/H} \mu_H\left(g_x^{-1}\bigcup_{i=1}^{n} A_i\right) d\overline{\mu}(x)$$

$$= \int_{G/H} \sum_{i=1}^{n} \mu_H(g_x^{-1}A_i)d\overline{\mu}(x) = \sum_{i=1}^{n} \int_{G/H} \mu_H^*(x^{-1}A_i)d\overline{\mu}(x) = \sum_{i=1}^{n} \nu(A_i),$$

con $g_x^{-1} \in x^{-1}$.

Para no perder consistencia veamos que todos los elementos de la colección $(g_x^{-1}A_i)_{i=1}^n$ son ajenos por pares.
Si existieran $1 \leq i, j \leq n$ tales que $g_x^{-1}A_i \cap g_x^{-1}A_j \neq \varnothing$, entonces existen $z \in g_x^{-1}A_i \cap g_x^{-1}A_j$ y $(a_i, a_j) \in A_i \times A_j$ tales que

$$g_x^{-1}a_i = z = g_x^{-1}a_j,$$

pero eso significaría que $a_i = a_j$, lo cual es absurdo ya que $A_i \cap A_j = \varnothing$.

Además,

$$\nu(G) = \int_{G/H} \mu_H^*(x^{-1}G)d\overline{\mu}(x) = \int_{G/H} \mu_H(G)d\overline{\mu}(x)$$
$$= \int_{G/H} \mu(H)d\overline{\mu}(x) = \int_{G/H} d\overline{\mu}(x) = 1.$$

Finalmente, si tomamos $A \subseteq G$, $g, g_0 \in G$, $x = \overline{g}$ y $y = g_0 x$, entonces

$$\mu_H^*(y^{-1}g_0A) = \mu_H((g_0g)^{-1}g_0A) = \mu_H(g^{-1}g_0^{-1}g_0A) = \mu_H(g^{-1}A) = \mu_H^*(x^{-1}A).$$

Por lo tanto

$$\nu(g_0A) = \int_{G/H} \mu_H^*(y^{-1}g_0A)d\overline{\mu}(x) = \int_{G/H} \mu_H^*(x^{-1}A)d\overline{\mu}(x) = \nu(A).$$

Concluimos que ν es el promedio G-invariante sobre G que buscabamos.

$\square$

Antes de probar el siguiente teorema de grupos promediables, necesitaremos introducir dos lemas auxiliares que nos serán de gran ayuda.

Lema 2.2.7. *[Teorema del punto fijo de Kakutani]. Sea K un subconjunto no vacío, convexo y compacto de un espacio vectorial topológico X, y $T : K \to K$ una función continua y afín. Entonces T tiene un punto fijo.*

Demostración. Tomemos $a \in K$. Definimos para $n \geq 1$

$$u_n = \frac{1}{n}\sum_{k=0}^{n-1} T^k(a).$$

Como K es convexo, y $T : K \to K$, entonces $u_n \in K$. Más aún,

$$T(u_n) - u_n = \frac{1}{n}\left(T^n(a) - a\right). \tag{2.1}$$

Con $T^n(a) - a \in K - K$, siendo $K - K$ un subconjunto compacto de X, pues la suma de compactos es compacta.
Como $(T(u_n) - u_n)_{n\in\mathbb{N}}$ está contenida en un compacto, entonces tiene una subsucesión convergente, de manera que la correspondiente subsucesión $\left(\frac{1}{n_k}(T^{n_k}(a) - a)\right)_{k\in\mathbb{N}}$ es convergente y converge a 0, pues $\lim_{k\to\infty} \dfrac{1}{n_k} = 0$. Por otro lado, si denotamos $v_k := u_{n_k}$, como K es compacto, entonces

$$\bigcap_{n\geq 0} \overline{\{v_k : k \geq n\}} \neq \varnothing,$$

lo que significa que existe $\ell \in \bigcap_{n\geq 0} \overline{\{v_k : k \geq n\}}$ y una subsucesión $(v_k^r)_{r\in\mathbb{N}}$ que converge a ℓ.
Como T es continua, entonces $(T(v_k^r))_{r\in\mathbb{N}}$ converge a $T(\ell)$. Luego, $\lim_{r\to\infty} T(v_k^r) - v_k^r = T(\ell) - \ell$, pero por la igualdad (2.1) se tiene también que $\lim_{k\to\infty} T(v_k^r) - v_k^r = 0$. Así, $T(\ell) - \ell = 0$. $\square$

Teorema 2.2.8. *Conservamos la notación del lema anterior. Sea $\mathscr{A}$ una colección de funciones continuas y afines de K en sí mismo, que conmutan por pares. Entonces los elementos de $\mathscr{A}$ tienen un punto fijo común.*

Demostración. Sean $T \in \mathscr{A}$ y F_T el conjunto de puntos fijos de T. Notemos que el lema anterior nos garantiza que $F_T \neq \varnothing$. Además F_T es convexo, pues $T : K \to K$ es de la forma $x \mapsto Ax + b$, donde A es una transformación lineal y $+b$ es una traslación, y si tomamos $x, y \in F_T$, entonces para $t \in [0, 1]$ se tiene que

$$\begin{aligned}
T((1-t)x + ty) &= A((1-t)x + ty) + b \\
&= (1-t)Ax + tAy + b \\
&= (1-t)(x-b) + t(y-b) + b \\
&= x - b - tx + tb + ty - tb + b \\
&= x - tx + ty \\
&= (1-t)x + ty.
\end{aligned}$$

Veamos que F_T es un subocnjunto cerrado de K. Sea $(x_n)_{n\in\mathbb{N}} \subseteq F_T$ una sucesión convergente. Como T es continua, entonces

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} T(x_n) = T\left(\lim_{n\to\infty} x_n\right).$$

Por lo tanto $\lim\limits_{n\to\infty} x_n \in F_T$. Además, F_T es compacto por ser un subconjunto cerrado del compacto K.
Como para cualesquiera $S, T \in \mathscr{A}, \quad S \circ T = T \circ S$, se tiene que F_T es estable bajo todos los elementos de $\mathscr{A}$, pues si $x \in F_T$, entonces

$$T(S(x)) = S(T(x)) = S(x).$$

Probaremos por inducción que la colección $(F_T)_{T\in\mathscr{A}}$ tiene la propiedad de intersección finita:

Sean $T_1, T_2 \in \mathscr{A}$ y $K_i := F_{T_i}$ con $i = 1, 2$. Gracias al lema anterior, sabemos que K_1 y K_2 son no vacíos. Además, ambos son convexos y compactos, por lo que si tomamos $T_1|_{K_2} : K_2 \to K_2$, el lema anterior nos dice que $T_1|_{K_2}$ tiene un punto fijo en K_2, es decir

$$K_1 \cap K_2 \neq \varnothing.$$

Si suponemos que cualquier subfamilia de $n-1$ elementos de $\mathscr{A}$ satisface que la intersección de los respectivos K_i es no vacía, podemos probar que cualquier familia de n elementos de $\mathscr{A}$ satisface la misma propiedad.

Tomemos $T_1, \ldots, T_n \in \mathscr{A}$, por la hipótesis inductiva tenemos que

$$\bigcap_{i=1}^{n-1} K_i \neq \varnothing.$$

Además, el conjunto anterior es compacto y convexo. Para facilitar la escritura denotaremos $W := \bigcap\limits_{i=1}^{n-1} K_i$. Tomando $T_n|_W : W \to W$ y aplicando el lema anterior se sigue que $T_n|_W$ tiene un punto fijo en W. Por consiguiente, $(F_T)_{T\in\mathscr{A}}$ tiene la propiedad de intersección finita. Como K es compacto, podemos concluir que

$$\bigcap_{T\in\mathscr{A}} F_T \neq \varnothing.$$

□

Ahora sí, vamos con la proposición de grupos promediables.

Proposición 2.2.9. Todo grupo abeliano es promediable.

Demostración. Sea $\mathscr{K}$ el conjunto de todas las funcionales lineales positivas φ sobre $\ell^\infty(G)$ que satisfacen $\varphi(1) = 1$. Recordemos que si $f \in \ell^\infty(G)$, tenemos que

$$|\varphi(f)| \leq \|f\|_\infty.$$

Por lo tanto, el conjunto $\mathscr{K}$ está contenido en la bola unitaria cerrada del espacio dual fuerte[3] $\ell^\infty(G)'$ de $\ell^\infty(G)$, además es cerrado en la topología débil-$\star$.

El Teorema de Banach-Alaoglu 1.3.4 nos garantiza que la bola unitaria cerrada es compacta y por lo tanto $\mathscr{K}$, equipado con la topología inducida por la topología débil-$\star$ sobre $\ell^\infty(G)'$, es compacto. Más aún, $\mathscr{K}$ es no vacío, ya que las funciones de evaluación están en él y es convexo, pues dados $\varphi, \eta \in \mathscr{K}, \quad f \in \ell^\infty(G)$ tal que $f \geq 0$ y $t \in [0,1]$, entonces

- $(1-t)\varphi(f) + t\eta(f) \geq 0$;
- $$\begin{aligned} ((1-t)\varphi + t\eta)(1) &= (1-t)\varphi(1) + t\eta(1) \\ &= \varphi(1) - t\varphi(1) + t\eta(1) \\ &= 1 - t + t \\ &= 1. \end{aligned}$$

Ahora tomamos un $g \in G$ fijo. Para $\varphi \in \mathscr{K}$ y $f \in \ell^\infty(G)$ definimos la funcional

$$\begin{aligned} T_g\varphi : \ell^\infty(G) &\to \mathbb{R} \\ f &\mapsto \varphi(f_g), \end{aligned}$$

donde

$$\begin{aligned} f_g : G &\to G \\ h &\mapsto f(gh). \end{aligned}$$

Notemos que $T_g\varphi \in \mathscr{K}$, pues si $f \in \ell^\infty(G)$ es tal que $f \geq 0$, entonces $T_g\varphi(f) = \varphi(f_g) \geq 0$. Además, $T_g\varphi(1) = \varphi(1_g) = \varphi(1) = 1$.

Por lo anterior podemos definir la función

$$\begin{aligned} T_g : \mathscr{K} &\to \mathscr{K} \\ \varphi &\mapsto T_g\varphi. \end{aligned}$$

Tomando en cuenta el Teorema 2.1.4 y la Proposición 2.1.5 resulta que probar la existencia de un promedio G-invariante sobre G se traduce a demostrar que todos los elementos de la familia $(T_g)_{g\in G}$ tienen un punto fijo en común, pues si encontramos $\varphi \in \mathscr{K}$ tal que $T_g\varphi = \varphi \quad \forall g \in G$, y por lo tanto

$$\begin{aligned} T_g\varphi(f) &= \varphi(f) \quad \forall g \in G \quad \forall f \in \ell^\infty(G) \\ \varphi(f_g) &= \varphi(f) \quad \forall g \in G \quad \forall f \in \ell^\infty(G) \\ \varphi(f(gh)) &= \varphi(f(h)) \quad \forall g \in G \quad \forall f \in \ell^\infty(G) \quad \forall h \in G. \end{aligned}$$

[3]Dado un espacio vectorial topológico X, definimos a X', el *espacio dual fuerte de* X, como el espacio dual de X equipado con la topología fuerte, la cual es la topología de la convergencia uniforme sobre el álgebra de Banach $\ell^\infty(X)$.

Así, φ es G-invariante y por consiguiente μ_φ también lo es.

Cada T_g es lineal y por lo tanto afín, además es continua, ya que por definición de topología débil-$\star$, para $f \in \ell^\infty(G)$ fijo la función

$$\begin{aligned} \mathscr{K} &\to \mathbb{R} \\ \varphi &\mapsto T_g(\varphi)(f) = \varphi({}_g f) \end{aligned}$$

es continua.

Finalmente notemos que como G es abeliano, entonces las T_g conmutan por pares, pues si tomamos $g, h \in G$ y $\varphi \in \mathscr{K}$, entonces

$$T_g(T_h(\varphi))(f) = T_g(T_h\varphi(f)) = T_g(\varphi(f_h)) = \varphi(f_{hg}) = \varphi(f_{gh}) = T_h(\varphi(f_g)) = T_h(T_g(\varphi))(f).$$

Del lema anterior se sigue que si F_{T_g} es el conjunto de puntos fijos de T_g, entonces

$$\bigcap_{g \in G} F_{T_g} \neq \varnothing.$$

Se concluye el resultado deseado. $\square$

Definición 2.2.10. Decimos que un grupo G es soluble si existe una cadena finita

$$\{1\} = G_0 \subseteq G_1 \subseteq \cdots \subseteq G_{p-1} \subseteq G_p = G$$

de subgrupos G_i de G, donde G_i es subgrupo normal de G_{i+1} y tal que los cocientes G_{i+1}/G_i son abelianos.

Corolario 2.2.11. *Cualquier grupo soluble es promediable.*

Demostración. Es consecuencia de las Proposiciones 2.2.6 y 2.2.9. $\square$

2.3. El problema de la completación de la medida de Lebesgue

En esta sección atacaremos el problema de la existencia de promedios invariantes bajo grupos de isometrías (o subgrupos de estos) definidos sobre $\mathscr{P}(\mathbb{R}^d)$. Veremos que para dimensiones 1 y 2, la solución a este problema se sigue de la teoría estudiada hasta el momento.

Teorema 2.3.1. *Sea $d \in \{1, 2\}$. El grupo $Is(\mathbb{R}^d)$ de las isometrías afines de $\mathbb{R}^d$ es soluble y por consiguiente promediable. Por lo tanto, existe un promedio $Is(\mathbb{R}^d)$-invariante definido sobre $\mathscr{P}(\mathbb{R}^d)$. De manera más general: dado $d \geq 1$ arbitrario, si G es un subgrupo promediable de $Is(\mathbb{R}^d)$, entonces existe un promedio G-invariante sobre $\mathbb{R}^d$.*

Demostración. Consideremos la cadena

$$\{id_{\mathbb{R}^d}\} \subseteq \mathscr{T}_d \subseteq Is^+(\mathbb{R}^d) \subseteq Is(\mathbb{R}^d), \tag{2.2}$$

donde $\mathscr{T}_d$ denota el grupo de traslaciones de $\mathbb{R}^d$ e $Is^+(\mathbb{R}^d)$ el grupo de isometrías directas.

Recordemos que cada isometría se puede representar de manera única como una transformación ortogonal seguida de una traslación, es decir, para cada $f \in Is(\mathbb{R}^d)$ existen $\rho \in O(\mathbb{R}^d)$ y $\tau \in \mathscr{T}_d$ únicos, tales que

$$f = \tau\rho.$$

Considerando la proyección

$$\begin{aligned} \pi : Is(\mathbb{R}^d) &\to O(\mathbb{R}^d) \\ f = \tau\rho &\mapsto \rho \end{aligned}$$

del primer teorema de ismorfismo se sigue que

$$\ker(\pi) = \mathscr{T}_d \trianglelefteq Is(\mathbb{R}^d) \tag{2.3}$$

y

$$Is(\mathbb{R}^d)/\mathscr{T}_d \cong O(\mathbb{R}^d). \tag{2.4}$$

Análogamente se tiene,

$$\mathscr{T}_d \trianglelefteq Is^+(\mathbb{R}^d) \tag{2.5}$$

y

$$Is^+(\mathbb{R}^d)/\mathscr{T}_d \cong SO(\mathbb{R}^d). \tag{2.6}$$

Notemos también que como el determinante es un homomorfismo, del primer teorema de isomorfismo se sigue que

$$SO(\mathbb{R}^d) \trianglelefteq O(\mathbb{R}^d) \tag{2.7}$$

y

$$O(\mathbb{R}^d)/SO(\mathbb{R}^d) \cong \mathbb{Z}_2. \tag{2.8}$$

Análogamente,

$$Is^+(\mathbb{R}^d) \trianglelefteq Is(\mathbb{R}^d). \tag{2.9}$$

Por el tercer teorema de isomorfismo, de (2.3), (2.4), (2.6), (2.8) y (2.9) se sigue que

$$(Is(\mathbb{R}^d)/\mathscr{T}_d)/(Is^+(\mathbb{R}^d)/\mathscr{T}_d) \cong Is(\mathbb{R}^d)/Is^+(\mathbb{R}^d) \cong O(\mathbb{R}^d)/SO(\mathbb{R}^d) \cong \mathbb{Z}_2.$$

Como $SO(\mathbb{R}^d)$ (para $d \in \{1, 2\}$) y $\mathbb{Z}_2$ son abelianos, de (2.2), (2.5) y (2.9) se sigue que $\mathrm{Is}(\mathbb{R}^d)$ es soluble, y por lo tanto promediable, es decir, existe un promedio μ $\mathrm{Is}(\mathbb{R}^d)$-invariante sobre $\mathrm{Is}(\mathbb{R}^d)$.

Consideremos un vector fijo $e \in \mathbb{R}^d$ y la función

$$\begin{aligned} f : Is(\mathbb{R}^d) &\to \mathbb{R}^d \\ u &\mapsto u(e). \end{aligned}$$

Aplicando la Proposición 2.2.3 se tiene que

$$\mu_f(X) = \mu(f^{-1}(X)) \quad \text{para } X \subseteq \mathbb{R}^d$$

es un promedio $\mathrm{Is}(\mathbb{R}^d)$-invariante sobre $\mathbb{R}^d$.

El resultado anterior también se satisface para un subgrupo promediable G del grupo $\mathrm{Is}(\mathbb{R}^d)$ (con $d \geq 1$) aplicando el mismo razonamiento a la función $f|_G$. $\square$

Las medidas anteriores no tienen relación alguna con la medida de Lebesgue, pues tienen masa total 1. Esto nos lleva a otra pregunta: *¿es posible extender la medida de Lebesgue a una medida finitamente aditiva, invariante bajo isometrías, y definida sobre todos los subconjuntos de $\mathbb{R}^d$?*

El estudio realizado en la sección anterior nos muestra que esto es equivalente a estudiar el problema de extensión de funcionales lineales invariantes bajo la acción de un grupo.

La herramienta adecuada para atacar este problema es el teorema de Hahn-Banach, el cual enunciaremos en una versión G-invariante, pero antes de llegar a eso necesitaremos definir algunos conceptos.

Definición 2.3.2. Dado un grupo G que actúa linealmente sobre un espacio lineal real V, diremos que un subespacio lineal V_0 de V es G-invariante si

$$gx \in V_0 \quad \forall (g, x) \in G \times V_0.$$

Observación 2.3.3. Sea G un grupo y V un espacio lineal. Una *acción* de G sobre V es un homomorfismo de grupos $G \to \mathscr{G}(V)$, donde $\mathscr{G}(V)$ es el grupo de permutaciones de V.

Teorema 2.3.4. *[Hahn-Banach G-invariante] Sea G un grupo promediable que actúa linealmente sobre un $\mathbb{R}$-espacio lineal V, φ_0 una funcional lineal G-invariante definida sobre un subespacio lineal V_0 de V G-invariante y $p : V \to \mathbb{R}$ una función sublineal que satisface*

- *$p(gx) = p(x)$ para $(g, x) \in G \times X$;*
- *$\varphi_0(x) \leq p(x)$ para $x \in V_0$.*

Entonces existe una funcional lineal $\varphi : V \to \mathbb{R}$ G-invariante que extiende a φ_0 y tal que

$$\varphi(x) \leq p(x) \quad \forall x \in V.$$

Demostración. Sea μ un promedio G-invariante. El teorema de Hahn-Banach nos garantiza la existencia de una extensión $\Phi : V \to \mathbb{R}$ de φ_0 tal que

$$\Phi(x) \leq p(x) \quad x \in V.$$

Sin embargo, no hay razón aparente para que Φ sea G-invariante. Para arreglar este inconveniente, promediaremos los trasladados de Φ sobre G. De manera más precisa, sea $v \in V$ y notemos que para $h \in G$ tenemos

$$({}_h\Phi)(v) = \Phi(h^{-1}v) \leq p(h^{-1}v) = p(v)$$

y
$$-({}_h\Phi)(v) = -\Phi(h^{-1}v) = \Phi(h^{-1}(-v)) \leq p(h^{-1}(-v)) = p(-v),$$
pues la acción de G sobre V es lineal.
Por lo tanto, la función
$$h \mapsto ({}_h\Phi)(v)$$
es un elemento de $\ell^\infty(G)$, lo cual nos permite definir la funcional lineal sobre V:
$$\varphi(v) = \int_G ({}_h\Phi)(v)d\mu(h) = \int_G \Phi(h^{-1}v)d\mu(h).$$
Veamos que en efecto, φ extiende a φ_0: Si $v \in V_0$ tenemos que
$$\varphi(v) = \int_G \Phi(h^{-1}v)d\mu(h) = \int_G \varphi_0(h^{-1}v)d\mu(h) = \int_G \varphi_0(v)d\mu(h) = \varphi_0(v),$$
pues V_0 y φ_0 son G-invariantes.

Más aún, para $g \in G$ y $v \in V$ se tiene que
$$\begin{aligned}
{}_g\varphi(v) &= \varphi(g^{-1}v) \\
&= \int_G \Phi(h^{-1}g^{-1}v)d\mu(h) \\
&= \int_G \Phi((gh)^{-1}v)d\mu(h) \\
&= \int_G \Phi(h^{-1}v)d\mu(h) \qquad \text{(pues } \mu \text{ es } G\text{-invariante)} \\
&= \varphi(v).
\end{aligned}$$
Por lo tanto φ es G-invariante.

Finalmente, para $v \in V$ se tiene
$$\varphi(v) = \int_G \Phi(h^{-1}v)d\mu(h) \leq \int_G p(h^{-1}v)d\mu(h) = \int_G p(v)d\mu(h) = p(v) \quad \forall v \in V.$$
□

Teorema 2.3.5. *Sea G un subgrupo promediable del grupo de isometrías de $\mathbb{R}^d$. Entonces la medida de Lebesgue admite una extensión finitamente aditiva y G-invariante al conjunto de todos los subconjuntos de $\mathbb{R}^d$.*

Demostración. Sean λ la medida de Lebesgue, $V_0 := L^1(\mathbb{R}^d)$ el espacio de las funciones $f : \mathbb{R}^d \to \mathbb{R}$ Lebesgue integrables y V el espacio de las funciones $f : \mathbb{R}^d \to \mathbb{R}$ tales que existe $u \in V_0$ con $|f| \leq u$. Nótese que V_0 es un subespacio vectorial de V. Definamos
$$\varphi_0(f) = \int_{\mathbb{R}^d} f(x)d\lambda(x) \quad \text{para } f \in V_0$$

y

$$p(f) = \inf_{u \in V_0} \left\{ \int_{\mathbb{R}^d} u(x) d\lambda(x) : f \leq u \right\} \quad \text{para } f \in V.$$

Tomando $g \in G$ y $f \in V$, definimos la acción de G sobre V como

$$\begin{aligned} gf := ({}_g f) : \mathbb{R}^d &\to \mathbb{R} \\ x &\mapsto f(g^{-1}x). \end{aligned}$$

Veamos que dicha acción es lineal:
Para $t \in \mathbb{R}$, $f \in V$ y $g \in G$ tenemos que

$${}_g(tf)(x) = (tf)(g^{-1}x) = t(f(g^{-1}(x))) = t({}_g f(x)),$$

por lo tanto $g(tf) = t(gf)$.

Por otro lado, para $f, h \in V$ y $g \in G$ tenemos que

$${}_g(f+h)(x) = (f+h)(g^{-1}x) = f(g^{-1}x) + h(g^{-1}x) = {}_g f(x) + {}_g h(x),$$

por lo tanto $g(f+h) = gf + gh$.

Verificamos que las hipótesis del Teorema 2.3.4 se satisfacen.

1. φ_0 es una funcional lineal G-invariante. Sean $f \in V_0$ y $g \in G$,

$$\varphi_0(f) = \int_{\mathbb{R}^d} f(x) d\lambda(x) = \int_{\mathbb{R}^d} f(g^{-1}x) d\lambda(x) = \int_{\mathbb{R}^d} {}_g f(x) d\lambda(x) = \varphi_0({}_g f).$$

2. $L^1(\mathbb{R}^d)$ es un espacio lineal G-invariante.
   Si $f \in L^1(\mathbb{R}^d)$ y $g \in G$, entonces

$$\int_{\mathbb{R}^d} |gf(x)| d\lambda(x) = \int_{\mathbb{R}^d} |f(g^{-1}x)| d\lambda(x) = \int_{\mathbb{R}^d} |f(x)| d\lambda(x) < \infty.$$

3. *a*) Dados $f, f_1, f_2 \in V, \quad t \in \mathbb{R}^+$ y $g \in G$,

$$\begin{aligned} p(f_1 + f_2) &= \inf_{u \in V_0} \left\{ \int_{\mathbb{R}^d} u(x) d\lambda(x) : f_1 + f_2 \leq u \right\} \\ &\leq \inf_{u \in V_0} \left\{ \int_{\mathbb{R}^d} u(x) d\lambda(x) : f_1 \leq u \right\} + \inf_{u \in V_0} \left\{ \int_{\mathbb{R}^d} u(x) d\lambda(x) : f_2 \leq u \right\} \\ &= p(f_1) + p(f_2), \end{aligned}$$

$$p(tf) = \inf_{u \in V_0} \left\{ \int_{\mathbb{R}^d} u(x) d\lambda(x) : tf \leq u \right\} = t \inf_{u \in V_0} \left\{ \int_{\mathbb{R}^d} u(x) d\lambda(x) : f \leq u \right\} = tp(f).$$

b) Como
$$\int_{\mathbb{R}^d} f(g^{-1}x)d\lambda(x) = \int_{\mathbb{R}^d} f(x)d\lambda(x),$$
entonces
$$\inf_{u\in V_0}\left\{\int_{\mathbb{R}^d} u(x)d\lambda(x) : f(g^{-1}x) \leq u(x)\right\} = \inf_{u\in V_0}\left\{\int_{\mathbb{R}^d} u(x)d\lambda(x) : f(x) \leq u(x)\right\},$$
y así
$$p(gf) = p(f).$$

c) Finalmente notemos que si tomamos $f \in V$ fijo y $u \in V_0$ es tal que $f(x) \leq u(x)$ para todo $x \in \mathbb{R}^d$, entonces
$$\varphi_0(f) = \int_{\mathbb{R}^d} f(x)d\lambda(x) \leq \int_{\mathbb{R}^d} u(x)d\lambda(x).$$
Tomando el ínfimo sobre el conjunto $\{u \in V_0 : f \leq u\}$ se sigue que
$$\varphi_0(f) \leq p(f).$$

Del Teorema 2.3.4 se sigue que la funcional lineal φ_0 admite una extensión lineal φ a V G-invariante y tal que
$$\varphi(f) \leq p(f) \quad \forall f \in V.$$

Ahora, para $A \subseteq \mathbb{R}^d$ definimos
$$\mu(A) = \begin{cases} \varphi(\mathbb{1}_A) & \text{si} \quad \mathbb{1}_A \in V \\ +\infty & \text{en otro caso.} \end{cases}$$

Notemos que μ es una extensión de la medida de Lebesgue.
Más aún, si tomamos $A_1, \ldots, A_n \subseteq \mathbb{R}^d$ ajenos por pares y denotamos $A := \bigcup_{k=1}^{n} A_k$, entonces
$$\mathbb{1}_A \in V \text{ si y sólo si } \mathbb{1}_{A_k} \in V \text{ para todo } 1 \leq k \leq n,$$
pues
$$\mathbb{1}_A = \sum_{k=1}^{n} \mathbb{1}_{A_k} \text{ y } \mathbb{1}_{A_k} \leq \mathbb{1}_A \text{ para todo } 1 \leq k \leq n$$
y por lo tanto
$$\mu\left(\bigcup_{k=1}^{n} A_k\right) = \mu(A) = \int_{\mathbb{R}^d} \mathbb{1}_A d\lambda = \int_{\mathbb{R}^d} \sum_{k=1}^{n} \mathbb{1}_{A_k} d\lambda = \sum_{k=1}^{n} \int_{\mathbb{R}^d} \mathbb{1}_{A_k} d\lambda = \sum_{k=1}^{n} \mu(A_k),$$
lo que prueba que μ es finitamente aditiva.

Para probar que μ es positiva, tomamos $A \subseteq \mathbb{R}^d$ tal que $\mathbb{1}_A \in V$. Como $-\mathbb{1}_A \leq 0$ y $0 \in V$, tenemos entonces que

$$-\mu(A) = -\varphi(\mathbb{1}_A) = \varphi(-\mathbb{1}_A) \leq p(-\mathbb{1}_A) \leq p(0) = \int_{\mathbb{R}^d} 0 d\lambda(x) = 0,$$

se sigue que $\mu(A) \geq 0$.

Ahora veamos que μ es G-invariante. Dado $g \in G$ y $A \subseteq \mathbb{R}^d$ con $\mathbb{1}_A \in V$,

$$\mu(gA) = \varphi(\mathbb{1}_{gA}) = \varphi({}_g\mathbb{1}_A) = \varphi(\mathbb{1}_A) = \mu(A),$$

pues φ es G-invariante y ${}_g\mathbb{1}_A = \mathbb{1}_{gA}$.

Así, μ es la extensión de la medida de Lebesgue que estabamos buscando. □

Corolario 2.3.6. *Para $d \in \{1, 2\}$, la medida de Lebesgue sobre $\mathbb{R}^d$ admite una extensión a todos los subconjuntos de $\mathbb{R}^d$ que es finitamente aditiva e invariante bajo isometrías.*

Demostración. Ya probamos que para $d \in \{1, 2\}$, el grupo $\mathrm{Is}(\mathbb{R}^d)$ es promediable. El resultado se sigue del Teorema 2.3.5. □

Corolario 2.3.7. *Sea $d \geq 1$. Si G es un subgrupo promediable de $Is(\mathbb{R}^d)$, la medida de Lebesgue sobre $\mathbb{R}^d$ admite una extensión a todos los subconjuntos de $\mathbb{R}^d$ que es finitamente aditiva y G-invariante. Este es el caso, por ejemplo, si G es el grupo de traslaciones sobre $\mathbb{R}^d$.*

Demostración. El grupo de traslaciones de $\mathbb{R}^d$ es abeliano y por lo tanto promediable. Aplicando el Teorema 2.3.5 obtenemos el resultado deseado. □

Capítulo 3

Paradojas

En 1918, Bertrand Russell hizo la siguiente observación [9]:

> El punto de la filosofía es empezar con algo tan simple que no parezca valer la pena siquiera empezar, y terminar con algo tan paradójico que nadie pueda creer.

Aunque las paradojas no son el objeto de estudio de las matemáticas, no cabe duda que el llegar a comprender estos resultados contraintuitivos nos ayuda a fortalecer nuestro entendimiento y nos motiva a realizar un estudio más profundo. En este capítulo explicaremos dos construcciones paradójicas. La primera de ellas data del año 1914 y se le atribuye a Hausdorff, es comúnmente conocida como *la duplicación de la esfera.* La segunda es el tema central de esta tesis, *la paradoja de Banach-Tarski.*

Estas paradojas nos permiten responder la pregunta de la existencia de una extensión de la medida de Lebesgue sobre todos los subconjuntos de $\mathbb{R}^d$ cuando $d \geq 3$, que sea finitamente aditiva e invariante bajo isometrías.

A partir de este punto, G denotará un grupo (la mayoría de las veces lo podremos pensar como el grupo de isometrías de $\mathbb{R}^d$) que actúa sobre un conjunto no vacío E.

Definición 3.0.1. Decimos que $A, B \subseteq E$ son G-*congruentes* si existe $g \in G$ tal que $B = gA = \{ga : a \in A\}$ y lo escribimos como

$$A \equiv_G B.$$

Proposición 3.0.2. La G-congruencia es una relación de equivalencia.

Demostración.

1. Sea $A \subseteq E$. Entonces $A \equiv_G A$, pues $A = eA$, donde e es el elemento neutro de G.

2. Sean $A, B \subseteq E$ tales que $A \equiv_G B$. Entonces existe $g \in G$ tal que $B = gA$, luego $A = g^{-1}B$ y así $B \equiv_G A$.

3. Sean $A, B, C \subseteq E$ tales que $A \equiv_G B$ y $B \equiv_G C$, entonces existen $g, h \in G$ tales que

$$B = gA \quad \text{y} \quad C = hB.$$

Luego $C = hgA$, con $hg \in G$. Por lo tanto $A \equiv_G C$. Concluimos que la G-congruencia es una relación de equivalencia.

□

Definición 3.0.3. Dados $A, B \subseteq E$ decimos que son *equidescomponibles* si son *congruentes por pedazos*, en otras palabras, si existen descomposiciones finitas

$$A = \bigsqcup_{i=1}^{n} A_i \quad \text{y} \quad B = \bigsqcup_{i=1}^{n} B_i$$

tales que $A_i, B_i \subseteq E$ y $A_i \equiv_G B_i$ para toda $1 \leq i \leq n$. En este caso escribimos

$$A \sim_G B.$$

Proposición 3.0.4. La equidescomponibilidad es una relación de equivalencia.

Demostración. La reflexividad y la simetría se siguen del hecho de que la G-congruencia es relación de equivalencia. Para la transitividad tenemos el siguiente argumento:

Sean $A, B, C \subseteq E$ tales que $A \sim_G B$ y $B \sim_G C$.
Entonces existen $X_1, \ldots, X_n \subseteq$ y $g_1, \ldots, g_n \in G$ tales que

$$A = \bigsqcup_{i=1}^{n} X_i \quad \text{y} \quad B = \bigsqcup_{i=1}^{n} g_i X_i.$$

Análogamente, existen $Y_1, \ldots, Y_m \subseteq E$ y $h_1, \ldots h_m \in G$ tales que

$$B = \bigsqcup_{j=1}^{m} h_j Y_j \quad \text{y} \quad C = \bigsqcup_{j=1}^{m} Y_j.$$

Definimos $Z_{ij} = g_i X_i \cap h_j Y_j$ con $1 \leq i \leq n$ y $1 \leq j \leq m$.

Notemos que

$$\begin{aligned}
\bigcup_{i=1}^{n} \bigcup_{j=1}^{m} g_i^{-1} Z_{ij} &= \bigcup_{i=1}^{n} \bigcup_{j=1}^{m} g_i^{-1}(g_i X_i \cap h_j Y_j) \\
&= \bigcup_{i=1}^{n} \bigcup_{j=1}^{m} X_i \cap g_i^{-1} h_j Y_j \\
&= \bigcup_{i=1}^{n} X_i \cap g_i^{-1} B \\
&= \bigcup_{i=1}^{n} X_i \cap \left(g_i^{-1} \bigsqcup_{i=1}^{n} g_i X_i \right) \\
&= \bigsqcup_{i=1}^{n} X_i \\
&= A.
\end{aligned}$$

Por otro lado,

$$\begin{aligned}
\bigcup_{j=1}^{m}\bigcup_{i=1}^{n} h_j^{-1}Z_{ij} &= \bigcup_{j=1}^{m}\bigcup_{i=1}^{n} h_j^{-1}(g_iX_i \cap h_jY_j)\\
&= \bigcup_{j=1}^{m}\bigcup_{i=1}^{n} h_j^{-1}(g_iX_i \cap Y_j)\\
&= \bigcup_{j=1}^{m} h_j^{-1}B \cap Y_j\\
&= \bigcup_{j=1}^{m} h_j^{-1}\left(\bigsqcup_{j=1}^{m} h_jY_j\right) \cap Y_j\\
&= \bigsqcup_{j=1}^{m} Y_j\\
&= C.
\end{aligned}$$

Además, si $(i,j),(k,l) \in [1,n]\times[1,m]$ son distintos (sin pérdida de generalidad, supongamos que $i \neq k$), entonces

$$Z_{ij} \cap Z_{kl} = (g_iX_i \cap h_jY_j) \cap (g_kX_k \cap h_lY_l) \subseteq g_iX_i \cap g_kX_k = \varnothing$$

y $(g_i^{-1}h_j)h_j^{-1}Z_{ij} = g_i^{-1}Z_{ij}$.
Se concluye que $A \sim_G C$ y por lo tanto $\sim_G$ es relación de equivalencia. □

Ejemplo 3.0.5. *El grupo $\mathbb{S}^1$ de los números complejos con módulo 1 actúa sobre sí mismo mediante traslaciones izquierdas. Veamos que*

$$\mathbb{S}^1\backslash\{1\} \sim_{\mathbb{S}^1} \mathbb{S}^1.$$

Sea $z \in \mathbb{S}^1$ un elemento de orden infinito y consideremos el conjunto $D = \{z^n : n \geq 0\}$. Notemos que $zD = D\backslash\{1\}$.

Luego,

$$\mathbb{S}^1 = (\mathbb{S}^1\backslash D) \sqcup D \sim_{\mathbb{S}^1} (\mathbb{S}^1\backslash D) \sqcup zD = (\mathbb{S}^1\backslash D) \sqcup (D\backslash\{1\}) = \mathbb{S}^1\backslash\{1\}.$$

Se sigue el resultado deseado.

Para explicar la paradoja de Banach-Tarski, utilizaremos la idea del hotel de Hilbert, donde *recorremos a los huéspedes para hacer espacio.*

Definición 3.0.6. Un subconjunto $A \subseteq E$ se dice *G-paradójico* si existen $A_1, A_2 \subseteq A$ ajenos tales que

$$A \sim_G A_1 \quad \text{y} \quad A \sim_G A_2.$$

La definición anterior se puede interpretar como que A contiene dos copias disjuntas de si mismo.

Las siguientes observaciones son cruciales para la teoría a estudiarse.

Observación 3.0.7. Si existe una medida μ finitamente aditiva, G-invariante, y $A, B \subseteq E$ son tales que $A \sim_G B$, entonces $\mu(A) = \mu(B)$.

Demostración. Como $A \sim_G B$, entonces existen $A_1, \ldots, A_n \subseteq E$ y $g_1, \ldots, g_n \in G$ tales que

$$A = \bigsqcup_{i=1}^{n} A_i \quad \text{y} \quad B = \bigsqcup_{i=1}^{n} g_i A_i.$$

Luego,

$$\mu(B) = \mu\left(\bigsqcup_{i=1}^{n} g_i A_i\right) = \sum_{i=1}^{n} \mu(g_i A_i) = \sum_{i=1}^{n} \mu(A_i) = \mu\left(\bigsqcup_{i=1}^{n} A_i\right) = \mu(A).$$

□

Observación 3.0.8. Si X es un subconjunto G-paradójico de E, no existe una medida μ finitamente aditiva, G-invariante y tal que $\mu(X) = 1$.

Demostración. Supongamos por el contrario que existe dicha medida μ. Como X es G-paradójico existen $X_1, X_2 \subseteq X$ ajenos tales que $X \sim_G X_1$ y $X \sim_G X_2$. Como $X_1, X_2 \subseteq X$, entonces $X_1 \cup X_2 \subseteq X$ y así, $\mu(X_1 \cup X_2) \leq \mu(X)$, lo cual es absurdo, pues $\mu(X_1 \cup X_2) = \mu(X_1) + \mu(X_2) = 2$. Por lo tanto, la existencia de dicha medida μ es imposible. □

Un resultado sorprendente, pero difícil de probar es que el recíproco de la observación anterior es verdadero. Se le atribuye a Alfred Tarski y muestra, de cierta manera, que la existencia de medidas invariantes y la ausencia de paradojas es consustancial.

Teorema 3.0.9. *[La alternativa de Tarski] Sea G un grupo que actúa sobre un conjunto E mediante traslaciones izquierdas, y X un subconjunto de E. Entonces se satisface alguna de las siguientes afirmaciones:*

- *X es paradójico.*
- *Existe una medida finitamente aditiva y G-invariante μ sobre E tal que $\mu(X) = 1$.*

La parte difícil de probar de este teorema es muy larga, esta es la existencia de una medida finitamente aditiva, G-invariante μ sobre E tal que $\mu(X) = 1$ cuando X no es G-paradójico, pero se puede consultar en [3] (p.7-14).

La existencia de una medida finitamente aditiva y G-invariante sobre E no excluye por completo la existencia de subconjuntos paradójicos de E. Simplemente, la medida de un subconjunto paradójico será cero o infinito.

Para ilustrar esta afirmación veamos el siguiente ejemplo:

Ejemplo 3.0.10. *[La paradoja de Sierpinski-Mazurkiewicz] Dado $u \in \mathbb{C}$ un número trascendente con $|u| = 1$. Definimos*

$$\mathbb{N}[u] = \{p(u) : p \in \mathbb{N}[x]\}\,.$$

Entonces $\mathbb{N}[u]$ es paradójico bajo la acción del grupo de isometrías directas. De manera más precisa, $\mathbb{N}[u]$ es paradójico bajo la acción del grupo generado por la rotación

$$r : z \mapsto uz$$

y la traslación

$$\tau : z \mapsto z + 1.$$

Demostración. Consideremos los subconjuntos de $\mathbb{N}[u]$ dados por

$$A = \{p(u) : p \in \mathbb{N}[u], \quad p(0) = 0\}$$

y

$$B = \{p(u) : p \in \mathbb{N}[u], \quad p(0) \geq 1\}$$

y denotemos por G al subgrupo de isometrías directas $\mathbb{C} \to \mathbb{C}$ generado por r y τ.

Notemos que $A \cap B = \varnothing$, además

$$r^{-1}A = \mathbb{N}[u],$$

pues dado $p \in \mathbb{N}[u]$,

$$\begin{aligned} p(u) &= a_0 + a_1 u + \cdots + a_n u^n \\ &= u^{-1}(a_0 u + a_1 u^2 + \cdots + a_n u^{n+1}) \\ &= u^{-1} q(u) \qquad \text{con } q \in A \\ &= r^{-1} q(u) \in r^{-1}A. \end{aligned}$$

Similarmente se tiene que

$$\tau^{-1}B = \mathbb{N}[u],$$

pues dado $p \in \mathbb{N}[u]$

$$\begin{aligned} p(u) &= a_0 + a_1 u + \cdots + a_n u^n \\ &= ((a_0 + 1) + a_1 u + \cdots + a_n u^n) - 1 \\ &= q(u) - 1 \qquad \text{con } q \in B \\ &= \tau^{-1} q(u). \end{aligned}$$

Así, es posible obtener $\mathbb{N}[u]$ a partir de A y B mediante una rotación o una traslación, respectivamente. Es decir

$$A \sim_G \mathbb{N}[u] \quad \text{y} \quad B \sim_G \mathbb{N}[u].$$

Como A y B son ajenos, se concluye que $\mathbb{N}[u]$ es G-paradójico. $\square$

3.1. La paradoja de la esfera

Denotamos por $SO(3)$ al grupo de las rotaciones lineales de $\mathbb{R}^3$, es decir, el conjunto de isometrías lineales con determinante 1 del espacio euclidiano $\mathbb{R}^3$. Denotamos a la identidad de $\mathbb{R}^3$ por 1.

Teorema 3.1.1. *[Hausdorff, 1914] La esfera unitaria de $\mathbb{R}^3$ es $SO(3)$-paradójica.*

Siguiendo el método de Von Neumann y Klein, la prueba consistirá en la construcción de un subconjunto paradójico del grupo $SO(3)$ que posteriormente transportaremos a la esfera.

Antes necesitamos introducir algunas nociones básicas de grupos libres.

Definición 3.1.2. Sea G un grupo.

- Dados $a, b \in G$, diremos que el conjunto

$$\mathscr{A} = \{a, a^{-1}, b, b^{-1}\}$$

  es un *alfabeto*.

- Dado $n \geq 1$, llamaremos una *palabra reducida de longitud n construida sobre el alfabeto $\mathscr{A}$* a cada elemento de G de la forma $a_1a_2 \cdots a_n$ con $a_i \in \mathscr{A}$ tal que $a_ia_{i+1} \neq 1$ para cada $1 \leq i \leq n-1$ (en otras palabras, no hay una simplificación evidente de la palabra). Por convención, la unidad de G es la palabra reducida de longitud cero, también llamada *la palabra vacía.*

- Diremos que el subgrupo de G generado por a y b es *libre de rango 2* si se satisfacen las siguientes condiciones.

  1. El alfabeto $\mathscr{A}$ tiene cuatro elementos distintos;
  2. Cada palabra reducida de longitud $n \geq 1$ construida sobre $\mathscr{A}$ es distinta de 1.

  Dicho de otra manera, lo anterior significa que no existe una relación no trivial entre a y b, y eso implica que el subgrupo $\langle a, b\rangle$ es *lo menos conmutativo posible*, en el sentido de que dos palabras reducidas distintas se consideran como elementos distintos del grupo.

Proposición 3.1.3. Un grupo libre de rango 2 es paradójico.

Demostración. Sea G el grupo libre de rango 2 en cuestión, con generadores a y b, para cualquier $g \in \{a, a^{-1}b, b^{-1}\} = \mathscr{A}$, denotamos por $I(g)$ al conjunto de todas las palabras reducidas que empiezan con la letra g.
Como G es grupo libre tenemos que

- Si $g, h \in \mathscr{A}$ son distintos, entonces $I(g) \cap I(h) = \varnothing$;

- $$\begin{aligned} G = \langle a, b\rangle &= \{1\} \sqcup I(a) \sqcup I(a^{-1}) \sqcup I(b) \sqcup I(b^{-1}) \\ &= I(a) \sqcup aI(a^{-1}) \\ &= I(b) \sqcup bI(b^{-1}) \end{aligned}$$

  pues dado $h \in G \backslash I(a)$, $a^{-1}h \in I(a^{-1})$ y $h = aa^{-1}h \in aI(a^{-1})$. Análogamente para b.

Finalmente, como

$$I(a) \sqcup aI(a^{-1}) = a\{1\} \sqcup aI(a) \sqcup aI(b) \sqcup aI(b^{-1}) \sqcup aI(a^{-1}),$$

entonces

$$I(a) \sqcup I(a^{-1}) \sim_G G.$$

Análogamente,

$$I(b) \sqcup I(b^{-1}) \sim_G G.$$

Se concluye que G es paradójico. □

Observación 3.1.4. Es posible probar que cuando $SO(3)$ está equipado con su topología usual, el conjunto de parejas $(a, b) \in SO(3)^2$ que generan un subgrupo libre es residual[1] y por lo tanto denso en $SO(3)$. Ver [2].

Con base en lo anterior, podemos enunciar el siguiente resultado.

Teorema 3.1.5. *El grupo $SO(3)$ contiene un grupo libre de rango 2 y por lo tanto es paradójico.*

Demostración. Es posible dar una construcción directa de dicho subgrupo (ver [2]), pero conlleva mucho cálculos complicados. El método que daremos aquí, aunque indirecto, es más sencillo, se construirá un subgrupo de $SO(3)$ que no es libre, pero dentro del cual se puede encontrar fácilmente el grupo buscado.

La idea original de Hausdorff consiste en usar dos rotaciones de ángulos $\frac{2\pi}{3}$ y π, escogiendo el ángulo θ definido entre sus ejes de rotación de tal manera que el número $\cos(2\theta)$ sea trascendente.

Sin embargo, nosotros seguiremos un ejemplo más reciente y sencillo dado por Osofsky y Adams (ver [4]) en el cual se usa el ángulo $\theta = \frac{\pi}{4}$. Consideremos la base canónica de $\mathbb{R}^3$, $\{e_1, e_2, e_3\}$, y las rotaciones con ángulos π y $\frac{2\pi}{3}$ sobre los ejes con direcciones $\frac{e_1+e_3}{\sqrt{2}}$ y e_3, respectivamente con matrices:

[1] Sea X un espacio topológico.

- Diremos que X es *denso en ninguna parte* si el interior de su cerradura es el conjunto vacío.
- Diremos que $E \subseteq X$ es *magro* si se puede escribir como unión numerable de conjuntos densos en ninguna parte.
- Diremos que $A \subseteq X$ es *residual* si $X \backslash A$ es magro.

$$\varphi = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{y} \quad \psi = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Nótese que $\varphi^2 = \psi^3 = 1$, lo cual nos muestra que el subgrupo $\langle \varphi, \psi \rangle$ de $SO(3)$ no es libre.

Sin embargo, veremos que $\langle \varphi, \psi \rangle$ es *casi libre*, en el sentido de que φ y ψ satisfacen dentro de $\langle \varphi, \psi \rangle$ únicamente las relaciones que satisfacen por separado, es decir, $\varphi^2 = \psi^3 = 1$.

De manera más precisa, podemos dar la siguiente proposición:

Proposición 3.1.6. Si $n \geq 1$, $\varepsilon \in \{0,1\}, \varepsilon' \in \{-1,0,1\}$ y $\varepsilon_i \in \{-1,1\}$, $i \in \{1,\ldots,n\}$. Entonces

$$\varphi^{\varepsilon}\psi^{\varepsilon_1}\varphi\psi^{\varepsilon_2}\varphi\cdots\psi^{\varepsilon_n}\varphi\psi^{\varepsilon'} \neq 1. \tag{3.1}$$

Demostración. Sea w una palabra como (3.1). Veamos que $w \neq 1$. Conjugando por ψ o ψ^{-1}, de ser necesario, podemos suponer que w termina con φ. Por lo tanto, tendremos disponibles dos tipos de palabra para w:

$$w = \psi^{\varepsilon_1}\varphi\psi^{\varepsilon_2}\varphi\cdots\psi^{\varepsilon_n}\varphi \tag{3.2}$$

y

$$w = \varphi\psi^{\varepsilon_1}\varphi\psi^{\varepsilon_2}\varphi\cdots\psi^{\varepsilon_n}\varphi \tag{3.3}$$

con $\varepsilon_1, \ldots, \varepsilon_n \in \{-1,1\}$.

Empezamos con las palabras de la forma (3.2).
Primero notemos que

- si $\varepsilon = 1$, entonces

$$\psi^{\varepsilon}\varphi = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 1 & 0 & 0 \end{pmatrix};$$

- si $\varepsilon = -1$, entonces

$$\psi^{\varepsilon}\varphi = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 1 & 0 & 0 \end{pmatrix}.$$

Por lo que para $\varepsilon \in \{-1,1\}$ se tiene que

$$\psi^{\varepsilon}\varphi = \begin{pmatrix} 0 & \varepsilon\frac{\sqrt{3}}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & \varepsilon\frac{\sqrt{3}}{2} \\ 1 & 0 & 0 \end{pmatrix} \tag{3.4}$$

Ahora probemos mediante un argumento inductivo que si $n \geq 1$, entonces

$$\psi^{\varepsilon_1}\varphi\psi^{\varepsilon_2}\varphi\cdots\psi^{\varepsilon_n}\varphi = \frac{1}{2^n}\begin{pmatrix} p_1 & i_1\sqrt{3} & i_2 \\ p_2\sqrt{3} & i_3 & i_4\sqrt{3} \\ p_3 & p_4\sqrt{3} & p_5 \end{pmatrix}, \tag{3.5}$$

donde i_1, i_2, i_3, i_4 son impares y p_1, p_2, p_3, p_4, p_5 son pares. En (3.4) se ve que el resultado se satisface para $n = 1$. Si suponemos que el resultado es cierto para $n - 1 \geq 1$, entonces a partir de (3.4) y (3.5) podemos ver que

$$\begin{aligned} w = \psi^{\varepsilon_1}\varphi\psi^{\varepsilon_2}\varphi\cdots\psi^{\varepsilon_n}\varphi &= \frac{1}{2^{n-1}}\begin{pmatrix} p_1 & i_1\sqrt{3} & i_2 \\ p_2\sqrt{3} & i_3 & i_4\sqrt{3} \\ p_3 & p_4\sqrt{3} & p_5 \end{pmatrix}\begin{pmatrix} 0 & \varepsilon_n\frac{\sqrt{3}}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & \varepsilon_n\frac{\sqrt{3}}{2} \\ 1 & 0 & 0 \end{pmatrix} \\ &= \frac{1}{2^n}\begin{pmatrix} 2i_2 & (\varepsilon_n p_1 + i_1)\sqrt{3} & -p_1 + 3\varepsilon_n i_1 \\ 2i_4\sqrt{3} & 3\varepsilon_n p_2 + i_3 & -p_2 + \varepsilon_n i_3\sqrt{3} \\ 2p_5 & (\varepsilon_n p_3 + p_4)\sqrt{3} & -p_3 + 3\varepsilon_n p_4 \end{pmatrix} \end{aligned}$$

lo cual prueba el paso inductivo.

De (3.5) se sigue que cuando w es del tipo (3.2), entonces no es la identidad y también es distinta de φ. De manera que, a partir de (3.5) también podemos deducir que si w es del tipo (3.3), entonces w es distinta de la identidad, lo cual prueba que en ambos casos

$$\varphi^{\varepsilon}\psi^{\varepsilon_1}\varphi\psi^{\varepsilon_2}\varphi\cdots\psi^{\varepsilon_n}\varphi\psi^{\varepsilon'} \neq 1.$$

□

Para terminar la prueba del Teorema 3.1.5, definimos

$$a = \psi\varphi\psi \quad \text{y} \quad b = \varphi\psi\varphi\psi\varphi.$$

Por la Proposición 3.1.6 se tiene que $\langle a, b\rangle$ es un grupo libre, pues todos los elementos del alfabeto $\mathscr{A} = \{a, a^{-1}, b, b^{-1}\}$ son distintos, y cada palabra construida a partir de $\mathscr{A}$ es distinta de 1 por ser de la forma (3.1). □

Ahora necesitamos llevar a la esfera unitaria de $\mathbb{R}^3$ la paradoja que acabamos de mostrar en $SO(3)$. Para esto usamos el siguiente lema:

Definición 3.1.7. Decimos que un grupo G actúa libremente sobre un conjunto E si no tiene puntos fijos, es decir, si $gx \neq x \quad \forall g \in G\backslash\{1\}\ \forall x \in E$.

Lema 3.1.8. *Sea G un grupo paradójico que actúa libremente sobre un conjunto E. Entonces E es G-paradójico.*

Demostración. Una acción de G sobre E puede asociarse con la relación de equivalencia orbital sobre E, definida por

$$x \sim y \Leftrightarrow \exists g \in G \text{ tal que } y = gx.$$

Sea $M \subseteq E$ conformado exactamente por un representante de cada clase de equivalencia. Como G actúa libremente sobre E, si $A, B \subseteq G$ son ajenos, entonces $AM \cap BM = \varnothing$.

Como G es paradójico, entonces existen $H, K \subseteq G$ ajenos tales que

$$H \sim_G G \quad \text{y} \quad K \sim_G G,$$

es decir, existen $G_1, \ldots, G_n \subseteq G$ y $g_1, \ldots, g_n \in G$ tales que

$$G = \bigsqcup_{i=1}^{n} G_i \quad \text{y} \quad H = \bigsqcup_{i=1}^{n} g_i G_i.$$

Sea $A_i = G_i M$, entonces

$$E = GM = \bigsqcup_{i=1}^{n} G_i M = \bigsqcup_{i=1}^{n} A_i$$

y

$$HM = \bigsqcup_{i=1}^{n} g_i G_i M = \bigsqcup_{i=1}^{n} g_i A_i.$$

Así, $E \sim_G HM$. Análogamente $E \sim_G KM$, y como $HM \cap KM = \varnothing$, entonces se tiene que E es G-paradójico. □

Desafortunadamente, el grupo $\langle a, b \rangle$ dado al final de la prueba del Teorema 3.1.5 no actúa libremente sobre $\mathbb{S}^2$, pues cada elemento distinto de 1 tiene exactamente dos puntos fijos en $\mathbb{S}^2$. Así que, de momento, nos conformaremos con un resultado más débil.

Lema 3.1.9. *Todo elemento de $SO(3)$ distinto de la identidad fija 2 puntos de $\mathbb{S}^2$.*

Proposición 3.1.10. Existe $D \subseteq \mathbb{S}^2$ numerable, tal que $\mathbb{S}^2 \backslash D$ es $SO(3)$-paradójico.

Demostración. Sabemos que existe $G \leqslant SO(3)$ subgrupo libre de rango 2. Naturalmente, G es numerable. Denotaremos por D al conjunto de puntos fijos en $\mathbb{S}^2$ por algún elemento de $G \backslash \{1\}$. D es numerable por el Lema 3.1.9 y por lo tanto $\mathbb{S}^2 \backslash D \neq \varnothing$. Veamos que el grupo G actúa sobre D:

Si tomamos $g \in G \backslash \{1\}$ y $x \in D$ es un punto fijo de $h \in G$, entonces $gx \in \mathbb{S}^2$ es un punto fijo de la rotación $ghg^{-1} \in G$, pues

$$ghg^{-1}(gx) = ghx = gx.$$

De manera similar vemos que G preserva a $\mathbb{S}^2 \backslash D$, ya que resulta imposible la existencia de $x \in \mathbb{S}^2 \backslash D$ y $g \in G$ tales que $gx \in D$ por un argumento similar al que acabamos de dar:

Si $gx \in D$, entonces existe $h \in G$ tal que $hgx = gx$ y agx es un punto fijo de la rotación aha^{-1} para todo $a \in G \backslash \{1\}$. En particular si tomamos $a = g^{-1}$ y así $agx = g^{-1}gx = x \in D$.

Por consiguiente, G actúa sobre $\mathbb{S}^2 \backslash D$ y esta acción es libre, pues le estamos quitando los puntos fijos. Por el Lema 3.1.8 se tiene que $\mathbb{S}^2 \backslash D$ es G-paradójico. Por lo tanto $\mathbb{S}^2 \backslash D$ es $SO(3)$-paradójico (pues G es subgrupo de $SO(3)$), que es lo que queríamos probar. □

Ahora sólo falta deshacernos del subconjunto D para mejorar este resultado, pero antes de seguir, será necesario dar un lema auxiliar.

Lema 3.1.11. *Si un grupo G actúa sobre un conjunto E, dos subconjuntos equidescomponibles de E son simultáneamente G-paradójicos.*

Demostración. Supongamos que $A, B \subseteq E$ son equidescomponibles y que A es G-paradójico, entonces existen $C, D \subseteq A$ ajenos y equidescomponibles a A.
Por hipótesis existen $A_1, \ldots, A_n \subseteq E$ y $g_1, \ldots, g_n \in G$ tales que

$$A = \bigsqcup_{i=1}^{n} A_i \quad \text{y} \quad B = \bigsqcup_{i=1}^{n} g_i A_i.$$

Definamos,

$$C' = \bigsqcup_{i=1}^{n} g_i(C \cap A_i) \quad \text{y} \quad D' = \bigsqcup_{i=1}^{n} g_i(D \cap A_i).$$

Por construcción, los conjuntos C' y D' son subconjuntos ajenos de B.
Además,

$$C' \sim_G \bigsqcup_{i=1}^{n} (C \cap A_i) = C \sim_G A \sim_G B.$$

Análogamente,

$$D' \sim_G \bigsqcup_{i=1}^{n} (D \cap A_i) = D \sim_G A \sim_G B.$$

Se sigue que B es G-paradójico. □

Proposición 3.1.12. El subconjunto $\mathbb{S}^2 \backslash D$ es $SO(3)$-equidescomponible a $\mathbb{S}^2$. Por lo tanto, $\mathbb{S}^2$ es $SO(3)$-paradójico.

Demostración. Tomemos $a \in \mathbb{S}^2 \backslash D$ fijo. Denotemos por E al conjunto de las rotaciones $r \in SO(3)$ con eje $\mathbb{R}a$ para las cuales existen $n \geq 1$ y $x, y \in D$ tales que $r^n(x) = y$. Veamos que E es numerable:
Sea $(x, y) \in D \times D$ tal que existe $r \in E$ que satisface la igualdad

$$r(x) = y.$$

Supongamos además que el ángulo de rotación de r es θ. Si queremos llegar de x a y en $n \geq 1$ pasos simplemente tomamos la rotacicón ρ de ángulo $\frac{\theta}{n}$ y así

$$\rho^n(x) = y.$$

Por lo tanto, para cada $n \geq 1$ hay una rotación que lleva x a y en n pasos y por consiguiente hay tantas rotaciones como naturales que satisfacen $r^n(x) = y$ para $n \geq 1$. Finalmente

$$|E| \leq |D \times D| \cdot |\mathbb{N}| = |\mathbb{N}| \cdot |\mathbb{N}| = |\mathbb{N}|.$$

Si $r \in SO(3) \backslash E$, entonces $r^n(D) \cap D = \varnothing$ para $n \geq 1$, pues para cualquier $x \in D$, si $r^n(x) \in D$ para algún $n \geq 1$, entonces $r \in E$. Veamos que los $r^n(D)$ con $n \in \mathbb{N}$ son ajenos por pares. Sean $n, m \geq 1$ con $n > m$ y $p = n - m$. Supongamos que existe $z \in r^n(D) \cap r^m(D)$, entonces existen $x, y \in D$ tales que $r^n(x) = z = r^m(y)$. Por lo tanto

$$\begin{aligned} r^{-m}(r^n(x)) &= r^{-m}(r^m(y)) \\ r^p(x) &= y, \end{aligned}$$

lo cual resulta absurdo, pues $r \notin E$.
Definimos $D' := \bigsqcup_{n \in \mathbb{N}} r^n(D)$. Note que $r(D') = D' \backslash D$ y $r(\mathbb{S}^2 \backslash D') = \mathbb{S}^2 \backslash D'$. Por lo que

$$\mathbb{S}^2 = (\mathbb{S}^2 \backslash D') \sqcup D' \sim_{SO(3)} (\mathbb{S}^2 \backslash D') \sqcup (D' \backslash D) = \mathbb{S}^2 \backslash D.$$

Finalmente, como $\mathbb{S}^2$ y $\mathbb{S}^2 \backslash D$ son equidescomponibles, y ya se probó que $\mathbb{S}^2 \backslash D$ es $SO(3)$-paradójico, entonces del Lema 3.1.11 se sigue que $\mathbb{S}^2$ es $SO(3)$-paradójico, que es lo que queríamos demostrar. □

Observación 3.1.13. El Lema 3.1.11 nos permite *iterar* la definición de la naturaleza paradójica. En efecto, si $A \subseteq E$ es G-paradójico, se sigue que $\forall n \geq 2$ existen $A_1, \ldots, A_n \subseteq E$ ajenos por pares, todos equidescomponibles a A.

Demostración. Como A es G-paradójico, existen $A_1, A_2 \subseteq A$ tales que $A_1 \sim_G A$ y $A_2 \sim_G A$. Por el Lema 3.1.11 se tiene que A_1 y A_2 son G-paradójicos. Entonces existen $A_3, A_4 \subseteq A_1$ ajenos y G-equidescomponibles a A_1, y por transitividad $A_3 \sim_G A$ y $A_4 \sim_G A$. Además, como $A_1 \cap A_2 = \varnothing$, entonces $A_3 \cap A_2 = \varnothing A_3 = \varnothing$ y $A_2 \cap A_4 = \varnothing$. Por lo tanto $\{A_2, A_3, A_4\}$ es una colección de 3 subconjuntos de A ajenos por pares y G-equidescomponibles a A. Repitiendo este proceso de manera inductiva, podemos obtener n subconjuntos de A ajenos por pares y equidescomponibles a A para cualquier $n \geq 2$. □

Proposición 3.1.14. La bola unitaria cerrada B en la topología euclidiana de $\mathbb{R}^3$ es $Is^+(\mathbb{R}^3)$-paradójica.

Demostración. Primero probaremos que $B \backslash \{0\}$ es $SO(3)$-paradójica. Como $\mathbb{S}^2$ es $SO(3)$-paradójica, existen $C, D \subseteq \mathbb{S}^2$ subconjuntos ajenos $SO(3)$-equidescomponibles a $\mathbb{S}^2$. Entonces existen $A_1, \ldots A_n \subseteq \mathbb{S}^2$ y $g_1, \ldots, g_n \in SO(3)$ tales que

$$\mathbb{S}^2 = \bigsqcup_{i=1}^{n} A_i \quad \text{y} \quad C = \bigsqcup_{i=1}^{n} g_i A_i.$$

Para cada subconjunto P de $\mathbb{S}^2$ definimos el conjunto

$$P^* = \{tx : x \in P, \quad 0 < t \leq 1\}.$$

Note que si P y Q son dos subconjuntos ajenos de $\mathbb{S}^2$, entonces los respectivos P^* y Q^* son subconjuntos ajenos de $B \backslash \{0\}$.

Además, como la acción de G es lineal se tiene que

$$C^* = \bigsqcup_{i=1}^{n} g_i A_i^* \sim_{SO(3)} \bigsqcup_{i=1}^{n} A_i^* = (\mathbb{S}^2)^* = B\backslash\{0\}^2.$$

Análogamente,

$$D^* \sim_{SO(3)} B\backslash\{0\},$$

con $C^* \cap D^* = \varnothing$. Por lo tanto, $B\backslash\{0\}$ es $SO(3)$-paradójico y en consecuencia $B\backslash\{0\}$ es $Is^+(\mathbb{R}^3)$-paradójico.

Ahora probaremos que B es $Is^+(\mathbb{R}^3)$-paradójico.
Sea $r \in Is^+(\mathbb{R}^3)$ una rotación afin de ángulo θ, escogido de tal manera que $\frac{\theta}{\pi} \neq \mathbb{Q}$ y cuyo eje contiene al punto $(0, 0, \frac{1}{12})$, pero no al origen. Tomando el conjunto

$$D = \{r^n(0) : n \geq 0\} \subseteq B$$

tenemos que $r(D) = \{r^n(0) : n \geq 1\} = D\backslash\{0\}$.
Además,

$$B = (B\backslash D) \sqcup D \sim_{Is^+(\mathbb{R}^3)} (B\backslash D) \sqcup r(D) = (B\backslash D) \sqcup (D\backslash\{0\}) = B\backslash\{0\}.$$

Así, $B\backslash\{0\} \sim_{Is^+(\mathbb{R}^3)} B$, y como $B\backslash\{0\}$ es $Is^+(\mathbb{R}^3)$-paradójico, del Lema 3.1.11 se sigue que B es $Is^+(\mathbb{R}^3)$-paradójico. □

3.2. La paradoja de Banach-Tarski en $\mathbb{R}^3$

La idea de cortar una figura y reacomodar los pedazos para formar otra figura es al menos tan antigua como la geometría clásica de los griegos, donde usaron este método para calcular áreas de regiones tales como los paralelogramos. Como ya se probó en la sección anterior, la equidescomponibilidad es una relación de equivalencia. Al estudiar las propiedades de esta relación, en 1924, Banach y Tarski lograron mejorar la paradoja de Hausdorff al eliminar la necesidad de excluir un subconjunto numerable de la esfera.

Teorema 3.2.1. *[Banach-Tarski] Dos subconjuntos acotados de $\mathbb{R}^3$ con interior no vacío son $Is^+(\mathbb{R}^3)$-equidescomponibles.*

El enunciado anterior es paradójico en el sentido de que pareciera contradecir la conservación del volumen bajo isometrías: por ejemplo, dos bolas con radio distinto o incluso una o n copias de esta bola son equidescomponibles.

En realidad, esto último no representa un problema, pues al menos una de las piezas de la descomposición no es Lebesgue medible.

---

[2] $(gA)^* = gA^*$ y $\left(\bigsqcup_{i=1}^{n} A_i\right)^* = \bigsqcup_{i=1}^{n} A_i^*$.

La prueba consiste en usar la naturaleza $Is^+(\mathbb{R}^3)$-paradójica de $\mathbb{S}^2$ para replicarla, así como un ingrediente extra: una adaptación del teorema de Cantor-Bernstein que veremos a continuación.

Teorema 3.2.2. *[Cantor-Bernstein] Sean E y F dos conjuntos tales que existen funciones inyectivas $f : E \to F$ y $g : F \to E$. Entonces existe una biyección entre E y F.*

Demostración. Consideremos la aplicación

$$\begin{aligned} \Phi : \mathscr{P}(E) &\to \mathscr{P}(E) \\ A &\mapsto E \backslash g(F \backslash f(A)). \end{aligned}$$

Veamos que Φ tiene un punto fijo. Para ello consideremos la colección

$$\mathscr{U} = \{U \in \mathscr{P}(E) : U \subseteq \Phi(U)\}$$

y la unión

$$W := \bigcup_{U \in \mathscr{U}} U \subseteq \bigcup_{U \in \mathscr{U}} \Phi(U) \subseteq \Phi(W).$$

Queremos probar ahora que $\Phi(W) \subseteq W$.
Notemos que como $W \subseteq \Phi(W)$, entonces $\Phi(W) \subseteq \Phi(\Phi(W))$, lo que significa que $\Phi(W) \in \mathscr{U}$ y así $\Phi(W) \subseteq W$, por definición de W. Por lo tanto, W es un punto fijo de Φ.

Consideremos la función $h : E \to F$ dada por

$$h(x) = \begin{cases} f(x) & \text{si} \quad x \in W \\ g^{-1}(x) & \text{si} \quad x \notin W \end{cases} .$$

Notemos que h está bien definida, pues g es inyectiva y si $x \notin W$, entonces $x \in g(F \backslash f(W))$.

Para probar que h es la biyección buscada primero veamos que es suprayectiva.
Sea $y \in F$, entonces $g(y) \in E$.

- Si $g(y) \notin W$, entonces $h(g(y)) = g^{-1}(g(y)) = y$.

- Si $g(y) \in W$, entonces $g(y) \in E \backslash g(F \backslash f(W))$, de donde $g(y) \notin g(F \backslash f(W))$. Luego $y \notin g^{-1}(g(F \backslash f(W))) = F \backslash f(W)$ y así $y \in f(W)$, lo que significa que existe $x \in W$ tal que $f(x) = h(x) = y$.

Se sigue que h es suprayectiva.

Ahora veamos que h es inyectiva.
Sean $x, y \in E$ distintos.

- Si $x, y \in W$, entonces $h(x) = f(x) \neq f(y) = h(y)$.
- Si $x, y \notin W$, entonces $h(x) = g^{-1}(x) \neq g^{-1}(y) = h(y)$.

- Si $x \in W$ y $y \notin W$, entonces $y \in g(F\backslash f(W))$, de manera que existe $a \in F\backslash f(W)$ tal que $g(a) = y$, por lo que $g^{-1}(y) = a$. Como $f(x) \in F(W)$ y $a \notin f(W)$, entonces $h(x) \neq h(y)$.

Concluimos que h es inyectiva y por ende, la biyección buscada. □

Ahora introducimos la siguiente notación:

Dado un grupo G que actúa sobre un conjunto E, y $A, B \subseteq E$ escribimos

$$A \preceq_G B$$

para indicar que A es equidescomponible a algún subconjunto de B. Luego definimos una relación binaria sobre $\mathscr{P}(E)$, la cual es reflexiva y transitiva. El teorema de Banach-Cantor-Bernstein afirma que, módulo equidescomponibilidad, ésta es una relación de orden.

Teorema 3.2.3. *[Banach-Cantor-Bernstein] Sea G un grupo que actúa sobre un conjunto E y $A, B \subseteq E$ tales que $A \preceq_G B$ y $B \preceq_G A$. Entonces $A \sim_G B$.*

Demostración. Por hipótesis, existen $A' \subseteq B$ y $B' \subseteq A$ tales que

$$A = \bigsqcup_{i=1}^{m} A_i, \quad A' = \bigsqcup_{i=1}^{m} g_i A_i,$$

$$B = \bigsqcup_{j=1}^{n} B_j, \quad B' = \bigsqcup_{j=1}^{n} h_j B_j.$$

Esto nos permite definir, por pedazos, las funciones inyectivas

$$\begin{aligned} f : A &\to B \\ x &\mapsto g_i x \quad \text{si } x \in A_i \end{aligned}$$

y

$$\begin{aligned} g : B &\to A \\ x &\mapsto h_j x \quad \text{si } x \in B_j. \end{aligned}$$

Estas funciones tienen la siguiente propiedad que es simple, pero muy poderosa:

Si $X \subseteq A$, entonces $X \sim_G f(X)$.

En efecto, bastará con expresar

$$X = X \cap A = X \cap \left(\bigsqcup_{i=1}^{m} A_i \right) = \bigsqcup_{i=1}^{m} X \cap A_i,$$

de manera que

$$f(X) = \bigsqcup_{i=1}^{m} g_i(X \cap A_i) \sim_G \bigsqcup_{i=1}^{m} (X \cap A_i) = X.$$

De manera análoga se prueba que si $X \subseteq B$, entonces

$$X \sim_G g(X).$$

Luego, por lo visto en la prueba del Teorema de Cantor-Bernstein 3.2.2 se tiene que existe un subconjunto $X \subseteq A$ tal que $A \backslash X = g(B \backslash f(X))$.
Por lo tanto

$$X \sim_G f(X) \quad \text{y} \quad A \backslash X \sim_G B \backslash f(X)$$

y finalmente

$$A = X \sqcup (A \backslash X) \sim_G f(X) \sqcup (B \backslash f(X)) = B.$$

□

Para finalizar con esta sección, procedemos a probar el Teorema de Banach-Tarski 3.2.1.

Demostración. Sean A y A' dos subconjuntos de $\mathbb{R}^3$, acotados y con interior no vacío. Entonces existen B y B' bolas cerradas tales que

$$B \subseteq A \quad \text{y} \quad A' \subseteq B' .$$

Escogemos $n \in \mathbb{N}$ suficientemente grande tal que existen $B_1, \ldots, B_n$ copias trasladadas de B que cubren a B'. Dichas copias no necesariamente son ajenas por pares.

Luego, existen $C_1, \ldots, C_n \subseteq B_1, \ldots, B_n$, respectivamente, tales que $C_i \cap C_j = \varnothing$ si $i \neq j$ y que satisfacen

$$A' = \bigsqcup_{i=1}^{n} C_i.$$

Como B es $Is^+(\mathbb{R}^3)$-paradójico (ver la Proposición 3.1.14), entonces por la Observación 3.1.13 existen subconjuntos $D_1, \ldots, D_n \subseteq B$ ajenos por pares, tales que todos son equidescomponibles a B.

Después tomamos $B'_1, \ldots, B'_n$ traslaciones de B ajenas por pares. Notemos que como $C_i \subseteq B_i$ y B_i es una traslación de B'_i (pues tanto B_i como B'_i son traslaciones de B), entonces existen $g_i \in \mathscr{T}_3$ y $C'_i \subseteq B'_i$ tales que $B_i = g_i B'_i$ y $C_i = g_i C'_i$. Por lo tanto

$$C_i \preceq_{Is^+(\mathbb{R}^3)} B_i \text{ para toda } 1 \leq i \leq n.$$

También es importante mencionar que $B'_i \sim_{Is^+(\mathbb{R}^3)} D_i$, ya que , por construcción, $D_i \sim_{Is^+(\mathbb{R}^3)} B$ y B'_i es una traslación de B.

Finalmente, notemos que

$$A' = \bigsqcup_{i=1}^{n} C_i \preceq_{Is^+(\mathbb{R}^3)} \bigsqcup_{i=1}^{n} B'_i \sim_{Is^+(\mathbb{R}^3)} \bigsqcup_{i=1}^{n} Di \subseteq B \subseteq A.$$

Por lo tanto,

$$A' \preceq_{Is^+(\mathbb{R}^3)} A.$$

Intercambiando los papeles de A y A', y repitiendo el argumento anterior, se llega a que

$$A \preceq_{Is^+(\mathbb{R}^3)} A'$$

y gracias al Teorema de Banach-Cantor-Bernstein 3.2.3 concluimos que

$$A' \sim_{Is^+(\mathbb{R}^3)} A.$$

$\square$

¿Cuáles son las consecuencias de la paradoja de Hausdorff-Banach-Tarski?

- No existe un promedio $Is^+(\mathbb{R}^3)$-invariante sobre $\mathbb{S}^2$ o sobre la bola cerrada unitaria de $\mathbb{R}^3$.

- No existe una medida μ sobre $\mathbb{R}^3$ finitamente aditiva, invariante bajo isometrías y tal que $\mu([0,1]^3) = 1$.
  En efecto, como $[0,1]^3$ y $[0,1]^3 \cup [2,3]^3$ son $Is^+(\mathbb{R}^3)$-equidescomponibles, el Teorema de Banach-Tarski 3.2.1 nos dice que si tal μ existiera, entonces

  $$2 = \mu([0,1]^3 \cup [2,3]^3) = \mu([0,1]^3) = 1,$$

  que es claramente una contradicción. En particular, la medida de Lebesgue no admite una extensión a todos los subconjuntos de $\mathbb{R}^3$ que sea finitamente aditiva e invariante bajo isometrías. El resultado anterior se puede extender a $\mathbb{R}^d$ para $d \geq 3$, aunque no es cierto en $\mathbb{R}$ y $\mathbb{R}^2$. Se pueden replicar bolas, pero no discos.

Apéndice A

Conjuntos no medibles

Al principio de esta tesis se habló un poco acerca de cómo Vitali descubrió que hay conjuntos que no son medibles en el sentido de Lebesgue. En este apéndice continuamos sobre esa línea de investigación, exponiendo un resultado que muestra que todo conjunto de medida positiva contiene un conjunto no medible.

A.1. El teorema de Steinhaus

Teorema A.1.1. *[Steinhaus] Sea λ la medida de Lebesgue en $\mathbb{R}^n$ y A un subconjunto medible tal que $\lambda(A) > 0$. Entonces $A - A$ contiene una vecindad del origen.*

Demostración. Veamos primero que si $f \in L^1(\mathbb{R}^n)$ y $g \in L^\infty(\mathbb{R}^n)$, entonces la convolución $f * g$ definida por

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dy$$

es una función continua sobre $\mathbb{R}^n$.

Notemos primero que para $x \in \mathbb{R}^n$

$$\int_{\mathbb{R}^n} |f(x-y)g(y)|dy \leq \int_{\mathbb{R}^n} |f(x-y)|dy \, \|g\|_\infty \leq \|f\|_1 \, \|g\|_\infty < +\infty.$$

Por lo tanto, $(f * g)(x)$ está definida para toda $x \in \mathbb{R}^n$, y además

$$\|f * g\|_\infty \leq \|f\|_1 \, \|g\|_\infty \, .$$

Recordemos que el espacio de funciones continuas con soporte compacto $C_c(\mathbb{R}^n)$ es denso en $L^1(\mathbb{R}^n)$. Por consiguiente, existe una sucesión $(f_k)_{k\in\mathbb{N}} \subseteq C_c(\mathbb{R}^n)$ convergente a f en $L^1(\mathbb{R}^n)$.

Del teorema de convergencia dominada de Lebesgue se sigue que la función

$$(f_k * g)(x) = \int_{\mathbb{R}^n} f_k(x-y)g(y)dy$$

es continua.

Finalmente,

$$\|(f_k * g) - (f * g)\|_\infty = \|(f_k - f) * g\|_\infty \leq \|f_k - f\|_1 \|g\|_\infty \xrightarrow[k\to\infty]{} 0.$$

Lo anterior significa que $(f_k * g)_{k\in\mathbb{N}}$ converge uniformemente a $f * g$, y por lo tanto $f * g$ es continua sobre $\mathbb{R}^n$.

Veamos que si A y B son subconjuntos medibles de $\mathbb{R}^n$, tales que $\lambda(A), \lambda(B) \in (0, +\infty)$, entonces $\mathbb{1}_A$ y $\mathbb{1}_B$ son funciones integrables y por lo tanto

$$\int_{\mathbb{R}^n} \mathbb{1}_A * \mathbb{1}_B d\lambda = \left(\int_{\mathbb{R}^n} \mathbb{1}_A d\lambda\right)\left(\int_{\mathbb{R}^n} \mathbb{1}_B d\lambda\right) = \lambda(A)\lambda(B) \in (0, +\infty).$$

De manera que existe $x \in \mathbb{R}^n$ tal que $(\mathbb{1}_A * \mathbb{1}_B)(x) > 0$.

Denotemos por P al conjunto de los $x \in \mathbb{R}^n$ tales que $(\mathbb{1}_A * \mathbb{1}_B)(x) \neq 0$. Veamos que $P \subseteq A + B$.

Sea $x \in \mathbb{R}^n \backslash (A + B)$. Notemos que $A \cap (-B + x) = \varnothing$, pues si existiera $y \in A \cap (-B + x)$, entonces por un lado se tendría que existe $a \in A$ tal que $y = a$, y por otro lado que existe $b \in B$ tal que $y = -b + x$. Se sigue que $x = a + b$, lo cual es absurdo.

Del análisis anterior podemos deducir que

$$\mathbb{R}^n \backslash (A + B) \subseteq \{x \in \mathbb{R}^n : (\mathbb{1}_A * \mathbb{1}_B)(x) = 0\},$$

y así

$$P \subseteq A + B.$$

Notese que esto significa que $A + B$ tiene interior no vacío, pues

$$P^+ := \{x \in \mathbb{R}^n : (\mathbb{1}_A * \mathbb{1}_B)(x) > 0\} \subseteq P,$$

y como $(\mathbb{1}_A * \mathbb{1}_B)$ es continua, entonces existe un abierto no vacío U tal que

$$U \subseteq P^+ \subseteq P \subseteq A + B. \tag{A.1}$$

Veamos ahora que el resultado se sigue cumpliendo si $\lambda(A) = +\infty$.

Para $k \in \mathbb{N}$ definimos $I_k = [-k, k]^n$.

Para probar que $A + B$ tiene interior no vacío bastará con tomar $A_0 = A \cap I_k$ para $k \in \mathbb{N}$ suficientemente grande tal que $\lambda(A_0) > 0$. Repitiendo el proceso anterior, veremos que es posible encontrar un abierto no vacío U_0 tal que

$$U_0 \subseteq A_0 + B \subseteq A + B.$$

En particular veremos que si se toma $B = -A$, entonces $A - A$ contiene una vecindad del origen, pues como la convolución $\mathbb{1}_A * \mathbb{1}_{-A}$ es continua, entonces

$$(\mathbb{1}_A * \mathbb{1}_{-A})(0) = \int_{\mathbb{R}^n} \mathbb{1}_A(y)\mathbb{1}_{-A}(y)dy = \int_{\mathbb{R}^n} (\mathbb{1}_A(-y))^2 dy = \lambda(A) > 0$$

implica que existe V vecindad del 0 tal que $\mathbb{1}_A * \mathbb{1}_{-A}$ es positiva en todo V, y por lo que ya probamos en (A.1), se concluye que

$$V \subseteq A - A.$$

□

A.2. Otro teorema tipo Vitali

Teorema A.2.1. *Sean λ la medida de Lebesgue en $\mathbb{R}$ y $A \subseteq \mathbb{R}$ tal que $\lambda(A) > 0$, entonces existe un subconjunto de A que no es medible.*

Demostración. Consideremos la relación de equivalencia sobre $\mathbb{R}$ definida a continuación:

$$x \sim y \Leftrightarrow x^{-1}y \in \mathbb{Q}.$$

El axioma de elección nos permite escoger un conjunto $M \subseteq \mathbb{R}$ que contiene exactamente un representante de cada clase de equivalencia.

Para cada $r \in \mathbb{Q}$ definimos el conjunto

$$A_r = A \cap (r + M).$$

Notemos lo siguiente:

$$\bigcup_{r\in\mathbb{Q}} A_r = \bigcup_{r\in\mathbb{Q}} A \cap (r + M) = A \cap \bigcup_{r\in\mathbb{Q}} (r + M) = A \cap \mathbb{R} = A. \tag{A.2}$$

Si $r, s \in \mathbb{Q}$ son distintos, entonces $A_r \cap A_s = \varnothing$, pues si existe $x \in A_r \cap A_s$, entonces existen $m_1, m_2 \in M$ tales que

$$m_1 + r = x = m_2 + s.$$

Luego $m_1 - m_2 = s - r \in \mathbb{Q}$, y entonces $m_1 \sim m_2$, pero por la manera en la que se eligió M se tendría que $m_1 = m_2$, y entonces $r = s$, lo cual es absurdo.

Supongamos ahora que todos los A_r son medibles. Por (A.2) se tiene que

$$\sum_{r\in\mathbb{Q}} \lambda(A_r) = \lambda\left(\bigcup_{r\in\mathbb{Q}} A_r\right) = \lambda(A) > 0. \tag{A.3}$$

Por lo tanto, necesariamente alguno de los A_r tiene medida estrictamente positiva.

Luego,

$$A_r - A_r = A \cap (r + M) - A \cap (r + M) \subseteq (r + M) - (r + M) = M - M. \tag{A.4}$$

Notemos también que

$$(M - M) \cap (\mathbb{Q} \backslash \{0\}) = \varnothing,$$

pues si tomamos $m_1 - m_2 \in (M - M) \cap \mathbb{Q}$, entonces $m_1 = m_2$, por la manera en la que se eligió M, y así $m_1 - m_2 = 0$.

Sea A_ℓ el conjunto de medida positiva, cuya existencia se mostró en (A.3). Entonces $A_\ell - A_\ell$ es medible y el teorema de Steinhaus A.1.1 nos garantiza la existencia de una vecindad del cero contenida en $A_\ell - A_\ell$, pero eso es imposible, pues $\mathbb{Q} \backslash \{0\}$ es denso en $\mathbb{R}$ y ya se vio en (A.4) que

$$(A_\ell - A_\ell) \cap (\mathbb{Q} \backslash \{0\}) = \varnothing.$$

La contradicción vino de suponer que todos los A_r son medibles, por lo que forzosamente entre ellos se encuentra un subconjunto de A que no es medible, que es lo que queríamos probar.

□

Apéndice B

Un teorema de Agnew y Morse

En este apéndice se dará una versión alternativa del Teorema de Hahn-Banach G-invariante 2.3.4.

Teorema B.0.1. *[Hahn-Banach invariante de acuerdo con Agnew-Morse] Sean X un espacio normado sobre $\mathbb{R}$, $\mathscr{A}$ una colección de funciones lineales y continuas de X en sí mismo. Supongamos que $p : X \to \mathbb{R}$ es una función positivamente homogénea*[1]*, subaditiva y $\mathscr{A}$-invariante.*[2]

Sea Y un subespacio lineal de X, estable bajo todos los elementos de $\mathscr{A}$, y $\ell : Y \to \mathbb{R}$ una funcional lineal $\mathscr{A}$-invariante y tal que

$$\ell(y) \leq p(y) \quad \forall y \in Y.$$

Entonces ℓ admite una extensión lineal $L : X \to \mathbb{R}$ tal que L es $\mathscr{A}$-invariante y

$$L(x) \leq p(x) \quad \forall x \in X.$$

Demostración. Sea $\mathscr{C}$ la envolvente convexa[3] del semigrupo generado por $\mathscr{A}$. Para $x \in X$ definimos

$$q(x) = \inf_{u \in \mathscr{C}} p(ux).$$

Veamos que para $x \in X$,

$$q(x) = \inf_{u \in \mathscr{C}} p(ux) \leq p\left(\sum_k a_k A_k x\right) \leq \sum_k p(a_k A_k x) = \sum_k a_k p(A_k x) = \sum_k a_k p(x) = p(x).$$

Así,

$$q(x) \leq p(x) \tag{B.1}$$

. Para $x, y \in X$ se tiene que

$$q(x+y) = \inf_{u \in \mathscr{C}} p(u(x+y))$$

[1]Esto es: $p(\lambda x) = \lambda p(x) \ \ \forall (x, \lambda) \in X \times \mathbb{R}^+$.

[2]Esto es: $p(Ax) = p(x) \ \ \forall (x, A) \in X \times \mathscr{A}$.

[3]Recordemos que ésta se define como $\mathscr{C} = \left\{ \sum_k a_k A_k : \sum_k a_k = 1, a_k > 0 \text{ y } A_k \in \mathscr{A} \right\}$.

$$\leq p\left(\sum_k a_kA_k(x+y)\right) = p\left(\sum_k a_kA_kx + \sum_k a_kA_ky\right) \leq p\left(\sum_k a_kA_kx\right) + p\left(\sum_k a_kA_ky\right).$$

Tomando ínfimo sobre $\mathscr{C}$ se sigue que

$$q(x+y) \leq q(x) + q(y). \tag{B.2}$$

Para $\lambda \in \mathbb{R}^+$ y $x \in X$,

$$q(\lambda x) = \inf_{u\in\mathscr{C}} p(u\lambda x) = \inf_{u\in\mathscr{C}} \lambda p(ux) = \lambda \inf_{u\in\mathscr{C}} p(ux) = \lambda q(x).$$

Por consiguiente,

$$q(\lambda x) \leq \lambda q(x). \tag{B.3}$$

Sea $u \in \mathscr{C}$ con $u = \sum_k a_kA_k$ y $y \in Y$. Notemos que

$$\ell(uy) = \ell\left(\sum_k a_kA_ky\right) = \sum_k a_k\ell(A_ky) = \sum a_k\ell(y) = \ell(y).$$

Luego, por la definición de ℓ se tiene que

$$\ell(uy) = \ell(y) \leq p(uy).$$

Tomando el ínfimo sobre $\mathscr{C}$ se sigue que

$$\ell(y) \leq \inf_{u\in\mathscr{C}} p(uy),$$

y así

$$\ell(y) \leq q(y) \quad \forall y \in Y.$$

Por (B.1),(B.2) y (B.3), del Teorema de Hahn-Banach 1.3.1 se sigue que ℓ admite una extensión lineal $L : X \to \mathbb{R}$ tal que

$$L(x) \leq q(x) \quad \forall x \in X.$$

Sólo nos falta probar que L es $\mathscr{A}$-invariante.

Para $n \geq 1$ y $A \in \mathscr{A}$ definimos $C_n^A \in \mathscr{C}$ como

$$C_n^A = \frac{1}{n}\sum_{k=0}^{n-1} A^k.$$

Observemos que

$$C_n^A(x - A) = \frac{1}{n}(x - A^n).$$

Por definición de q se tiene que

$$q(x - Ax) \leq p(C_n^A(x - Ax)) = p\left(\frac{1}{n}\sum_{k=0}^{n-1} A^k(x - Ax)\right) = \frac{1}{n}p(x - Ax) \xrightarrow[k\to\infty]{} 0.$$

Por lo tanto,

$$q(x - Ax) \leq 0 \quad \forall (x, A) \in X \times \mathscr{A}.$$

Finalmente notemos que para $x \in X$ y $A \in \mathscr{A}$,

$$L(x) - L(Ax) = L(x - Ax) \leq q(x - Ax) \leq 0.$$

Por consiguiente

$$L(x) \leq L(Ax).$$

Tomando $y = -x$ y repitiendo el proceso anterior se tiene que

$$L(y) \leq L(Ay),$$

entonces

$$-L(x) \leq -L(Ax),$$

y así

$$L(Ax) \leq L(x).$$

Concluimos que

$$L(Ax) = L(x) \quad \forall (x, A) \in X \times \mathscr{A}.$$

□

Bibliografía

[1] Choimet, Denis y Queffélec, Hervé, *Twelve landmarks of twentieth-century analysis.* Cambridge University Press, Cambridge, UK, 2015.

[2] Tomkowicz, Grzegorz y Wagon, Stan, *The Banach-Tarski Paradox, 2nd edition.* Cambridge University Press, 2016.

[3] Runde, Volker *Lectures on Ameanability.* Lecture Notes in Mathematics No. 1774, Springer, 2002.

[4] B. Osofsky y S. Adams, *Some rotations of* $\mathbb{R}^3$, solution of problem 6102, Amer. Math. Monthly 85, 6, 1978, pp. 504–505.

[5] S. Banach y A. Tarski, *Sur la décomposition des ensembles de points en parties respectivement congruentes*, Fund. Math. 6, 1924, pp. 244–277.

[6] Brezis, Haim, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, 2010.

[7] Salomon, Dietmar A., *Measure and Integration*, European Mathematical Society, Zürich, Switzerland, 2016.

[8] Vitali, Giuseppe, *Sul problema della misura dei gruppi di punti di una retta*, Bologna, 1905.

[9] Russell, Bertrand, *The Philosophy of Logical Atomism*, Routledge, 1918.