Universidad Nacional Autónoma de México

Facultad de Ciencias

LA PARADOJA DE BANACH-TARSKI

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
MATEMÁTICO

PRESENTA:
MANUEL GERARDO ZORRILLA NORIEGA

DIRECTOR DE TESIS:
DR. JOSÉ ALFREDO AMOR Y MONTAÑO

2008

Índice general

Introducción

En 1924, el alemán Stefan Banach y el polaco Alfred Tarski probaron un resultado sumamente contraintuitivo. Demostraron que dada una bola cerrada en $\mathbb{R}^3$, es posible partirla en una cantidad finita de pedazos y reacomodar estos pedazos, usando movimientos rígidos, para formar *dos* bolas cerradas, cada una del mismo tamaño que la original.

La demostración de este notable teorema, que ahora se conoce con el nombre de Paradoja de Banach-Tarski, emplea el Axioma de Elección. Por ello, se ha considerado como un argumento *en contra* de aceptar tal axioma.

La Paradoja de Banach-Tarski choca fuertemente con la intuición natural, porque los movimientos rígidos de $\mathbb{R}^3$ preservan el volumen. La paradoja se resuelve a la luz de que el Axioma de Elección implica la existencia de conjuntos no Lebesgue-medibles, y no tiene sentido esperar que el volumen de éstos se preserve, pues no tienen volumen.

Desde luego, el procedimiento descrito en el enunciado de este teorema no se puede llevar a cabo en el mundo físico, ya que los elementos de la partición indicada son conjuntos arbitrarios de puntos (y el concepto matemático de punto carece de realidad física).

En realidad, lo que la Paradoja de Banach-Tarski pone en evidencia es que, si se acepta el Axioma de Elección, no se puede definir el volumen de cualquier conjunto de puntos. Así, esta paradoja, como todas, representa una oportunidad de afinar la intuición y lograr con ello un avance epistémico.

El objetivo de este trabajo es hacer clara, para cualquier estudiante de pregrado que maneje bien el material de geometría analítica y de álgebra lineal, que haya llevado un curso elemental de álgebra abstracta y uno de teoría de conjuntos y que tenga nociones de análisis matemático, la demostración de la Paradoja de Banach-Tarski.

La estructura general de los cuatro capítulos principales sigue a [W], aunque los detalles de las demostraciones son originales. Las ideas generales de los apéndices F y H, y de la sección de comentarios, fueron tomadas

de [M-S], de una conferencia y de [W], respectivamente, pero el desarrollo preciso de estas secciones, así como el contenido de los otros apéndices, son aportaciones originales.

A lo largo de este trabajo, el lector observará que se tuvo el cuidado de decir exactamente cuáles resultados dependen del Axioma de Elección, especialmente cuando se habla de cardinalidades, a partir del Capítulo 3.

Resumen

En primer lugar, se definirán algunos conceptos fundamentales. Si un grupo G actúa sobre un conjunto X, decimos que $E \subseteq X$ es G-paradójico si y sólo si hay $A, B \subseteq E$ ajenos tales que A se puede partir en una cantidad finita de subconjuntos que se pueden transformar uno por uno con elementos de G para cubrir a E, y lo mismo sucede con B. En tal situación, en el Capítulo 2 veremos que A y B, y sus particiones, pueden ser tomados de manera que los subconjuntos de A, ya transformados, queden ajenos dos a dos, que pase lo mismo con B, y que además $A \cup B$ sea todo E. Si G es un grupo y pensamos en la traslación, que es una acción de G sobre G mismo, decimos que G es paradójico si y sólo si G es G-paradójico. El primer resultado que demostraremos es que todo grupo libre de rango 2 es paradójico. A continuación veremos que, aceptando el Axioma de Elección, si un grupo paradójico G actúa sobre un conjunto X, y la acción cumple una cierta condición, entonces X resulta ser G-paradójico. Tomando en cuenta que se prueba que SO_3 coincide con el grupo de rotaciones de $\mathbb{R}^3$ por el origen, pensemos ahora en la acción natural de SO_3 sobre $\mathbb{R}^3$. En el Capítulo 3, será demostrado, con ayuda de un cierto lema técnico, que SO_3 tiene un subgrupo libre de rango 2, digamos F, y que hay $D \subseteq \mathbb{S}^2$, D contable, tal que F actúa sobre $\mathbb{S}^2 \backslash D$ de la manera adecuada para que podamos asegurar que $\mathbb{S}^2 \backslash D$ es F-paradójico, y por tanto SO_3-paradójico. Eso se conoce con el nombre de Paradoja de Hausdorff. La contabilidad de D permitirá que entonces obtengamos que $\mathbb{S}^2$ es SO_3-paradójico. Sea ahora $\mathscr{B}$ una bola cerrada en $\mathbb{R}^3$ centrada en $\mathbf{0}$ (el origen). La linealidad de los elementos de SO_3 nos permitirá engrosar a $\mathbb{S}^2$ hasta obtener $\mathscr{B} \backslash \mathbf{0}$, de manera que sea evidente que $\mathscr{B} \backslash \mathbf{0}$ es también SO_3-paradójico. Si denotamos con G_3 al grupo de isometrías de $\mathbb{R}^3$, y consideramos la acción natural de G_3 sobre $\mathbb{R}^3$, será fácil llegar, partiendo de lo que ya tendremos, a que cualquier bola cerrada en $\mathbb{R}^3$ es un conjunto G_3-paradójico, es decir, la Paradoja de Banach-Tarski. Revisaremos por último la demostración de la

forma fuerte de este resultado, que afirma que si A y B son subconjuntos de $\mathbb{R}^3$ acotados y con interior no vacío, A se puede partir en una cantidad finita de subconjuntos que se pueden transformar con rotaciones y traslaciones para formar una partición de B. Se incluyen varios apéndices que explican conceptos básicos y otros en los que se dan diversos resultados relacionados, tales como la existencia de un conjunto numerable y paradójico en el plano, una consecuencia directa de la Paradoja de Banach-Tarski en el área de la teoría de la medida, y un acertijo cuya solución es otra consecuencia contraintuitiva del Axioma de Elección.

Capítulo 1

Descomposiciones Paradójicas

Se sabe, desde la antigüedad, que el concepto de infinito conduce rápidamente a construcciones que desafían la intuición. Algunas de estas construcciones parecen cambiar el tamaño de ciertos objetos mediante operaciones que cabría esperar que lo preservaran. Consideremos, por ejemplo, la biyectabilidad, ya observada por Galileo, de $\mathbb{Z}^+$ con $\{n^2 \mid n \in \mathbb{Z}\}$. El famoso científico renacentista dedujo de este fenómeno que "los atributos de 'igual', 'mayor' y 'menor' no son aplicables a cantidades infinitas", anticipándose a varios desarrollos que tendrían lugar en el siglo XX, en los que resultados paradójicos de esta índole serían utilizados para demostrar la no existencia de ciertas medidas.

Una característica importante de la observación de Galileo es que su construcción muestra cómo, a partir de $\mathbb{Z}^+$, se pueden dar dos conjuntos ajenos (a saber, los números que son cuadrados y los que no lo son), cada uno de los cuales tiene el mismo tamaño que $\mathbb{Z}^+$. Así, hay una idea de duplicación inherente en este ejemplo. La razón por la que el concepto es tan fascinante es que, poco tiempo después de que las paradojas como la de Galileo fuesen aclaradas por la teoría de Cantor sobre cardinalidad, se descubrió que era posible producir duplicaciones aún más contraintuitivas usando movimientos rígidos. Tal es el caso del resultado que nos ocupa.

Los conceptos que se manejan en este primer capítulo son de naturaleza fundamentalmente algebraica. Comenzaremos con una definición que probablemente no le sea familiar al lector. La idea es tratar de capturar la noción de duplicar un conjunto usando ciertas transformaciones. La teoría

general se simplifica si utilizamos permutaciones de un conjunto (es decir, biyecciones de él en sí mismo), y la manera más fácil de hacer esto es trabajar en el contexto de las acciones de grupos.[1]

A lo largo de todo este trabajo, $\mathbb{N}\backslash\mathbb{Z}^+ = \{0\}$.

Para cada $n \in \mathbb{Z}+$, denotaremos con G_n al grupo de isometrías[2] de $\mathbb{R}^n$, y consideraremos la acción natural de este grupo sobre $\mathbb{R}^n$.

Convengamos en que, si G es un grupo que actúa sobre un conjunto X y $A \subseteq X$, entonces para cada $g \in G$, $gA = \{gx \mid x \in A\}$.

Definición 1.1. *Sea G un grupo que actúe sobre un conjunto X, y sea $E \subseteq X$. Decimos que E es **G-paradójico** (o bien, **paradójico con respecto a la acción de G**) si y sólo si para algunos $m, n \in \mathbb{Z}^+$ hay $A_1, \ldots, A_m$, $B_1, \ldots, B_n \subseteq E$ ajenos dos a dos y $g_1, \ldots, g_m, h_1, \ldots, h_n \in G$ tales que $E = \bigcup_{i=1}^{m} g_i A_i$ y además $E = \bigcup_{i=1}^{n} h_i B_i$.*

Intuitivamente, E es G-paradójico si y sólo si tiene dos subconjuntos ajenos ($\bigcup_{i=1}^{m} A_i$, $\bigcup_{i=1}^{n} B_i$) cada uno de los cuales se puede descomponer y reacomodar mediante G para cubrir todo E.

Observemos que, si G es un grupo que actúa sobre cualquier conjunto X, $\emptyset$ es G-paradójico.

En el Apéndice F se da un ejemplo de un subconjunto de $\mathbb{R}^2$ numerable (es decir, biyectable con $\mathbb{N}$) y G_2-paradójico.

El primer resultado que probaremos proporciona el ejemplo por excelencia de paradojicidad. Desempeña, además, un papel medular en la estructura de este trabajo.

Si G es un grupo, la función $\phi : G \times G \to G$ dada por $(g, h) \mapsto gh$ es una acción de G sobre G, como es fácil verificar. Decimos que G actúa sobre sí mismo por **traslación** para referirnos a esa acción. Decimos que G es **paradójico** si y sólo si G es paradójico con respecto a la traslación.

Si el lector no está familiarizado con el concepto de grupo libre, conviene revisar ahora el Apéndice C.

Teorema 1.2. *Sea F un grupo libre de rango 2. Entonces, F es paradójico.*

Demostración. Podemos suponer que F es $F_{\{\sigma,\tau\}}$, el grupo libre generado por $\{\sigma, \tau\}$. Definimos $W(\sigma) = \{\alpha \in F \mid \alpha \text{ comienza con } \sigma\}$, y análogamente $W(\tau)$, $W(\sigma^{-1})$ y $W(\tau^{-1})$. Tenemos que $F = \{1\} \cup W(\sigma) \cup W(\tau) \cup$

[1] Véase el Apéndice A.
[2] Véase el Apéndice D.

$W(\sigma^{-1}) \cup W(\tau^{-1})$, y que los cinco unendos son ajenos dos a dos. Afirmamos que $F = W(\sigma) \cup \sigma W(\sigma^{-1})$ y que $F = W(\tau) \cup \tau W(\tau^{-1})$. Basta probar la primera igualdad, pues la segunda se obtiene de manera análoga. Sea $\alpha \in F \backslash W(\sigma)$. Entonces $\sigma^{-1}\alpha \in W(\sigma^{-1})$, y así $\alpha = \sigma(\sigma^{-1}\alpha) \in \sigma W(\sigma^{-1})$. La otra contención es trivial. □

El siguiente resultado es también fundamental. Nos permite trasladar o "subir" la paradojicidad de un grupo a un conjunto sobre el cual actúe. Es el único momento en el que apelaremos al Axioma de Elección. Antes necesitamos una definición.

Definición 1.3. *Sea G un grupo que actúe sobre un conjunto X. Decimos que la acción es **sin puntos fijos no triviales (sin p.f.n.t.)** si y sólo si no hay $g \in G$, $g \neq 1$, y $x \in X$ tales que $gx = x$.*

Por ejemplo, si $n \in \mathbb{Z}^+$, el grupo de traslaciones de $\mathbb{R}^n$ actúa de manera natural sobre $\mathbb{R}^n$, y esa acción es sin p.f.n.t.

En lo sucesivo, aquellos resultados en cuya prueba se emplee el Axioma de Elección se denotarán con (AE).

Proposición 1.4 (AE)**.** *Sea G un grupo paradójico que actúe sobre un conjunto X sin p.f.n.t. Entonces, X es G-paradójico.*

Demostración. Supongamos que los elementos de $\{A_i\}_{i=1}^m \cup \{B_j\}_{j=1}^n \subseteq \mathscr{P}(G)$ y de $\{g_i\}_{i=1}^m \cup \{h_j\}_{j=1}^n \subseteq G$ atestiguan que G es paradójico. La colección de G-órbitas[3] de X es una partición de X, así que el Axioma de Elección garantiza la existencia de un conjunto M que tenga exactamente un elemento de cada G-órbita de X.

Probemos que los elementos de $\{gM \mid g \in G\}$ son subconjuntos de X ajenos dos a dos. Si $g_1M \cap g_2M \neq \emptyset$ para algunos $g_1, g_2 \in G$, tenemos $g_1m_1 = g_2m_2$ para algunos $m_1, m_2 \in M$, por lo que $m_1 = m_2$ (ya que están en la misma G-órbita) y entonces $g_1m_1 = g_2m_1$, es decir, $m_1 = g_1^{-1}g_2m_1$, por lo que $g_1^{-1}g_2 = 1$ (por ser la acción sin p.f.n.t.), es decir, $g_1 = g_2$ y por tanto $g_1M = g_2M$.

Definimos, para cada $i \in \{1, \ldots, m\}$, $A_i^* = \bigcup\{gM \mid g \in A_i\}$, y para cada $j \in \{1, \ldots, n\}$, $B_j^* = \bigcup\{gM \mid g \in B_j\}$. Verifiquemos que $\{A_i^*\}_{i=1}^m \cup \{B_j^*\}_{j=1}^n$ es una colección de subconjuntos de X ajenos dos a dos. Tomemos, para algunos $i, j \in \{1, \ldots, m\}$, A_i^*, A_j^* tales que $A_i^* \cap A_j^* \neq \emptyset$. Tenemos que, para algunos $r \in A_i$ y $s \in A_j$, $rM \cap sM \neq \emptyset$, de lo que, siguiendo el razonamiento del párrafo anterior, se tiene $r = s$, y como los elementos de

[3] Véase el Apéndice B.

la colección $\{A_i\}_{i=1}^m \cup \{B_j\}_{j=1}^n$ son ajenos dos a dos, tenemos que $A_i = A_j$, de donde $A_i^* = A_j^*$. Se observa que nuestra elección de elementos del conjunto $\{A_i^*\}_{i=1}^m \cup \{B_j^*\}_{j=1}^n$ fue sin pérdida de generalidad.

Por último, probaremos que $X = \bigcup_{i=1}^{m} g_i A_i^*$ (y con eso se tendrá también que $X = \bigcup_{j=1}^{n} h_j B_j^*$, por un razonamiento similar). Sea pues $x \in X$. Como $x \in Gm$ para alguna $m \in M$, tenemos $x = gm$ para alguna $g \in G$. Pero como $G = \bigcup_{i=1}^{m} g_i A_i$, $g = g_i a_i$ para algunos $i \in \{1, \ldots, m\}$ y $a_i \in A_i$, y entonces $x = g_i a_i m \in g_i(a_i M) \subseteq g_i A_i^*$. □

Corolario 1.5 (AE)**.** *Sea F un grupo libre de rango 2 que actúe sobre un conjunto X sin p.f.n.t. Entones, X es F-paradójico.*

Demostración. Se sigue de 1.2 y de 1.4. □

Los siguientes tres resultados no son relevantes para los fines de este trabajo, pero se incluyen por su valor fundamental para el estudio de los grupos paradójicos.

Corolario 1.6 (AE)**.** *Sean G un grupo, y $H \leq G$. Si H es paradójico, G también lo es.*

Demostración. La función $\phi : H \times G \to G$ tal que $\phi : (h, g) \mapsto hg$ es una acción, como es fácil verificar. (Podemos decir, por extensión del concepto, que H actúa sobre G por traslación.) Esta acción es sin p.f.n.t., ya que si $hg = g$ para algunos $h \in H$, $g \in G$, tenemos que $h = 1$. Entonces, si H es paradójico, 1.4 asegura que G es H-paradójico, y por tanto paradójico. □

Si G y H son grupos, denotaremos como $G \cong H$ la condición de que G y H sean isomorfos.

Proposición 1.7. *Sea G un grupo libre no abeliano. Entonces, existe $H \leq G$ tal que H es un grupo libre de rango 2.*

Demostración. Supongamos que G es libre no abeliano. Entonces, hay un conjunto libre de símbolos, M, tal que $G \cong F_M$, donde F_M es el grupo libre generado por M. Como H es no abeliano, F_M es no abeliano. Luego,[4] M tiene al menos dos símbolos distintos, digamos α y β. Podemos estar seguros

[4] Véase el Apéndice C.

de que $\{\alpha, \beta\}$ es un conjunto libre de símbolos porque M es un conjunto libre de símbolos. Así, podemos hablar de $F_{\{\alpha,\beta\}}$. Como $\{\alpha, \beta\} \subseteq M$, tenemos evidentemente que $F_{\{\alpha,\beta\}} \leq F_M$. Como $F_M \cong G$, hay $F \leq G$ tal que $F \cong F_{\{\alpha,\beta\}}$, con lo que se tiene que F es un grupo libre de rango 2. □

Corolario 1.8 (AE)**.** *Sean G un grupo, y $H \leq G$. Si H es libre no abeliano, G es paradójico.*

Demostración. Por 1.7, H (y por tanto G) tiene un subgrupo libre de rango 2. En virtud de 1.2 y de 1.6, G es paradójico. □

Capítulo 2

Equidescomponibilidad

En este capítulo desarrollaremos, empleando técnicas elementales algebraicas y conjuntistas, herramientas que, aplicadas a la noción de paradojicidad, luego nos serán muy útiles. Comenzaremos con algunas definiciones.

Definición 2.1. *Sea G un grupo que actúe sobre un conjunto X, y sean $A, B \subseteq X$. Decimos que A y B son* ***G-congruentes*** *si y sólo si hay $g \in G$ tal que $gA = B$. En tal caso, escribimos $A \simeq_G B$.*

Observación. Si G es un grupo que actúa sobre un conjunto X, entonces $\simeq_G$ es una relación de equivalencia sobre $\mathscr{P}(X)$.

Por ejemplo, si X es un conjunto, S_X (el grupo de permutaciones de X) actúa de manera natural sobre X. Entonces, para cualesquiera $A, B \subseteq X$, $A \simeq_{S_X} B$ si y sólo si $A = B$.

Definición 2.2. *Sea G un grupo que actúe sobre un conjunto X, y sean $A, B \subseteq X$. Decimos que A y B son* ***G-equidescomponibles*** *(o* ***G-congruentes a trozos****) si y sólo si para algún $n \in \mathbb{N}$ hay una partición de A, $\{A_i\}_{i=1}^n$, y una partición de B, $\{B_i\}_{i=1}^n$, tales que, para cada $i \in \{1, \ldots, n\}$, $A_i \simeq_G B_i$. En tal caso, escribimos $A \sim_G B$.*

Observación. Sea G un grupo que actúe sobre un conjunto X, y sean $A, B \subseteq X$. Si $A \simeq_G B$, entonces $A \sim_G B$.

Por ejemplo, la versión fuerte de la Paradoja de Banach-Tarski, que fue planteada informalmente en el resumen de este trabajo, afirma que, para cualesquiera $A, B \subseteq \mathbb{R}^3$, ambos acotados y con interior no vacío, ocurre que $A \sim_{G_3} B$.

Daremos ahora una caracterización de la G-paradojicidad en términos de la G-equidescomponibilidad.

Proposición 2.3. *Sea G un grupo que actúe sobre un conjunto X, y sea $E \subseteq X$. Entonces, E es G-paradójico si y sólo si hay $A, B \subseteq E$ ajenos tales que $A \sim_G E$ y $B \sim_G E$.*

Demostración. Supongamos que E es G-paradójico. Tenemos, para algunos $m, n \in \mathbb{Z}^+$, $\{A_i\}_{i=1}^m \cup \{B_i\}_{i=1}^n$, una colección de subconjuntos de E ajenos dos a dos, y $\{g_i\}_{i=1}^m \cup \{h_i\}_{i=1}^n \subseteq G$ tales que $E = \bigcup_{i=1}^{m} g_i A_i$ y $E = \bigcup_{i=1}^{n} h_i B_i$. Definimos $A_1' = A_1$, $A_2' = A_2 \backslash g_2^{-1}(g_1 A_1')$, $A_3' = A_3 \backslash g_3^{-1}((g_1 A_1') \cup g_2 A_2')$, y en general, para cada $j \in \{1, \ldots, m\}$, $A_j' = A_j \backslash g_j^{-1} \bigcup_{1 \leq i < j} g_i A_i'$.

Afirmamos que $\{g_i A_i'\}_{i=1}^m$ es una colección de subconjuntos de E ajenos dos a dos. Para cada $i \in \{1, \ldots, m\}$, $A_i' \subseteq A_i$, de donde $g_i A_i' \subseteq g_i A_i \subseteq E$. Sean $i, j \in \{1, \ldots, m\}$ tales que $i < j$. Tenemos que $A_j' \cap g_j^{-1} \bigcup_{1 \leq k < j} g_k A_k' = \emptyset$, porque $A_j' = A_j \backslash g_j^{-1} \bigcup_{1 \leq k < j} g_k A_k'$. Entonces, ocurre lo siguiente:

$$\begin{aligned}
(g_j A_j') \cap g_i A_i' &\subseteq (g_j A_j') \cap \bigcup_{1 \leq k < j} g_k A_k' \\
&= (g_j A_j') \cap g_j g_j^{-1} \bigcup_{1 \leq k < j} g_k A_k' \\
&= g_j (A_j' \cap g_j^{-1} \bigcup_{1 \leq k < j} g_k A_k') \\
&= g_j \emptyset \\
&= \emptyset
\end{aligned}$$

Por lo tanto, $(g_j A_j') \cap g_i A_i' = \emptyset$.

Probemos ahora que $E = \bigcup_{i=1}^{m} g_i A_i'$. Basta ver que $E \subseteq \bigcup_{i=1}^{m} g_i A_i'$. Sea $x \in E$. Como $E = \bigcup_{i=1}^{m} g_i A_i$, $x \in g_i A_i$ para alguna $i \in \{1, \ldots, m\}$. Sea $k =$

mín$\{i \in \{1,\dots,m\} \mid x \in g_iA_i\}$. Tenemos lo siguiente:

$$\begin{aligned} g_kA'_k &= g_k(A_k \backslash g_k^{-1} \bigcup_{1\leq i<k} g_iA'_i) \\ &= (g_kA_k)\backslash g_kg_k^{-1} \bigcup_{1\leq i<k} g_iA'_i \\ &= (g_kA_k)\backslash \bigcup_{1\leq i<k} g_iA'_i \end{aligned}$$

Además, para cada i tal que $1 \leq i < k$, por la minimalidad de k, $x \notin g_iA_i$, pero como $A'_i \subseteq A_i$, tenemos que $g_iA'_i \subseteq g_iA_i$, y que por ello $x \notin g_iA'_i$. Por lo tanto, $x \notin \bigcup_{1\leq i<k} g_iA'_i$. Como, por la elección de k, $x \in g_kA_k$, tenemos que $x \in (g_kA_k)\backslash \bigcup_{1\leq i<k} g_iA'_i = g_kA'_k \subseteq \bigcup_{i=1}^{m} g_iA'_i$.

Si ahora definimos, para cada $j \in \{1,\dots,n\}$, $B'_j = B_j \backslash h_j^{-1} \bigcup_{1\leq i<j} h_iB'_i$, un razonamiento similar proporciona que $\{h_iB'_i\}_{i=1}^n$ es una colección de subconjuntos de E ajenos dos a dos que cubre a E.

Se observa que $\{A'_i\}_{i=i}^m \cup \{B'_i\}_{i=i}^n$ es una colección de subconjuntos de E ajenos dos a dos, pues para cada $i \in \{1,\dots,m\}$ y $j \in \{1,\dots,n\}$, $A'_i \subseteq A_i \subseteq E$ y $B'_j \subseteq B_j \subseteq E$, y $A_1,\dots,A_m,B_1,\dots,B_n$ son ajenos dos a dos.

Por supuesto, para cada $i \in \{1,\dots,m\}$ y $j \in \{1,\dots,n\}$, $A'_i = \emptyset$ si y sólo si $g_iA'_i = \emptyset$, y además $B'_j = \emptyset$ si y sólo si $h_jB'_j = \emptyset$. Eliminemos pues los conjuntos vacíos de la lista $A'_1,\dots,A'_m,B'_1,\dots,B'_n$ y renumeremos, para obtener, para algunos $r,s \in \mathbb{N}$, $A''_1,\dots,A''_r,B''_1,\dots,B''_s$ y $g''_1,\dots,g''_r,h''_1,\dots,h''_s$ tales que, si $A = \bigcup_{i=1}^{r} A''_i$ y $B = \bigcup_{i=1}^{s} B''_i$, entonces tenemos que:

(i) $A, B \subseteq E$ con $A \cap B = \emptyset$,

(ii) $\{A''_i\}_{i=1}^r$ es una partición de A y $\{B''_i\}_{i=1}^s$ es una partición de B, y

(iii) $\{g''_iA''_i\}_{i=1}^r$ y $\{h''_iB''_i\}_{i=1}^s$ son ambas particiones de E.

Por lo tanto, tenemos $A, B \subseteq E$ ajenos tales que $A \sim_G E$ y $B \sim_G E$.

La implicación en sentido contrario es inmediata. □

Proposición 2.4. *Sea G es un grupo que actúe sobre un conjunto X. Entonces, $\sim_G$ es una relación de equivalencia sobre $\mathscr{P}(X)$.*

Demostración. Es inmediato que $\sim_G$ es reflexiva y simétrica. Probemos que es transitiva. Sean $A, B, C \subseteq X$ tales que $A \sim_G B$ y $B \sim_G C$. Verifiquemos que $A \sim_G C$. Tenemos, para alguna $m \in \mathbb{N}$, una partición de A, $\{A_i\}_{i=1}^m$, una partición de B, $\{B_i\}_{i=1}^m$, y $\{g_i\}_{i=1}^m \subseteq G$ tales que, para cada $i \in \{1, \ldots, m\}$, $g_i A_i = B_i$. Además, tenemos, para alguna $n \in \mathbb{N}$, otra partición de B, $\{B'_j\}_{j=1}^n$, una partición de C, $\{C_j\}_{j=1}^n$, y $\{h_j\}_{j=1}^n \subseteq G$ tales que, para cada $j \in \{1, \ldots, n\}$, $h_j B'_j = C_j$. Sobreponemos las dos particiones de B que tenemos para obtener $S = \{B_i \cap B'_j \mid i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\}$. Sean $R = \{g_i^{-1}(B_i \cap B'_j) \mid i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\}$ y $T = \{h_j(B_i \cap B'_j) \mid i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\}$.

Afirmamos que R es una colección de subconjuntos de A ajenos dos a dos que cubre a A. Para cualesquiera $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$, tenemos que $g_i^{-1}(B_i \cap B'_j) = (g_i^{-1}B_i) \cap g_i^{-1}B'_j = A_i \cap g_i^{-1}B'_j \subseteq A_i \subseteq A$. Ahora, si $x \in A$, $x \in A_i$ para alguna $i \in \{1, \ldots, m\}$, por lo que $g_i x \in g_i A_i = B_i$. Pero $B_i \subseteq B$, así que $g_i x \in B'_j$ para alguna $j \in \{1, \ldots, n\}$. Por lo tanto, $g_i x \in B_i \cap B'_j$ y entonces $x \in g_i^{-1}(B_i \cap B'_j)$. Por último, sean $i, k \in \{1, \ldots, m\}$ y $j, l \in \{1, \ldots, n\}$ y supongamos que $(g_i^{-1}(B_i \cap B'_j)) \cap (g_k^{-1}(B_k \cap B'_l)) \neq \emptyset$. Tenemos que $(g_i^{-1}B_i) \cap (g_i^{-1}B'_j) \cap (g_k^{-1}B_k) \cap g_k^{-1}B'_l \neq \emptyset$, pero como $g_i^{-1}B_i = A_i$ y $g_k^{-1}B_k = A_k$, tenemos que $A_i \cap A_k \neq \emptyset$ y por lo tanto $i = k$. Ahora, hay $z \in (g_i^{-1}B'_j) \cap g_k^{-1}B'_l$, por lo que $g_i z \in B'_j$ y $g_k z \in B'_l$. Pero como $i = k$, tenemos que $g_i z = g_k z$, por lo que $B'_j \cap B'_l \neq \emptyset$, de donde $j = l$. Por lo tanto, $g_i^{-1}(B_i \cap B'_j) = g_k^{-1}(B_k \cap B'_l)$.

Análogamente, T es una colección de subconjuntos de C ajenos dos a dos que cubre a C. Ahora, sean $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$. Tenemos que $(h_j g_i)(g_i^{-1}(B_i \cap B'_j)) = h_j(B_i \cap B'_j)$. Esto nos permite observar que, para cualesquiera $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$, $g_i^{-1}(B_i \cap B'_j) = \emptyset$ si y sólo si $h_j(B_i \cap B'_j) = \emptyset$. Entonces, si de R y de T eliminamos los elementos vacíos, nos quedan particiones de A y de C, respectivamente, que atestiguan que $A \sim_G C$. □

Lema 2.5. *Sea G un grupo que actúe sobre un conjunto X. Si $A, B \subseteq X$ son tales que $A \sim_G B$, entonces hay $f : A \to B$ biyectiva tal que para todo $C \subseteq A$, $C \sim_G f[C]$. Además, para cualesquiera $A_1, A_2, B_1, B_2 \subseteq X$ tales que $A_1 \cap A_2 = \emptyset = B_1 \cap B_2$, si $A_1 \sim_G B_1$ y $A_2 \sim_G B_2$, entonces $A_1 \cup A_2 \sim_G B_1 \cup B_2$.*

Demostración. Tenemos, para alguna $n \in \mathbb{N}$, una partición de A, $\{A_i\}_{i=1}^n$, una partición de B, $\{B_i\}_{i=1}^n$, y $\{g_i\}_{i=1}^n \subseteq G$ tales que, para cada $i \in \{1, \ldots, n\}$, $g_i A_i = B_i$. Definimos $f : A \to B$ tal que para toda $x \in A$, $f(x) = g_i x$, donde $i \in \{1, \ldots, n\}$ es tal que $x \in A_i$. Es fácil ver que f es

biyectiva. Sea $C \subseteq A$. Consideremos $R = \{C \cap A_i\}_{i=1}^n$ y $S = \{f[C] \cap B_i\}_{i=1}^n$. Es de rutina probar que R es una colección de subconjuntos de C ajenos dos a dos que cubre a C, y que S es una colección de subconjuntos de $f[C]$ ajenos dos a dos que cubre a $f[C]$. Sea $i \in \{1, \ldots, n\}$. Verifiquemos que $g_i(C \cap A_i) = f[C] \cap B_i$. Sea $x \in C \cap A_i$. Tenemos que $g_i x = f(x) \in f[C]$ y además $g_i x \in g_i A_i = B_i$. Ahora, sea $x \in f[C] \cap B_i$. Tenemos que $g_i^{-1} x \in g_i^{-1} B_i = A_i$. Por lo tanto, $f(g_i^{-1} x) = g_i g_i^{-1} x = x$, y entonces $g_i^{-1} x = f^{-1}(x) \in f^{(-1)}[f[C]] = C$. Así, $g_i^{-1} x \in C \cap A_i$ y por tanto $x \in g_i(C \cap A_i)$. Obsérvese que, para cada $i \in \{1, \ldots, n\}$, $C \cap A_i = \emptyset$ si y sólo si $f[C] \cap B_i = \emptyset$. Así, si eliminamos de R y de S los elementos vacíos, nos quedan particiones de C y de $f[C]$, respectivamente, que atestiguan que $C \sim_G f[C]$.

La afirmación restante es evidente. □

Proposición 2.6. *Sea G un grupo que actúe sobre un conjunto X, y sean $E, E' \subseteq X$ tales que $E \sim_G E'$. Si E es G-paradójico, E' también lo es.*

Demostración. Supongamos que E es G-paradójico. Entonces, por 2.3, hay $A, B \subseteq E$ ajenos tales que $A \sim_G E$ y $B \sim_G E$. Ahora, como $E \sim_G E'$, 2.5 da una biyección $f : E \to E'$ tal que $A \sim_G f[A]$ y $B \sim_G f[B]$. Tenemos que $f[A]$ y $f[B]$ son subconjuntos de E' ajenos (por ser f biyectiva), y además que $F[A] \sim_G A \sim_G E \sim_G E'$ y que $F[B] \sim_G B \sim_G E \sim_G E'$. En virtud de 2.3, E' es G-paradójico. □

Si un grupo G actúa sobre un conjunto X, ya sabemos que $\sim_G$ es una relación de equivalencia sobre $\mathscr{P}(X)$. Observemos que 2.6 muestra que la G-paradojicidad se puede ver como una propiedad que compete a las clases de equivalencia.

Introduzcamos ahora algo de notación. Si G es un grupo que actúa sobre un conjunto X, y si $A, B \subseteq X$, escribiremos $A \preccurlyeq_G B$ para denotar la situación de que haya $B' \subseteq B$ tal que $A \sim_G B'$.

Proposición 2.7. *Sea G un grupo que actúe sobre un conjunto X, y sean $A, B \subseteq X$.*

(i) Si $A \sim_G B$, entonces $A \preccurlyeq_G B$.

(ii) Si $A \subseteq B$, entonces $A \preccurlyeq_G B$.

(iii) $\preccurlyeq_G$ es una relación reflexiva sobre $\mathscr{P}(X)$.

(iv) $\preccurlyeq_G$ es una relación transitiva sobre $\mathscr{P}(X)$.

Demostración. Las afirmaciones (i)-(iii) son inmediatas. Probemos (iv). Sean $A, B, C \subseteq X$ tales que $A \preccurlyeq_G B$ y $B \preccurlyeq_G C$. Entonces hay $B' \subseteq B$ tal que $A \sim_G B'$, y además hay $C' \subseteq C$ tal que $B \sim_G C'$. Por 2.5 hay una biyección $f : B \to C'$ tal que $B' \sim_G f[B']$. Entonces, tenemos que $A \sim_G B' \sim_G f[B'] \subseteq C' \subseteq C$, por lo que $A \preccurlyeq_G C$. □

Un resultado de la teoría de conjuntos, conocido como Teorema de Cantor-Schröder-Bernstein, o simplemente Teorema de Schröder-Bernstein, afirma que si A y B son conjuntos tales que hay $f : A \to B$ y $g : B \to A$ inyectivas, entonces A y B son biyectables. Daremos ahora una generalización, debida a Banach, de este famoso teorema.

Teorema 2.8 (Teorema de Banach-Schröder-Bernstein)**.** *Sea X un conjunto, y sea $\sim$ una relación de equivalencia sobre $\mathscr{P}(X)$ que satisfaga las siguientes dos condiciones:*

(i) para cualesquiera $A, B \subseteq X$, si $A \sim B$ entonces hay $f : A \to B$ biyectiva tal que para todo $C \subseteq A$, $C \sim f[C]$, y

(ii) para cualesquiera $A_1, A_2, B_1, B_2 \subseteq X$ tales que $A_1 \cap A_2 = \emptyset = B_1 \cap B_2$, si $A_1 \sim B_1$ y $A_2 \sim B_2$, entonces $A_1 \cup A_2 \sim B_1 \cup B_2$.

Entonces, para cualesquiera $A, B \subseteq X$, si hay $B' \subseteq B$ tal que $A \sim B'$ y hay $A' \subseteq A$ tal que $B \sim A'$, se tiene $A \sim B$.

Demostración. Sean $f : A \to B'$ y $g : B \to A'$ biyecciones dadas por (i). Sea $C_0 = A \backslash A'$, y definimos recursivamente $C_{n+1} = g[f[C_n]]$. Sea $C = \bigcup_{n \in \mathbb{N}} C_n$.

Afirmamos que $A \backslash C = g[B \backslash f[C]]$. Sea $x \in A \backslash C$. Como $A \backslash A' = C_0 \subseteq C$, $A \backslash C \subseteq A'$ y podemos hablar de $g^{-1}(x)$. Si ocurriese que $g^{-1}(x) \in f[C]$, tendríamos que $g^{-1}(x) \in f[C_n]$ para algún $n \in \mathbb{N}$, y que por lo tanto $x \in g[f[C_n]] = C_{n+1} \subseteq C$, cosa que no pasa. Así, $g^{-1}(x) \in B \backslash f[C]$, por lo que $x \in g[B \backslash f[C]]$. Inversamente, sea $x \in g[B \backslash f[C]]$. Entonces, $g^{-1}(x) \in B \backslash f[C]$. Si ocurriese que $x \in C$, tendríamos que $x \in C_n$ para algún $n \in \mathbb{N}$. Ahora, no podría pasar que $x \in C_0 = A \backslash A'$, pues $x \in g[B \backslash f[C]] \subseteq g[B] = A'$. Por lo tanto, estaría pasando que $x \in C_{m+1}$ para algún $m \in \mathbb{N}$, por lo que $x \in g[f[C_m]]$ y entonces $g^{-1}(x) \in f[C_m] \subseteq f[C]$, lo cual es una contradicción. Luego, $x \in A \backslash C$.

Ahora, por la elección de g, $A \backslash C = g[B \backslash f[C]] \sim B \backslash f[C]$, y por la elección de f, $C \sim f[C]$. Entonces, por (ii), $A = (A \backslash C) \cup C \sim (B \backslash f[C]) \cup f[C] = B$. □

Observemos que el teorema clásico de Schröder-Bernstein se sigue de la prueba de 2.8, ya que nunca se usó que $\sim$ fuese una relación sobre la potencia de un conjunto. Si hacemos que $\sim$ sea la biyectabilidad (que es un relacional[1] de equivalencia sobre la clase de todos los conjuntos), entonces se cumplen (i) y (ii) y es fácil ver que, para cualesquiera A, B conjuntos, hay $f : A \to B$ inyectiva si y sólo si A es biyectable con un subconjunto de B.

Apliquemos ahora el Teorema de Banach-Schröder-Bernstein.

Corolario 2.9. *Sea G un grupo que actúe sobre un conjunto X, y sean $A, B \subseteq X$. Si $A \preccurlyeq_G B$ y $B \preccurlyeq_G A$, entonces $A \sim_G B$.*

Demostración. Se sigue del hecho de que 2.5 afirma que $\sim_G$ cumple las condiciones (i) y (ii) del enunciado de 2.8. □

Con lo que tenemos hasta aquí, se observa que, si un grupo G actúa sobre un conjunto X, $\sim_G$ es una relación de equivalencia sobre $\mathscr{P}(X)$ y $\preccurlyeq_G$ se puede ver como un orden parcial reflexivo sobre la partición inducida. Se dice que $\preccurlyeq_G$ es un *preorden* sobre $\mathscr{P}(X)$ con respecto a $\sim_G$.

Concluiremos este capítulo con un resultado sobre G-paradojicidad que hace acopio de gran parte de lo que sabemos hasta ahora.

Corolario 2.10. *Sea G un grupo que actúe sobre un conjunto X, y sea $E \subseteq X$. Entonces E es G-paradójico si y sólo si hay $A, B \subseteq E$ ajenos tales que $A \cup B = E$, $A \sim_G E$ y $B \sim_G E$.*

Demostración. Si E es G-paradójico, por 2.3 hay $A', B' \subseteq E$ ajenos tales que $A' \sim_G E$ y $B' \sim_G E$. Entonces $E \sim_G B' \subseteq E \backslash A' \subseteq E$, de donde $E \preccurlyeq_G E \backslash A' \preccurlyeq_G E$. Entonces, en virtud de 2.9, $E \backslash A' \sim_G E$. Basta ahora notar que $A' \cap (E \backslash A') = \emptyset$ y que $A' \cup (E \backslash A') = E$. Tomemos pues $A = A'$ y $B = E \backslash A'$. La implicación en sentido contrario es inmediata. □

[1] Un relacional es una clase de pares ordenados de conjuntos.

Capítulo 3

La Paradoja de Hausdorff

El objetivo de este capítulo es demostrar un importante resultado que se conoce como Paradoja de Hausdorff. Para ello, probaremos primero un lema un tanto técnico que hace ver que SO_3, que coincide con el grupo de rotaciones de $\mathbb{R}^3$ cuyo eje pasa por el origen, tiene un subgrupo libre de rango 2. Como utilizaremos herramientas de ramas de las matemáticas que antes no han aparecido, revisemos en primer término algunos conceptos pertinentes.

Sea $n \in \mathbb{Z}^+$. De aquí en adelante, cada vez que mencionemos a $\mathbb{R}^n$ lo estaremos viendo como $\mathbb{R}$-espacio vectorial, y además como espacio métrico euclidiano, es decir, con la métrica usual. En el Apéndice D se incluye un repaso de los conceptos relacionados con los grupos de transformaciones, tanto lineales como rígidas, de $\mathbb{R}^n$. Si T es un operador lineal sobre $\mathbb{R}^n$, denotaremos como $[T]$ a la matriz que representa a T en la base canónica. Denotemos como GL_n al grupo de operadores lineales no singulares sobre $\mathbb{R}^n$. Formalmente, se define $SO_n = \{T \in GL_n \mid [T]^{-1} = [T]^t \text{ y } \det[T] = 1\}$, y es fácil ver que $SO_n \leq GL_n$. Lo que se afirma en D.8 es que los elementos de SO_3 son precisamente las rotaciones de $\mathbb{R}^3$ cuyo eje pasa por el origen.

Si G es un grupo y $A \subseteq G$, entonces se define el subgrupo de G generado por A, $\langle A \rangle$, como el $\subseteq$-mínimo subgrupo de G que contiene a A (es decir, $A \subseteq \langle A \rangle \leq G$ y para todo H tal que $A \subseteq H \leq G$, $\langle A \rangle \subseteq H$). Se prueba que, si $A \neq \emptyset$, entonces $\langle A \rangle = \{\alpha_1^{\pm 1} \cdots \alpha_n^{\pm 1} \mid n \in \mathbb{Z}^+,\ \alpha_i \in A \text{ para cada } i \in \{1, \ldots, n\}\}$. Sea G un grupo y $A \subseteq G$, y supongamos que A está escrito como un conjunto libre de símbolos.[1] Tiene sentido entonces hablar de F_A,

[1] Para esto es suficiente que no ocurrra que un elemento de A esté escrito como una sucesión finita de otros elementos de A, que si un elemento de A está escrito como α,

el grupo libre generado por A. Decimos que los elementos de A **generan libremente** a $\langle A\rangle$ (o bien, que A es un **conjunto libre de generadores**) si y sólo si la función $\Phi : F_A \to \langle A\rangle$ dada por $\Phi(\alpha) = \alpha$ es un isomorfismo. Obsérvese que para establecer tal hecho, basta demostrar la inyectividad de Φ (pues es claro que Φ sería un homomorfismo suprayectivo). Para esto, en virtud de que para todo f homomorfismo de grupos f es inyectivo si y sólo si $\ker f = \{1\}$ (donde $\ker f = f^{-1}[\{1\}]$), es suficiente hacer ver que toda B-palabra reducida no trivial es distinta de la identidad de G, donde $B = A \cup \{\rho^{-1} \mid \rho \in A\}$. Si esto sucede, tenemos que $\langle A\rangle$ es un grupo libre de rango $|A|$.[2]

Por último, establezcamos dos convenciones que utilizaremos. Consideraremos que la función arco coseno (arc cos) toma valores en $[0, \pi]$. Así, cuando digamos que un ángulo es de $\arccos(\frac{3}{5})$ radianes, estaremos pensando en un ángulo del primer cuadrante. Decimos que una rotación de $\mathbb{R}^3$ sobre el eje X con un ángulo θ es **en sentido antihorario** si y sólo si, situándose un observador sobre algún punto de la parte positiva del eje X mirando hacia el origen, ve girar al plano YZ en θ radianes en sentido contrario a las manecillas del reloj. Cuando hablemos de rotaciones de $\mathbb{R}^3$ sobre el eje Y (o Z) en sentido antihorario, será de acuerdo con convenciones análogas.

Lema 3.1. *Existen ϕ y ρ, rotaciones de $\mathbb{R}^3$ cada una de las cuales es sobre un eje que pasa por el origen, tales que ϕ y ρ generan libremente a $\langle\{\phi, \rho\}\rangle \leq SO_3$. Por lo tanto, SO_3 tiene un subgrupo libre de rango 2.*

Demostración. Sean ϕ y ρ las rotaciones de $\mathbb{R}^3$ sobre el eje Z y el eje X, respectivamente, ambas en sentido antihorario y con un ángulo de $\arccos(\frac{3}{5})$ radianes.

Dibujando los ejes coordenados de $\mathbb{R}^3$, se observa que

$$[\phi^{\pm 1}] = \begin{pmatrix} \frac{3}{5} & \mp\frac{4}{5} & 0 \\ \pm\frac{4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ y que } [\rho^{\pm 1}] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{5} & \mp\frac{4}{5} \\ 0 & \pm\frac{4}{5} & \frac{3}{5} \end{pmatrix}.$$

En virtud de lo ya explicado, basta ver que si $S = \{\phi^{\pm 1}, \rho^{\pm 1}\}$, entonces ninguna S-palabra reducida no trivial es la identidad. Sea pues w una S-palabra reducida no trivial, y supongamos que $w = 1$. Podemos suponer que w termina en $\phi^{\pm 1}$, ya que si no, podemos conjugar por ϕ para obtener $\phi^{-1}w\phi = \phi^{-1}1\phi = 1$, y entonces tendríamos una S-palabra reducida no trivial, igual a la identidad y que termina en ϕ.

---

ningún elemento de A esté escrito como α^{-1}, y que ningún elemento de A esté escrito como 1. Véase el Apéndice C.

[2]Para que esto tenga sentido se requiere que $|A|$ esté definido. Si A es finito, no hay problema. Véase el Apéndice C.

Afirmamos que, si $\lg(w)$ denota la longitud de w, $w(1,0,0)$ tiene la forma $\frac{1}{5^{\lg(w)}}(a,b,c)$, donde $a,b,c \in \mathbb{Z}$ y $5 \nmid b$. Esto mostrará que $w(1,0,0) \neq (1,0,0)$, lo cual es una contradicción. Probemos la afirmación por inducción sobre $\lg(w)$. Si $\lg(w) = 1$, $w = \phi^{\pm 1}$, y entonces $w(1,0,0) = (\frac{3}{5}, \pm\frac{4}{5}, 0) = \frac{1}{5^1}(3, \pm 4, 0)$, con lo que se tiene la base de la inducción.

Para el paso inductivo, pongamos $w = \phi^{\pm 1}w'$ o $w = \rho^{\pm 1}w'$, con $w'(1,0,0) = \frac{1}{5^{\lg(w')}}(a',b',c')$, con $a',b',c' \in \mathbb{Z}$ y $5 \nmid b'$. Hagamos los cálculos.

Si $w = \phi^{\pm 1}w'$, entonces tenemos lo siguiente:

$$\begin{aligned}
[w]\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} &= [\phi^{\pm 1}][w']\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\
&= \frac{1}{5^{\lg(w')}}\begin{pmatrix} \frac{3}{5} & \mp\frac{4}{5} & 0 \\ \pm\frac{4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} \\
&= \frac{1}{5^{\lg(w')}}\begin{pmatrix} \frac{3a' \mp 4b'}{5} \\ \frac{3b' \pm 4a'}{5} \\ c' \end{pmatrix} \\
&= \frac{1}{5^{\lg(w')+1}}\begin{pmatrix} 3a' \mp 4b' \\ 3b' \pm 4a' \\ 5c' \end{pmatrix}
\end{aligned}$$

Así, $w(1,0,0) = \frac{1}{5^{\lg(w)}}(a,b,c)$, donde $a = 3a' \mp 4b'$, $b = 3b' \pm 4a'$ y $c = 5c'$.

Si $w = \rho^{\pm 1}w'$, entonces tenemos lo siguiente:

$$\begin{aligned}
[w]\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} &= [\rho^{\pm 1}][w']\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\
&= \frac{1}{5^{\lg(w')}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{5} & \mp\frac{4}{5} \\ 0 & \pm\frac{4}{5} & \frac{3}{5} \end{pmatrix}\begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} \\
&= \frac{1}{5^{\lg(w')}}\begin{pmatrix} a' \\ \frac{3b' \mp 4c'}{5} \\ \frac{3c' \pm 4b'}{5} \end{pmatrix} \\
&= \frac{1}{5^{\lg(w')+1}}\begin{pmatrix} 5a' \\ 3b' \mp 4c' \\ 3c' \pm 4b' \end{pmatrix}
\end{aligned}$$

Así, $w(1,0,0) = \frac{1}{5^{\lg(w)}}(a,b,c)$, donde $a = 5a'$, $b = 3b' \mp 4c'$ y $c = 3c' \pm 4b'$.

En cualquier caso, $w(1,0,0) = \frac{1}{5^{\lg(w)}}(a,b,c)$, con $a,b,c \in \mathbb{Z}$. Probemos ahora que $5 \nmid b$. Seguimos suponiendo la misma hipótesis de inducción, es decir, que $w = \phi^{\pm 1}w'$ o $w = \rho^{\pm 1}w'$, con $w'(1,0,0) = \frac{1}{5^{\lg(w')}}(a',b',c')$, con $a',b',c' \in \mathbb{Z}$ y $5 \nmid b'$. Subdividiremos cada uno de los dos casos anteriores en dos subcasos.

Con v denotaremos alguna S-palabra reducida, trivial o no, pero que si no es trivial termina en $\phi^{\pm 1}$. Con lo que tenemos hasta ahora, podemos suponer que $v(1,0,0) = \frac{1}{5^{\lg(v)}}(a'',b'',c'')$, donde $a'',b'',c'' \in \mathbb{Z}$. (Si v es trivial, tenemos que $v(1,0,0) = (1,0,0)$ y que $a'' = 1, b" = 0, c'' = 0$.)

Si $w = \phi^{\pm 1}\rho^{\pm 1}v$, entonces $b = 3b' \pm 4a'$, donde $5 \mid a'$ (pues $a' = 5a''$).

Si $w = \rho^{\pm 1}\phi^{\pm 1}v$, entonces $b = 3b' \mp 4c'$, donde $5 \mid c'$ (pues $c' = 5c''$).

Si $w = \phi^{\pm 1}\phi^{\pm 1}v$, entonces $b = 3b' \pm 4a' = 3b' \pm (12a'' \mp 16b'') = 3b' + 9b'' \pm 12a'' - 16b'' - 9b'' = 3b' + 3(3b'' \pm 4a'') - 25b'' = 6b' - 25b''$.

Si $w = \rho^{\pm 1}\rho^{\pm 1}v$, entonces $b = 3b' \mp 4c' = 3b' \mp (12c'' \pm 16b'') = 3b' + 9b'' \mp 12c'' - 16b'' - 9b'' = 3b' + 3(3b'' \mp 4c'') - 25b'' = 6b' - 25b''$.

Obsérvese que, en los casos primero y cuarto, v no puede ser trivial, pues si lo fuera, tendríamos que w termina en $\rho^{\pm 1}$. □

Si A es un conjunto, decimos que A es **contable** si y sólo si existe $f : A \to \mathbb{N}$ inyectiva.

A partir de este momento, haremos referencia al Apéndice E cada vez que necesitemos un resultado relacionado con cardinalidad que no dependa del Axioma de Elección.

El siguiente resultado devuelve nuestra atención a la discusión sobre conjuntos paradójicos, pero antes de entrar de lleno, adoptemos algunas convenciones adicionales.

Para cada $n \in \mathbb{Z}^+$, denotaremos como $\mathbb{S}^{n-1}$ a la esfera unitaria centrada en el origen en $\mathbb{R}^n$, como es usual.

En lo que resta de este trabajo, consideraremos que SO_3 actúa de manera natural sobre $\mathbb{R}^3$.

Si $P_1, P_2 \in \mathbb{R}^3$, con $P_1 = (x_1, y_1, z_1)$, $P_2 = (x_2, y_2, z_2)$ podemos comparar lexicográficamente a P_1 con P_2 y decimos que $P_1 <_L P_2$ si y sólo si ocurre que $x_1 < x_2$, o bien que $x_1 = x_2$ y $y_1 < y_2$, o bien que $x_1 = x_2$ y $y_1 = y_2$ y $z_1 < z_2$. Es de rutina probar que $<_L$ es un orden estricto total sobre $\mathbb{R}^3$.

Teorema 3.2 (Paradoja de Hausdorff) (AE)**.** *Hay $D \subseteq \mathbb{S}^2$ tal que se puede dar $f : D \to \mathbb{N}$ inyectiva y que $\mathbb{S}^2 \backslash D$ es SO_3-paradójico.*

Demostración. Sea F el subgrupo libre de SO_3 construido en 3.1. Cada elemento distinto de la identidad de F, al ser una rotación no trivial de $\mathbb{R}^3$ cuyo eje pasa por el origen, deja fijos exactamente a dos puntos de $\mathbb{S}^2$, a

saber, las intersecciones de su eje con $\mathbb{S}^2$. Sea $D = \{P \in \mathbb{S}^2 \mid \text{hay } \alpha \in F\backslash\{1\} \text{ tal que } \alpha(P) = P\}$.

Demos ahora una función $f : D \to \mathbb{N}$ inyectiva. En primer lugar, consideremos al conjunto libre de símbolos $M = \{\phi, \rho\}$. Quedó establecido, en 3.1, que la función $\Phi : F_M \to F$ dada por $\Phi(\alpha) = \alpha$ es un isomorfismo, y por tanto una biyección. Ahora, como M es finito, se puede dar $g : M \to \mathbb{N}$ inyectiva. Entonces, en virtud de E.3, se puede dar $h : F_M \to \mathbb{N}$ inyectiva. Por lo tanto, $h \circ \Phi^{-1} : F \to \mathbb{N}$ es inyectiva. Definimos ahora, para cada $\alpha \in F\backslash\{1\}$, $A_\alpha = \{P \in \mathbb{S}^2 \mid \alpha(P) = P\}$. Para cada $\alpha \in F\backslash\{1\}$, tenemos que $|A_\alpha| = 2$ y que la función de A_α en $\mathbb{N}$ que asigna 0 al punto $<_L$-menor y 1 al $<_L$-mayor es inyectiva. Basta ahora observar que $D = \bigcup_{\alpha \in F\backslash\{1\}} A_\alpha$ para obtener que E.1 garantiza que se puede dar $f : D \to \mathbb{N}$ inyectiva.

Ahora, si $P \in \mathbb{S}^2\backslash D$ y $\alpha \in F$, $\alpha(P) \in \mathbb{S}^2\backslash D$, ya que si hubiera $\beta \in F\backslash\{1\}$ que fijara a $\alpha(P)$. tendríamos $\beta(\alpha(P)) = \alpha(P)$, es decir, $(\alpha^{-1}\beta\alpha)(P) = P$, con lo que P sería un punto fijo de $\alpha^{-1}\beta\alpha \in F\backslash\{1\}$ y P estaría en D. Acabamos de ver que F actúa de manera natural sobre $\mathbb{S}^2\backslash D$ y esta acción, por definición de D, es sin p.f.n.t. Entonces, por 1.5, $\mathbb{S}^2\backslash D$ es F-paradójico. Por supuesto, la acción de F sobre $\mathbb{S}^2\backslash D$ es la restricción de la acción natural de F sobre $\mathbb{R}^3$, con respecto a la cual, desde luego, tenemos que $\mathbb{S}^2\backslash D$ es paradójico. Por lo tanto, como $F \leq SO_3$, $\mathbb{S}^2\backslash D$ es SO_3-paradójico. □

Capítulo 4

La Paradoja de Banach-Tarski

Aplicaremos ahora todos los resultados y herramientas que tenemos. El primer resultado que probaremos habla de cardinalidad, así que, para mostrar que no depende del Axioma de Elección, nos referiremos varias veces al Apéndice E.

Proposición 4.1. *Sea $D \subseteq \mathbb{S}^2$ tal que se puede dar $f : D \to \mathbb{N}$ inyectiva. Entonces, $\mathbb{S}^2 \sim_{SO_3} \mathbb{S}^2 \backslash D$.*

Demostración. La función $A : D \to \{-x \mid x \in D\}$ tal que $A : x \mapsto -x$, es decir, la función antípoda, es claramente biyectiva. Definimos $g : D \cup \{-x \mid x \in D\} \to (\mathbb{N}\times\{0\}) \cup (\mathbb{N}\times\{1\})$ tal que $g : P \to \begin{cases} (f(P), 0) & \text{si } P \in D \\ ((f \circ A^{-1})(P), 1) & \text{si } P \notin D \end{cases}$. Se observa que g es inyectiva. Sea $s : (\mathbb{N} \times \{0\}) \cup (\mathbb{N} \times \{1\}) \to \mathbb{N}$ la función inyectiva que dimos en la prueba de E.3. Entonces, $s \circ g : D \cup \{-x \mid x \in D\} \to \mathbb{N}$ es inyectiva. Por lo tanto, $D \cup \{-x \mid x \in D\}$ es contable. Así, como $\mathbb{S}^2$ es biyectable con $\mathbb{R}$, por E.5 hay ℓ, una recta por el origen, que no toca a D. Usando el orden lexicográfico, demos a ℓ un sentido, y convengamos en que cada vez que hablemos de una rotación de $\mathbb{R}^3$ alrededor de ℓ, será en sentido antihorario, si vemos a ℓ como un eje dirigido con el sentido que le dimos. Si $\theta \in \mathbb{R}$, llamemos ρ_θ a la rotación de $\mathbb{R}^3$ alrededor de ℓ con un ángulo de θ radianes. Definimos, para cada $P \in D$ y $n \in \mathbb{Z}^+$, $A_{P,n} = \{\theta \in [0, 2\pi) \mid \rho_\theta^n(P) \in D\}$. Para cada $P \in D$ y $n \in \mathbb{Z}^+$, sea $g_{P,n} : A_{P,n} \to D$ tal que $g_{P,n} : \theta \mapsto \rho_\theta^n(P)$. Para cada $P \in D$ y $n \in \mathbb{Z}^+$, como $P \notin \ell$, $g_{P,n}$ es inyectiva, así que también lo es $f \circ g_{P,n} : A_{P,n} \to \mathbb{N}$. Por

lo tanto, si para cada $P \in D$, $A_P = \bigcup_{n\in\mathbb{Z}^+} A_{P,n}$, E.1 da $f_P : A_P \to \mathbb{N}$ inyectiva. Ahora, sea $A = \bigcup_{P\in D} A_P$. Una segunda aplicación de E.1 da $f : A \to \mathbb{N}$ inyectiva, por lo que podemos asegurar que A es contable.

Conviene observar aquí que $A = \{\theta \in [0, 2\pi) \mid$ hay $n \in \mathbb{Z}^+$ y $P \in D$ tales que $\rho_\theta^n(P) \in D\}$. Como A es contable y $[0, 2\pi)$ es biyectable con $\mathbb{R}$, E.5 da $\theta_0 \in [0, 2\pi)\backslash A$. Sea $\rho = \rho_{\theta_0}$. Por supuesto, $\rho \in SO_3$. Para cualquier $n \in \mathbb{Z}^+$, tenemos, por construcción de A, que $\rho^n[D] \cap D = \emptyset$, ya que si hubiera $P \in D$ con $\rho^n(P) \in D$, tendríamos $\theta_0 \in A$. Sea $\bar{D} = \bigcup_{n\in\mathbb{N}} \rho^n[D]$. Observemos que $\bar{D} = D \cup \bigcup_{n\in\mathbb{Z}^+} \rho^n[D]$ y que $\bigcup_{n\in\mathbb{Z}^+} \rho^n[D] = \rho[\bar{D}]$, con $D \cap \rho[\bar{D}] = \emptyset$, por lo que $\rho[\bar{D}] = \bar{D}\backslash D$. Entonces, tenemos que $\mathbb{S}^2 = \bar{D} \cup (\mathbb{S}^2\backslash\bar{D}) \sim_{SO_3} \rho[\bar{D}] \cup (\mathbb{S}^2\backslash\bar{D}) = (\bar{D}\backslash D) \cup (\mathbb{S}^2\backslash\bar{D}) = (\mathbb{S}^2\backslash D)$. □

Convengamos en que, para cualesquiera $X \subseteq \mathbb{R}^3$ y $r \in \mathbb{R}$, $rX = \{rP \mid P \in X\}$.

Teorema 4.2 (AE)**.** $\mathbb{S}^2$ *es SO_3-paradójico. En general, cualquier esfera en* $\mathbb{R}^3$ *centrada en el origen es SO_3-paradójica.*

Demostración. La Paradoja de Hausdorff (3.2) afirma que hay $D \subseteq \mathbb{S}^2$ tal que se puede dar $f : D \to \mathbb{N}$ inyectiva y que $\mathbb{S}^2\backslash D$ es SO_3-paradójico. Así, 4.1 garantiza que $\mathbb{S}^2\backslash D \sim_{SO_3} \mathbb{S}^2$. Entonces, en virtud de 2.6, resulta que $\mathbb{S}^2$ es SO_3-paradójico.

Ahora, sea $\mathscr{S}$ una esfera de radio r centrada en el origen. En virtud de la linealidad de los elementos de SO_3, podemos hacer corresponder a cada punto $P \in \mathbb{S}^2$ con $rP \in \mathscr{S}$ para obtener una descomposición SO_3-paradójica de $\mathscr{S}$ partiendo de la que ya teníamos para $\mathbb{S}^2$. Hagámoslo con detalle. Tenemos, para algunos $m, n \in \mathbb{Z}^+$, $\{A_i\}_{i=1}^m \cup \{B_i\}_{i=1}^n$, una colección de subconjuntos de $\mathbb{S}^2$ ajenos dos a dos, y $\{\alpha_i\}_{i=1}^m \cup \{\beta_i\}_{i=1}^n \subseteq SO_3$ tales que $\mathbb{S}^2 = \bigcup_{i=1}^m \alpha_i A_i$ y $\mathbb{S}^2 = \bigcup_{i=1}^n \beta_i B_i$. Claramente, $\{rA_i\}_{i=1}^m \cup \{rB_i\}_{i=1}^n$ es una colección de subconjuntos de $\mathscr{S}$ ajenos dos a dos, y además ocurre

lo siguiente:

$$\begin{aligned}\mathscr{S} &= r\mathbb{S}^2 \\ &= r\bigcup_{i=1}^{m}\alpha_i A_i \\ &= \bigcup_{i=1}^{m} r\alpha_i A_i \\ &= \bigcup_{i=1}^{m}\alpha_i r A_i\end{aligned}$$

Análogamente, $\mathscr{S} = \bigcup_{i=1}^{n}\beta_i r B_i$. □

Denotemos como $\mathbf{0}$ a $(0,0,0)$, el origen de $\mathbb{R}^3$.

Lema 4.3. *Sea $\mathscr{B}$ una bola cerrada en $\mathbb{R}^3$, centrada en $\mathbf{0}$. Entonces, $\mathscr{B} \sim_{G_3} \mathscr{B}\backslash\{\mathbf{0}\}$.*

Demostración. Sea $\mathscr{C}$ una circunferencia completamente contenida en $\mathscr{B}$ que pase por $\mathbf{0}$. Sea ℓ la recta normal al plano de $\mathscr{C}$ que pasa por el centro de $\mathscr{C}$. Notemos que, al rotar $\mathbb{R}^3$ alrededor de ℓ, $\mathbf{0}$ describe precisamente a $\mathscr{C}$. Usando el orden lexicográfico, demos a ℓ un sentido. Sea ρ_1 la rotación de $\mathbb{R}^3$ alrededor de ℓ con un ángulo de 1 radián en sentido antihorario, viendo a ℓ como un eje dirigido con el sentido que le dimos. Observemos que $\rho_1 \in G_3$, y que para todo $n \in \mathbb{Z}^+$, $\rho_1^n(\mathbf{0}) \neq \mathbf{0}$, ya que si para algún $n \in \mathbb{Z}^+$, $\rho_1^n(\mathbf{0}) = \mathbf{0}$, tendríamos que $n = 2\pi k$ para alguna $k \in \mathbb{Z}^+$, y entonces que $\pi = \frac{n}{2k}$. Entonces, si $A = \{\rho_1^n(\mathbf{0}) \mid n \in \mathbb{N}\}$, tenemos que $A\backslash\{\mathbf{0}\} = \{\rho_1^n(\mathbf{0}) \mid n \in \mathbb{Z}^+\} = \rho_1[A]$ y que $\mathscr{B} = A \cup (\mathscr{B}\backslash A) \sim_{G_3} \rho_1[A] \cup (\mathscr{B}\backslash A) = A\backslash\{\mathbf{0}\} \cup (\mathscr{B}\backslash A) = \mathscr{B}\backslash\{\mathbf{0}\}$. □

Teorema 4.4 (Paradoja de Banach-Tarski) (AE)**.** *Cualquier bola cerrada en $\mathbb{R}^3$ es G_3-paradójica.*

Demostración. Sea $\mathscr{B}$ una bola cerrada en $\mathbb{R}^3$. Como todas las traslaciones son elementos de G_3, podemos suponer que $\mathscr{B}$ está centrada en $\mathbf{0}$. Llamemos $\mathscr{S}$ a la frontera de $\mathscr{B}$. Por supuesto, $\mathscr{S}$ es una esfera centrada en $\mathbf{0}$. Por 4.2, $\mathscr{S}$ es SO_3-paradójico.

En virtud de la linealidad de los elementos de SO_3, podemos hacer corresponder a cada punto $P \in \mathscr{S}$ con el radio $\{\alpha P \mid 0 < \alpha \leq 1\} \subseteq \mathscr{B}\backslash\mathbf{0}$ para obtener una descomposición SO_3-paradójica de $\mathscr{B}\backslash\mathbf{0}$ partiendo de la que ya

teníamos para $\mathscr{S}$. Probemos esto con detalle. Tenemos, para algunos $m, n \in \mathbb{Z}^+$, $\{A_i\}_{i=1}^m \cup \{B_i\}_{i=1}^n$, una colección de subconjuntos de $\mathscr{S}$ ajenos dos a dos, y $\{\alpha_i\}_{i=1}^m \cup \{\beta_i\}_{i=1}^n \subseteq SO_3$ tales que $\mathscr{S} = \bigcup_{i=1}^m \alpha_i A_i$ y $\mathscr{S} = \bigcup_{i=1}^n \beta_i B_i$. Claramente, $\{\bigcup_{P\in A_i}\{\gamma P \mid 0<\gamma\leq 1\}\}_{i=1}^m \cup \{\bigcup_{P\in B_i}\{\gamma P \mid 0<\gamma\leq 1\}\}_{i=1}^n$ es una colección de subconjuntos de $\mathscr{B}\backslash\mathbf{0}$ ajenos dos a dos, y además ocurre lo siguiente:

$$\begin{aligned}
\mathscr{B}\backslash\mathbf{0} &= \bigcup_{Q\in\mathscr{S}}\{\gamma Q \mid 0<\gamma\leq 1\}\\
&= \bigcup_{i=1}^m \bigcup_{Q\in\alpha_i A_i}\{\gamma Q \mid 0<\gamma\leq 1\}\\
&= \bigcup_{i=1}^m \bigcup_{P\in A_i}\{\gamma\alpha_i P \mid 0<\gamma\leq 1\}\\
&= \bigcup_{i=1}^m \bigcup_{P\in A_i}\{\alpha_i\gamma P \mid 0<\gamma\leq 1\}\\
&= \bigcup_{i=1}^m \bigcup_{P\in A_i}\alpha_i\{\gamma P \mid 0<\gamma\leq 1\}\\
&= \bigcup_{i=1}^m \alpha_i \bigcup_{P\in A_i}\{\gamma P \mid 0<\gamma\leq 1\}
\end{aligned}$$

Análogamente, $\mathscr{B}\backslash\mathbf{0} = \bigcup_{i=1}^n \beta_i \bigcup_{P\in B_i}\{\gamma P \mid 0<\gamma\leq 1\}$.

Como $SO_3 \leq G_3$ (véase el Apéndice D), $\mathscr{B}\backslash\mathbf{0}$ es G_3-paradójico. Ahora, 4.3 da que $\mathscr{B}\backslash\mathbf{0} \sim_{G_3} \mathscr{B}$. Entonces, 2.6 proporciona el resultado. □

Corolario 4.5 (AE)**.** *Sean $\mathscr{B}_1$, $\mathscr{B}_2$ y $\mathscr{B}_3$ bolas cerradas en $\mathbb{R}^3$, las tres del mismo radio, tales que $\mathscr{B}_2$ y $\mathscr{B}_3$ son ajenos. Entonces, $\mathscr{B}_1 \sim_{G_3} \mathscr{B}_2 \cup \mathscr{B}_3$.*

Demostración. Por 4.4, $\mathscr{B}_1$ es G_3-paradójico. Entonces, por 2.10, hay $A, B \subseteq \mathscr{B}_1$ ajenos tales que $A \cup B = \mathscr{B}_1$, $A \sim_{G_3} \mathscr{B}_1$ y $B \sim_{G_3} \mathscr{B}_1$. Ahora, hay una traslación de $\mathbb{R}^3$ que atestigua que $\mathscr{B}_1 \simeq_{G_3} \mathscr{B}_2$, de donde $\mathscr{B}_1 \sim_{G_3} \mathscr{B}_2$ y por tanto $A \sim_{G_3} \mathscr{B}_2$. Simétricamente, $B \sim_{G_3} \mathscr{B}_3$. Por lo tanto, $\mathscr{B}_1 = A \cup B \sim_{G_3} \mathscr{B}_2 \cup \mathscr{B}_3$. □

Observación. La versión intuitiva (y mejor conocida) de la Paradoja de Banach-Tarski, planteada en la introducción de este trabajo, queda formalizada en 4.5.

Si $\mathbf{x} \in \mathbb{R}^3$ y $r \in \mathbb{R}^+$, llamemos $V_r(\mathbf{x})$ a la vecindad o bola abierta de radio r centrada en $\mathbf{x}$ y $B_r(\mathbf{x})$ a la bola cerrada de radio r centrada en $\mathbf{x}$.

Teorema 4.6 (Paradoja de Banach-Tarski, Forma Fuerte) (AE)**.** *Sean $A, B \subseteq \mathbb{R}^3$, ambos acotados y con interior no vacío. Entonces, $A \sim_{G_3} B$.*

Demostración. Sean K, L bolas cerradas tales que $A \subseteq K$ y $L \subseteq B$ (hay tal K por ser acotado A, y hay tal L por ser no vacío el interior de B). Digamos que K es de radio k y L es de radio l. Consideremos $\{V_l(\mathbf{x})\}_{\mathbf{x}\in K}$, que es una cubierta abierta de K. Como K es cerrado y acotado, es compacto, y por lo tanto podemos extraer una subcubierta finita, digamos $\{V_l(\mathbf{x}_i)\}_{i=1}^n$, para algún $n \in \mathbb{Z}^+$. Sea $R = \bigcup_{i=1}^{n} B_l(\mathbf{x}_i)$. Definimos, para cada $j \in \{1, \ldots, n\}$, $S_j = B_l(\mathbf{x}_j) \setminus \bigcup_{1\leq i<j} S_i$.

Verifiquemos que $\{S_i\}_{i=1}^n$ es una colección de subconjuntos de R ajenos dos a dos que cubre a R. En primer lugar, para cada $i \in \{1, \ldots, n\}$, tenemos que $S_i \subseteq B_l(\mathbf{x}_i) \subseteq R$. Ahora, para cualesquiera $i, j \in \{1, \ldots, n\}$ tales que $i < j$, $S_j \cap S_i \subseteq S_j \cap \bigcup_{1\leq k<j} S_k = \emptyset$. Por último, sea $\mathbf{x} \in R$. Por definición de R, hay $i \in \{1, \ldots, n\}$ tal que $\mathbf{x} \in B_l(\mathbf{x}_i)$. Sea $m = \text{mín}\{i \in \{1, \ldots, n\} \mid \mathbf{x} \in B_l(\mathbf{x}_i)\}$. Para cada i tal que $1 \leq i < m$, $\mathbf{x} \notin B_l(\mathbf{x}_i)$, por lo que $\mathbf{x} \notin S_i$. Así, $\mathbf{x} \notin \bigcup_{1\leq i<m} S_i$. Entonces, $\mathbf{x} \in B_l(\mathbf{x}_m) \setminus \bigcup_{1\leq i<m} S_i = S_m \subseteq \bigcup_{i=1}^{n} S_i$. Por lo tanto, $R \subseteq \bigcup_{i=1}^{n} S_i$.

Sean $T_1, \ldots, T_n$ bolas cerradas, todas de radio l y ajenas dos a dos. Por supuesto, para cada $i \in \{1, \ldots, n\}$, si α_i es la traslación tal que $\alpha_i B_l(x_i) = T_i$, tenemos que $\alpha_i S_i \subseteq \alpha_i B_l(x_i) = T_i$. Sea $T = \bigcup_{i=1}^{n} T_i$. Entonces, $R = \bigcup_{i=1}^{n} S_i \sim_{G_3} \bigcup_{i=1}^{n} \alpha_i S_i \subseteq T$. Aplicando $n - 1$ veces 4.5, obtenemos que $L \sim_{G_3} T$. Por lo tanto, $A \subseteq K \subseteq R \preccurlyeq_{G_3} T \sim_{G_3} L \subseteq B$. Entonces, $A \preccurlyeq_{G_3} B$. Por simetría, $B \preccurlyeq_{G_3} A$. Por 2.9, $A \sim_{G_3} B$. □

Observación. De los diferentes tipos de isometrías de $\mathbb{R}^3$, en toda la demostración sólo se emplearon rotaciones y traslaciones, que son las que más se acercan a la realidad física.

En el último resultado, A puede ser una bola cerrada del tamaño de un guisante y B una del tamaño del sol. De ahí el nombre de la Paradoja del Guisante y el Sol.

Apéndice A

Acciones

Sea G un grupo y X un conjunto cualquiera. Sea $\phi : G \times X \to X$. Decimos que ϕ es una **acción** de G sobre X si y sólo si para cualesquiera $x \in X$ y $g, h \in G$, ocurren las siguientes dos cosas:

(i) $\phi(1, x) = x$

(ii) $\phi(g, \phi(h, x)) = \phi(gh, x)$

Para cualesquiera $g \in G$ y $x \in X$, se acostumbra denotar a $\phi(g, x)$ como gx. Así, las condiciones (i) y (ii) se pueden reescribir de la sigiente manera:

(i) $1x = x$

(ii) $g(hx) = (gh)x$

Decimos que G actúa sobre X mediante ϕ si y sólo si ϕ es una acción de G sobre X. En general, diremos simplemente que G actúa sobre X y denotaremos a la acción mediante la cual esto sucede como $(g, x) \mapsto gx$.

Ahora bien, si un grupo G actúa sobre un conjunto X, se puede definir, para cada $g \in G$, una función $\phi_g : X \to X$ mediante $\phi_g(x) = gx$ para todo $x \in X$. Si $A \subseteq X$, y si, como en el texto, para cada $g \in G$, $gA = \{gx \mid x \in A\}$, entonces claramente $gA = \phi_g[A]$. Probar que ϕ_g es biyectiva es de rutina, pero lo haremos aquí por su valor didáctico. Verifiquemos primero que es inyectiva. Sean $x, y \in X$ tales que $\phi_g(x) = \phi_g(y)$. Entonces, $x = 1x = (g^{-1}g)x = g^{-1}(gx) = g^{-1}(\phi_g(x)) = g^{-1}(\phi_g(y)) = g^{-1}(gy) = (g^{-1}g)y = 1y = y$. Probemos ahora que es suprayectiva. Sea $y \in X$. Entonces, $g^{-1}y \in X$, y $\phi_g(g^{-1}y) = g(g^{-1}y) = (gg^{-1})y = 1y = y$.

Sea G un grupo que actúe sobre un conjunto X. Si denotamos con S_X al grupo de permutaciones de X con la composición de funciones, y si $\psi : G \to S_X$ está dada por $\psi(g) = \phi_g$ para toda $g \in G$ (definiendo ϕ_g como en el párrafo anterior), entonces ψ es un homomorfismo de grupos. Para obtener eso, basta tomar $g, h \in G$ y probar que $\psi(gh) = \psi(g) \circ \psi(h)$. Sea pues $x \in X$. Tenemos que $\psi(gh)(x) = \phi_{gh}(x) = (gh)x = g(hx) = \phi_g(\phi_h(x)) = \psi(g)(\psi(h)(x)) = (\psi(g) \circ \psi(h))(x)$, que es lo que estaba por demostrar. Observemos que, como todo homomorfismo manda el neutro en el neutro, para cualquier $g \in G$, $\phi_g \circ \phi_{g^{-1}} = \phi_{gg^{-1}} = \phi_1 = id_X$. Así, la inversa de ϕ_g es $\phi_{g^{-1}}$.

Recíprocanente, para un grupo G y un conjunto X, si $f : G \to S_X$ es un homomorfismo, $\phi : G \times X \to X$ dada por $\phi : (g, x) \mapsto f(g)(x)$ es una acción de G sobre X. Probémoslo. Sean $x \in X$, $g, h \in G$. Tenemos que $\phi(1, x) = f(1)(x) = id_X(x) = x$ y que $\phi(g, \phi(h, x)) = f(g)(f(h)(x)) = (f(g) \circ f(h))(x) = f(gh)(x) = \phi(gh, x)$. Podemos llamar a ϕ la acción inducida por f. En caso de que $G \subseteq S_X$, $i : G \to S_X$ (la inclusión) es, desde luego, un homomorfismo, y a la acción inducida por i se le llama la acción natural de G sobre X, con la cual $(g, x) \mapsto g(x)$.

En este apéndice quedó probado el siguiente importante resultado:

Teorema A.1. *Si G es un grupo y X es un conjunto, entonces las acciones de G sobre X están en correspondencia biyectiva con los homomorfismos de G en S_X.*

Apéndice B

Órbitas

Definición B.1. *Sea G un grupo que actúe sobre un conjunto X. Para cada $x \in X$, definimos Gx, la* ***G-órbita*** *de x (o simplemente la* ***órbita*** *de x), como $Gx = \{gx \mid g \in G\}$.*

Un ejemplo sencillo, que le da nombre al concepto de G-órbita, es el que daremos a continuación. Sea D el grupo de rotaciones de $\mathbb{R}^2$ alrededor del origen. Tenemos que D actúa de manera natural sobre $\mathbb{R}^2$. Para cualquier $\mathbf{x} \in \mathbb{R}^2$, Dx es la circunferencia con centro en el origen que pasa por $\mathbf{x}$.

Otro ejemplo es el siguiente. Si $n \in \mathbb{Z}^+$ y S es un alfabeto (es decir, un conjunto de símbolos), tenemos que S_n, el grupo de permutaciones del conjunto $\{1, \ldots, n\}$, actúa sobre ${}^nS$, el conjunto de S-palabras (es decir, sucesiones finitas de símbolos de S) de longitud n, de manera que para $\alpha \in S_n$ y $w \in {}^nS$, αw se obtiene permutando las letras de w según indica α.[1] Entonces, para cualquier $v \in {}^nS$, $S_n v = \{u \in {}^nS \mid u \text{ es un anagrama de } v\}$.

Proposición B.2. *Sea G un grupo que actúe sobre un conjunto X. Entonces, $\{Gx \mid x \in X\}$, el conjunto de G-órbitas del conjunto X, es una partición de X.*

Demostración. Como para cada $x \in X$, $x = 1x \in Gx$, se tiene que cada órbita es no vacía y que $X = \bigcup_{x \in X} Gx$. Ahora, supongamos que para algunos $x, y \in X$, $Gx \cap Gy \neq \emptyset$, y probemos que $Gx = Gy$. Tomemos $z \in Gx \cap Gy$, así que $g_0 x = z = g_1 y$ para algunos $g_0, g_1 \in G$. Lo que probaremos es que

[1] Formalmente, si pensamos en que $w : \{1, \ldots, n\} \to S$, entonces $\alpha w = w \circ \alpha$.

$Gx \subseteq Gy$. Para ello, tomemos $w \in Gx$. Por definición de Gx, $w = gx$ para alguna $g \in G$. Como $g_0 x = g_1 y$, tenemos que $x = g_0^{-1}(g_1 y)$, y entonces $w = gx = g(g_0^{-1}(g_1 y)) = (gg_0^{-1}g_1)y \in Gy$. La otra contención se tiene por simetría. □

Proposición B.3. *Sea G un grupo que actúe sobre un conjunto X. Para cualesquiera $x, y \in X$, las siguientes afirmaciones son equivalentes:*

(i) *x, y están en la misma órbita.*

(ii) *$Gx = Gy$.*

(iii) *Hay $g \in G$ tal que $x = gy$.*

Demostración. Para obtener que (i) implica (ii), notemos que si hay un elemento $z \in X$ tal que $x, y \in Gz$, entonces $Gx = Gz$, ya que x pertenece a ambas, y $Gz = Gy$, ya que y pertenece a ambas. El hecho de que (ii) implica (iii) se sigue de que $x \in Gx$ y de la definición de Gy. Por último, si tenemos (iii), $x, y \in Gy$, y entonces tenemos (i). □

Apéndice C

Grupos Libres

Se conviene en que, al decir que M es un conjunto de símbolos, se entiende que ningún símbolo de M es una sucesión finita de otros símbolos de M. Por ejemplo, $\{a, b, 0, \zeta, \aleph\}$ es un conjunto de símbolos. Hacemos énfasis en el hecho de que un conjunto arbitrario de símbolos M, visto como tal, no tiene estructura de grupo ni de nada más que de conjunto, por lo que si hablamos de, por ejemplo, $\rho^{-1} \in M$, sólo estamos refiriéndonos al *símbolo* ρ^{-1}, y no al "inverso" (?) de ρ (tal cosa no tendría sentido). Decimos que un conjunto de símbolos M es libre si y sólo si para todo $\alpha \in M$, α no es el símbolo 1, y además el símbolo $\alpha^{-1} \notin M$. Por ejemplo, $\{\sigma, \tau, \sigma^{-1}\}$ es un conjunto de símbolos que no es libre. Sea pues M un conjunto libre de símbolos. Consideremos el conjunto de símbolos $S = M \cup \{\rho^{-1} \mid \rho \in M\}$. Sea $\mathbb{E}_S$ el conjunto de todas las S-palabras, es decir, sucesiones finitas de símbolos de S. Por ejemplo, si $M = \{\xi, \mu\}$, un elemento típico de $\mathbb{E}_S$ es $\mu^{-1}\xi^{-1}\xi\mu\mu\xi^{-1}\xi\mu\mu^{-1}\mu^{-1}\xi^{-1}\xi\xi^{-1}$. Decimos que dos S-palabras son equivalentes si y sólo si una de ellas se puede obtener de la otra eliminando o insertando una cantidad finita de pares de letras adyacentes de la forma $\rho\rho^{-1}$ ó $\rho^{-1}\rho$. (Obsérvese que esta relación es de equivalencia.) Decimos que una S-palabra está reducida si y sólo si en ella no aparece ninguno de tales pares. Para no usar clases de equivalencia, tomemos $\mathbb{E}'_S = \{w \in \mathbb{E}_S \mid w$ está reducida$\}$. Por ejemplo, si $M = \{\beth, \mathscr{A}\}$, un elemento típico de $\mathbb{E}'_S$ es $\beth^{-1}\mathscr{A}\mathscr{A}\beth\mathscr{A}^{-1}\beth^{-1}\beth^{-1}\mathscr{A}\beth\mathscr{A}^{-1}\beth$, pero $\mathscr{A}\mathscr{A}^{-1}\mathscr{A} \in \mathbb{E}_S \backslash \mathbb{E}'_S$. Si w es una S-palabra, reducir a w significa tomar la única S-palabra reducida a la cual w es equivalente. Demos a $\mathbb{E}'_S$ la operación binaria que consiste en concatenar S-palabras reducidas y reducir la S-palabra resultante. Con esta operación, $\mathbb{E}'_S$ es un grupo, al cual llamaremos F_M, **el grupo libre generado por**

M. La identidad de F_M es la palabra vacía, que denotamos con 1, y si $\alpha \in \mathbb{E}'_S$, el inverso de α se obtiene invirtiendo el orden de los símbolos de α y reemplazando cada símbolo por el símbolo con el cual se cancela (es decir, reemplazando cada símbolo $\rho \in M$ por el símbolo ρ^{-1}, y cada símbolo ρ^{-1} con $\rho \in M$ por el símbolo ρ).

Si M es un conjunto libre de símbolos[1], llamaremos a $|M|$ el **rango** de F_M. Sean M y M' conjuntos libres de símbolos. Claramente, tenemos que M y M' son biyectables si y sólo si F_M y $F_{M'}$ son isomorfos. Decimos que un grupo G es libre si y sólo si es isomorfo a F_M para algún conjunto libre de símbolos M, y en tal caso diremos que su rango es $|M|$. Observemos que un grupo libre de rango 0 es trivial y uno de rango 1 es isomorfo a $(\mathbb{Z}, +)$. Ahora, si M es un conjunto libre de símbolos que tenga al menos dos elementos, hay en M dos símbolos distintos, digamos α y β, y entonces tenemos que $\alpha, \beta \in F_M$ no conmutan. Así, F_M es no abeliano. Se sigue que todo grupo libre de rango por lo menos 2 es no abeliano. Por lo tanto, un grupo libre es abeliano si y sólo si su rango es 0 ó 1.

[1]Formalmente, se necesita que M sea bien ordenable (es decir, que exista $R \subseteq M \times M$ tal que R sea un buen orden sobre M) para poder hablar de $|M|$. Si M es finito, no hay problema, pues M es bien ordenable (podemos dar a M el orden del número natural con el cual M es biyectable, y $|M|$ es ese número natural).

Apéndice D

Grupos de Transformaciones Euclidianas

Sea X un espacio métrico, digamos con métrica[1] d. Decimos que una función $f : X \to X$ **preserva distancias** si y sólo si para cualesquiera $x, y \in X$, $d(x, y) = d(f(x), f(y))$. Una **isometría** de X es una función $f : X \to X$ biyectiva que preserva distancias. Es de rutina probar que el conjunto de isometrías de X es un grupo con la composición de funciones. Para cualquier entero positivo n, el espacio métrico que consiste en $\mathbb{R}^n$ con la métrica euclidiana[2] se conoce como un espacio euclidiano. Denotemos con G_n al grupo de isometrías de $\mathbb{R}^n$. A los elementos de G_n se les llama comúnmente **movimientos rígidos**, o **transformaciones rígidas**.

Revisemos ahora algunas ideas propias del álgebra lineal. Se pueden ver demostraciones de los resultados que enunciaremos aquí en cualquier texto de álgebra lineal, por ejemplo [H]. Sean F un campo y V, W espacios vectoriales sobre F. Decimos que $T : V \to W$ es una **transformación lineal** si y sólo si para cualesquiera $u, v \in V$ y $\alpha \in F$, ocurre que $T(u+v) = Tu + Tv$ y que $T(\alpha v) = \alpha Tv$ (es usual escribir Tv en vez de $T(v)$). En tal caso, son resultados elementales del álgebra lineal que $T(0_V) = 0_W$, donde 0_V es el vector cero de V y 0_W el de W, y que T es inyectiva si y sólo

[1] Aquí utilizaremos las palabras *métrica* y *distancia* como sinónimos.

[2] Es decir, si $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ con $\mathbf{u} = (x_1, \ldots, x_n)$ y $\mathbf{v} = (y_1, \ldots, y_n)$, entonces $d(\mathbf{u}, \mathbf{v}) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$.

si $\ker T = \{0_V\}$, donde $\ker T = T^{-1}[\{0_W\}]$. Es de rutina probar que un composición de transformaciones lineales es una transformación lineal. Un **operador lineal** sobre V es una transformación lineal $T : V \to V$. Decimos que un operador lineal T sobre V es **no singular** si y sólo si $\ker T = \{0_V\}$. Un resultado de álgebra lineal afirma que, si T es un operador lineal sobre V y la dimensión de V es finita, entonces que T sea inyectivo, que T sea biyectivo y que T sea suprayectivo son tres condiciones equivalentes. Es fácil probar que el conjunto de operadores lineales no singulares sobre V es un grupo con la composición de funciones.

Para cada $n \in \mathbb{Z}^+$, convengamos en ver a $\mathbb{R}^n$ como espacio vectorial sobre $\mathbb{R}$. Denotemos con GL_n al grupo de operadores lineales no singulares sobre $\mathbb{R}^n$. Notemos que $G_n \nsubseteq GL_n$ (ya que si α es una traslación no trivial de $\mathbb{R}^n$, $\alpha \in G_n$ pero como α mueve al origen, $\alpha \notin GL_n$), y que $GL_n \nsubseteq G_n$ (ya que si $T : \mathbb{R}^n \to \mathbb{R}^n$ es tal que $T : \mathbf{v} \mapsto 2\mathbf{v}$, $T \in GL_n$ pero $T \notin G_n$). Observemos que cada transformación en $G_n \cap GL_n$ se puede restringir para obtener una isometría de $\mathbb{S}^{n-1}$, la esfera unitaria centrada en el origen vista como subespacio métrico de $\mathbb{R}^n$.

Adoptemos un poco de terminología y de notación. Denotemos como $\mathbf{0}$ al origen de $\mathbb{R}^n$. Para $\mathbf{v} \in \mathbb{R}^n$, digamos $\mathbf{v} = (x_1, \ldots, x_n)$, definimos el vector columna $[\mathbf{v}] = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, que es una matriz de $n \times 1$. Con esto, queda definido el vector renglón $[\mathbf{v}]^t = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}$, que es una matriz de $1 \times n$. (Si A es una matriz, A^t denota la traspuesta de A.) Llamemos $\mathcal{M}_n(\mathbb{R})$ al conjunto de matrices de $n \times n$ con entradas en $\mathbb{R}$. Si $i, j \in \{1, \cdots, n\}$ y $A \in \mathcal{M}_n(\mathbb{R})$, denotaremos como A_i a la matriz de $1 \times n$ cuyas entradas forman el i-ésimo renglón de A, como A^j a la matriz de $n \times 1$ cuyas entradas forman la j-ésima columna de A, y como A_{ij} a la entrada de A que se encuentra en el i-ésimo renglón y en la j-ésima columna. Adoptemos también la convención de que $|A| = \det A$. Definimos la delta de Kronecker, $\delta_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$, para cualesquiera $i, j \in \mathbb{Z}^+$. Como es usual, $I \in \mathcal{M}_n(\mathbb{R})$ es la matriz identidad, es decir, para cada $i, j \in \{1, \cdots, n\}$, $I_{ij} = \delta_i j$. Por último, si $(a) \in \mathcal{M}_1(\mathbb{R})$, en vez de (a) escribiremos simplemente a.

La base canónica de $\mathbb{R}^n$ es $\{e_1, \ldots, e_n\}$, donde para cada $i \in \{1, \ldots, n\}$, $e_i = (0, \ldots, 0, \underbrace{1}_{i\text{-ésimo lugar}}, 0, \ldots, 0)$. Para un operador lineal T sobre $\mathbb{R}^n$, denotemos con $[T]$ a la matriz que representa a T en la base canónica. Recordemos que $[T] \in \mathcal{M}_n(\mathbb{R})$ es tal que, para cada $j \in \{1, \ldots, n\}$, $[T]^j = [Te_j]$. Es

de suma importancia que la función tal que $T \mapsto [T]$ (con contradominio $\mathcal{M}_n(\mathbb{R})$) es una biyección, y que para un operador lineal T sobre $\mathbb{R}^n$ y un vector $\mathbf{v} \in \mathbb{R}^n$, se tiene que $[Tv] = [T][v]$. Además, para cualesquiera T, S operadores lineales sobre $\mathbb{R}^n$ se tiene que $[S \circ T] = [S][T]$, y que $T \in GL_n$ si y sólo si $[T]$ es invertible (y en tal caso $[T]^{-1} = [T^{-1}]$). Por supuesto, $[id_{\mathbb{R}^n}] = I$. Recordemos también que para cualesquiera $A, B \in \mathcal{M}_n(\mathbb{R})$, se tiene que $(AB)^t = B^t A^t$, que $|A| = |A^t|$ y además que A es invertible si y sólo si $|A| \neq 0$.

Antes de comenzar con los resultados, recordemos que para $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, decimos que $\mathbf{u}$ y $\mathbf{v}$ son **paralelos** si $\mathbf{u} = \alpha\mathbf{v}$ para alguna $\alpha \in \mathbb{R}\backslash\{0\}$. Si $\mathbf{u} = (x_1, \ldots, x_n)$ y $\mathbf{v} = (y_1, \ldots, y_n)$, se define el **producto punto** (o **producto escalar**) de $\mathbf{u}$ y $\mathbf{v}$ como $\mathbf{u} \cdot \mathbf{v} = x_1y_1 + \cdots + x_ny_n$, y la **norma** de $\mathbf{v}$ como $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$.[3] Supongamos, para lo que resta de este párrafo, que $\mathbf{u} \neq \mathbf{0} \neq \mathbf{v}$. Para $n \leq 3$, se prueba que, si θ es el menor ángulo entre los vectores $\mathbf{u}$ y $\mathbf{v}$, se tiene que $\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}$, mientras que para $n > 3$, el ángulo entre $\mathbf{u}$ y $\mathbf{v}$ se define como el número $\theta \in [0, \pi]$ que satisface esa igualdad. Para cualquier $n \in \mathbb{Z}^+$, decimos que $\mathbf{u}$ y $\mathbf{v}$ son **perpendiculares** si y sólo si entre ellos se forma un ángulo de $\frac{\pi}{2}$. Por supuesto, eso ocurre si y sólo si $\mathbf{u} \cdot \mathbf{v} = 0$.

Se define $O_n = \{T \in GL_n \mid [T]^{-1} = [T]^t\}$. (La O proviene de *orthogonal*, la palabra inglesa que significa ortogonal, ya que tradicionalmente las matrices cuadradas cuya inversa es su transpuesta se llaman ortogonales.) Es fácil verificar que $O_n \leq GL_n$, y a O_n se le llama el grupo ortogonal de dimensión n. Decimos que una función $f : \mathbb{R}^n \to \mathbb{R}^n$ **preserva el producto punto** si y sólo si para cualesquiera $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, $\mathbf{u} \cdot \mathbf{v} = f(\mathbf{u}) \cdot f(\mathbf{v})$.

Proposición D.1. *Sea T un operador lineal sobre $\mathbb{R}^n$. $T \in O_n$ si y sólo si T preserva el producto punto.*

Demostración. Si $T \in O_n$, tenemos, para $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, que $\mathbf{u} \cdot \mathbf{v} = [\mathbf{u}]^t[\mathbf{v}] = [\mathbf{u}]^t I[\mathbf{v}] = [\mathbf{u}]^t[T]^t[T][\mathbf{v}] = ([T][\mathbf{u}])^t[T][\mathbf{v}] = [T\mathbf{u}]^t[T\mathbf{v}] = T\mathbf{u} \cdot T\mathbf{v}$.

Recíprocamente, si T preserva el producto punto, tenemos, para cada $i, j \in \{1, \ldots, n\}$, que $([T]^t[T])_{ij} = [e_i]^t([T]^t[T])^j = [e_i]^t([T]^t[T])[e_j] = ([e_i]^t[T]^t)([T][e_j]) = ([T][e_i])^t([T][e_j]) = [Te_i]^t[Te_j] = [Te_i \cdot Te_j] = [e_i \cdot e_j] = \delta_{ij}$, de donde $[T]^t[T] = I$ (y por ello T es no singular, ya que tiene inversa por la izquierda). Por lo tanto, $[T]^{-1} = [T]^t$. □

La relación entre G_n, GL_n y O_n queda establecida en el resultado que se presenta a continuación.

[3] Así, por supuesto, $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$.

Proposición D.2. *Sea $n \in \mathbb{Z}^+$. Entonces, $O_n = G_n \cap GL_n$.*

Demostración. Si $T \in O_n$, es fácil ver que como preserva el producto punto, preserva normas (es decir, para todo $\mathbf{v} \in \mathbb{R}^n$, $\|Tv\| = \|v\|$), pues si $\mathbf{v} \in \mathbb{R}^n$, ocurre que $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{T\mathbf{v} \cdot T\mathbf{v}} = \|T\mathbf{v}\|$. Entonces, si $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ (y denotamos como d la distancia euclidiana), se tiene que $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \|T(\mathbf{u} - \mathbf{v})\| = \|T\mathbf{u} - T\mathbf{v}\| = d(T\mathbf{u}, T\mathbf{v})$.

Recíprocamente, supongamos que $n \leq 3$, y sea $T \in GL_n$ tal que T preserva distancias. Probemos que T preserva el producto punto. Sean pues $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Si $\mathbf{u} = \mathbf{0}$ o $\mathbf{v} = \mathbf{0}$, tenemos que $\mathbf{u} \cdot \mathbf{v} = 0 = T\mathbf{u} \cdot T\mathbf{v}$. Podemos entonces suponer que $\mathbf{u} \neq \mathbf{0} \neq \mathbf{v}$. Con esto en mente, consideremos el triángulo D, de lados a, b, c, donde el lado a es el vector $\mathbf{u}$, el lado b es el vector $\mathbf{v}$, y θ es el ángulo comprendido entre los lados a y b. Consideremos aparte otro triángulo, D', de lados a', b', c', donde el lado a' es el vector $T\mathbf{u}$, el lado b' es el vector $T\mathbf{v}$, y θ' es el ángulo comprendido entre los lados a' y b'. Como T preserva distancias, claramente preserva normas, y por lo tanto (si denotamos con lg la longitud) $\lg(a) = \|\mathbf{u}\| = \|T\mathbf{u}\| = \lg(a')$ y $\lg(b) = \|\mathbf{v}\| = \|T\mathbf{v}\| = \lg(b')$. Además tenemos que $\lg(c) = \|\mathbf{u} - \mathbf{v}\| = d(\mathbf{u}, \mathbf{v}) = d(T\mathbf{u}, T\mathbf{v}) = \|T\mathbf{u} - T\mathbf{v}\| = \lg(c')$. Por tanto, los triángulos D y D' son congruentes, y entonces $\theta = \theta'$. Tenemos ahora que $\mathbf{u} \cdot \mathbf{v} = \cos\theta \|\mathbf{u}\| \|\mathbf{v}\| = \cos\theta' \|T\mathbf{u}\| \|T\mathbf{v}\| = T\mathbf{u} \cdot T\mathbf{v}$.

Para $n \leq 3$, la prueba está completa. Para el caso general, es decir, $n \in \mathbb{Z}^+$, debemos tomar $T \in GL_n$ tal que T preserva distancias y hacer ver que T preserva el producto punto. Podemos apelar al hecho de que, para $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, $\mathbf{u} \cdot \mathbf{v} = \frac{1}{4}\|\mathbf{u} + \mathbf{v}\|^2 - \frac{1}{4}\|\mathbf{u} - \mathbf{v}\|^2$. Esta igualdad es una de las llamadas *identidades de polarización*, y es muy fácil de probar (usando que el producto punto es lineal en cada entrada). Ahora, como T preserva distancias, T preserva normas. Esto, junto con la linealidad de T y la identidad de polarización descrita, basta para obtener que T preserva el producto punto. □

Fácilmente, si T es un operador lineal sobre $\mathbb{R}^n$ que preserva el producto punto, T preserva ángulos, por la relación $\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}$ (y por que la función coseno es inyectiva en $[0, \pi]$).

Por supuesto, si $T \in O_n$, entonces para $i, j \in \{1, \ldots, n\}$ se tiene que $Te_i \cdot Te_j = e_i \cdot e_j = \delta_{ij}$. Recíprocamente, si T es un operador lineal sobre $\mathbb{R}^n$ tal que para cualesquiera $i, j \in \{1, \ldots, n\}$ se tiene que $Te_i \cdot Te_j = \delta_{ij}$, probemos que $T \in O_n$. Tenemos que, para cualesquiera $i, j \in \{1, \ldots, n\}$, $([T]^t[T])_{ij} = ([T]^t)_i[T]^j = ([T]^i)^t[T]^j = [Te_i]^t[Te_j] = \delta_{ij}$, por lo que $[T]^t[T] = I$ (y por ello T es no singular, ya que tiene inversa por la izquierda). Por lo tanto, $T \in O_n$.

Antes observamos que si $T \in O_n$, entonces $T \restriction_{\mathbb{S}^{n-1}}$ (la restricción de T a $\mathbb{S}^{n-1}$) es una isometría de $\mathbb{S}^{n-1}$. Probemos ahora que si α es una isometría de $\mathbb{S}^{n-1}$, hay una única $T \in O_n$ tal que $\alpha = T \restriction_{\mathbb{S}^{n-1}}$. Supongamos primero que $n \leq 3$. Como $\alpha : \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}$, tenemos que, para toda $i \in \{1, \ldots, n\}$, $\sqrt{\alpha(e_i) \cdot \alpha(e_i)} = \|\alpha(e_i)\| = 1$, por lo que $\alpha(e_i) \cdot \alpha(e_i) = 1$. Sean $i, j \in \{1, \ldots, n\}$, $i \neq j$. Como $\|e_i - e_j\| = d(e_i, e_j) = d(\alpha(e_i), \alpha(e_j)) = \|\alpha(e_i) - \alpha(e_j)\|$, el triángulo formado por los vectores e_i y e_j es congruente con el triángulo formado por los vectores $\alpha(e_i)$ y $\alpha(e_j)$, pues cada uno de los tres lados del primero mide lo mismo que el lado correspondiente del segundo. Entonces, como e_i y e_j son perpendiculares, $\alpha(e_i)$ y $\alpha(e_j)$ también lo son, y eso es equivalente a que $\alpha(e_i) \cdot \alpha(e_j) = 0$. Por lo tanto, para cualesquiera $i, j \in \{1, \ldots, n\}$, $\alpha(e_i) \cdot \alpha(e_j) = \delta_{ij}$. Sea T el único operador lineal sobre $\mathbb{R}^n$ tal que $Te_i = \alpha(e_i)$ para cada $i \in \{1, \ldots, n\}$. En virtud de lo visto en el párrafo anterior, $T \in O_n$. Falta probar que $\alpha = T \restriction_{\mathbb{S}^{n-1}}$, pero esto se sigue del hecho de que ambas son isometrías de $\mathbb{S}^{n-1}$ que coinciden en la base canónica. Para $n \in \mathbb{Z}^+$, sean $i, j \in \{1, \ldots, n\}$. Tenemos que $d(\alpha(-e_j), \alpha(e_j)) = d(-e_j, e_j) = 2$. Cortando a $\mathbb{R}^n$ con algún plano que pase por $\alpha(e_j)$, por $-\alpha(e_j)$ y por $\alpha(-e_j)$, es muy fácil convencerse de que $\alpha(-e_j) = -\alpha(e_j)$. Utilizando otra vez la identidad de polarización que empleamos en la prueba de D.2, tenemos lo siguiente:

$$\begin{aligned}
\alpha(e_i) \cdot \alpha(e_j) &= \frac{1}{4}\|\alpha(e_i) + \alpha(e_j)\|^2 - \frac{1}{4}\|\alpha(e_i) - \alpha(e_j)\|^2 \\
&= \frac{1}{4}\|\alpha(e_i) - (-\alpha(e_j))\|^2 - \frac{1}{4}\|\alpha(e_i) - \alpha(e_j)\|^2 \\
&= \frac{1}{4}\|\alpha(e_i) - \alpha(-e_j)\|^2 - \frac{1}{4}\|\alpha(e_i) - \alpha(e_j)\|^2 \\
&= \frac{1}{4}d(\alpha(e_i), \alpha(-e_j))^2 - \frac{1}{4}d(\alpha(e_i), \alpha(e_j))^2 \\
&= \frac{1}{4}d(e_i, -e_j)^2 - \frac{1}{4}d(e_i, e_j)^2 \\
&= \frac{1}{4}\|e_i - (-e_j)\|^2 - \frac{1}{4}\|e_i - e_j\|^2 \\
&= \frac{1}{4}\|e_i + e_j)\|^2 - \frac{1}{4}\|e_i - e_j\|^2 \\
&= e_i \cdot e_j \\
&= \delta_{ij}
\end{aligned}$$

Basta entonces extender a α hasta T, el único operador lineal sobre $\mathbb{R}^n$ tal que $Te_i = \alpha(e_i)$ para cada $i \in \{1, \ldots, n\}$. Como antes, $T \in O_n$ y $\alpha = T \restriction_{\mathbb{S}^{n-1}}$.

Acabamos de probar el siguiente resultado.

Proposición D.3. *Para $n \in \mathbb{Z}^+$, las restricciones a $\mathbb{S}^{n-1}$ de los elementos de O_n son precisamente las isometrías de $\mathbb{S}^{n-1}$.*

Para $T \in O_n$, se cumple que $|[T]|^2 = |[T]||[T]| = |[T]||[T]^t| = |[T][T]^t| = |I| = 1$, y por lo tanto que $|[T]| = \pm 1$. Se define $SO_n = \{T \in O_n \mid |[T]| = 1\}$. (La S proviene de *special*, la palabra inglesa que significa especial, ya que a las matrices cuadradas de determinante 1 se les llama especiales.) Claramente, $SO_n \leq O_n$.

El siguiente resultado no nos interesa para el desarrollo de este trabajo, pero se incluye por su valor estético y por que proporciona algo de perspectiva.

Teorema D.4. *$SO_2 = \{\alpha \mid \alpha$ es una rotación de $\mathbb{R}^2$ en torno al origen$\}$.*

Demostración. Afirmamos lo siguiente:

$$SO_2 = \{T \in GL_2 \mid [T] = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \theta \in \mathbb{R}\}$$

Probemos la afirmación. Sea $T \in SO_2$. Pongamos que $[T] = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Como $\begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = [T]^{-1} = [T]^t = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, tenemos que $c = -b$ y que $d = a$, por lo que $[T] = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. Además tenemos que $a^2 + b^2 = |[T]| = 1$, por lo que (a, b) satisface la ecuación $x^2 + y^2 = 1$, es decir, $(a, b) \in \mathbb{S}^1$. De ahí, $(a, b) = (\cos\theta, \sin\theta)$ para algún $\theta \in \mathbb{R}$ y terminamos.

Ahora, si $[T] = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, tenemos lo siguiente:

$$\begin{aligned} [T][T]^t &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \cos^2\theta + \sin^2\theta & \cos\theta\sin\theta - \cos\theta\sin\theta \\ \cos\theta\sin\theta - \cos\theta\sin\theta & \cos^2\theta + \sin^2\theta \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= I \end{aligned}$$

Además, $|[T]| = \cos^2\theta + \sin^2\theta = 1$. En vista de esto, T es no singular, y entonces $[T]^{-1} = [T]^t$.

Obtengamos ahora el teorema a partir de la afirmación. Tenemos que T es una rotación de $\mathbb{R}^2$ por el origen si y sólo si ocurre que $T \in GL_2$ y que $Te_1 = (\cos\theta, \sin\theta)$ y $Te_2 = (\cos(\theta + \frac{\pi}{2}), \sin(\theta + \frac{\pi}{2}))$, para algún $\theta \in \mathbb{R}$. Además, sabemos que, para cualquier $\theta \in \mathbb{R}$, $(\cos(\theta + \frac{\pi}{2}), \sin(\theta + \frac{\pi}{2})) = (-\sin\theta, \cos\theta)$. □

El caso de $\mathbb{R}^3$ es más delicado. Se puede ver un tratamiento riguroso de los elementos de geometría analitica que necesitamos en [E1, Caps. 9, 10].

Para $\mathbf{u}, \mathbf{v} \in \mathbf{R}^3$, digamos $\mathbf{u} = (x_1, y_1, z_1)$ y $\mathbf{v} = (x_2, y_2, z_2)$, se define el **producto vectorial** de $\mathbf{u}$ y de $\mathbf{v}$, que es un vector de $\mathbf{R}^3$, como $\mathbf{u} \times \mathbf{v} = (y_1z_2 - z_1y_2, z_1x_2 - x_1z_2, x_1y_2 - y_1x_2)$. Hay una regla sencilla para recordar esto: $\mathbf{u} \times \mathbf{v} = \begin{vmatrix} e_1 & e_2 & e_3 \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix}$. Tal cosa es una mera mnemotecnia; formalmente no tiene sentido. Se tiene que, si $\mathbf{u} = \mathbf{0}$, si $\mathbf{v} = \mathbf{0}$ o si $\mathbf{u}$ y $\mathbf{v}$ son paralelos, entonces $\mathbf{u} \times \mathbf{v} = \mathbf{0}$, y que si $\mathbf{u} \neq \mathbf{0} \neq \mathbf{v}$ y si $\mathbf{u}$ y $\mathbf{v}$ no son paralelos, entonces $\mathbf{u} \times \mathbf{v}$ es un vector distinto de $\mathbf{0}$ perpendicular a $\mathbf{u}$ y a $\mathbf{v}$, y que cumple que, si un observador se coloca sobre el punto final de $\mathbf{u} \times \mathbf{v}$ mirando hacia $\mathbf{0}$, él ve la rotación que lleva a $\mathbf{u}$ hasta $\mathbf{v}$ como una rotación en sentido contrario a las manecillas del reloj. Si $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$, tiene sentido hablar del triple producto $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$.

Proposición D.5. *Sean $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$. Los tres vectores son coplanares si y sólo si $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = 0$.*

Demostración. Tratemos primero el caso de que alguno de los tres vectores sea $\mathbf{0}$. Entonces, se verifica que $\mathbf{u}$, $\mathbf{v}$, y $\mathbf{w}$ son coplanares y que $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = 0$. Podemos entonces suponer que los tres vectores son distintos de $\mathbf{0}$.

Deshagámonos ahora del caso en que $\mathbf{v}$ y $\mathbf{w}$ son paralelos. Entonces, otra vez se cumple que $\mathbf{u}$, $\mathbf{v}$, y $\mathbf{w}$ son coplanares y que $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = 0$ (pues $\mathbf{v} \times \mathbf{w} = \mathbf{0}$). Podemos ahora suponer que $\mathbf{v}$ y $\mathbf{w}$ no son paralelos.

Sabemos que $\mathbf{v} \times \mathbf{w}$ es un vector normal al plano que contiene a $\mathbf{v}$ y a $\mathbf{w}$. Entonces, claramente, $\mathbf{u}$ está contenido en ese plano si y sólo si $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = 0$. □

Sean $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ vectores no coplanares. La **orientación** de la terna ordenada $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ se define como el signo de $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$. La idea geométrica subyacente es que, si identificamos a $\mathbf{u}$ con el dedo pulgar, a $\mathbf{v}$ con el índice y a $\mathbf{w}$ con el cordial, entonces tiene sentido preguntarnos si podemos representar la terna ordenada con la mano izquierda o con la derecha. Convenzámonos de que podemos usar la mano derecha si y sólo si la orientación de la terna ordenada es positiva. Como los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ no

son coplanares, podemos estar seguros de que ninguno de los tres es $\mathbf{0}$, de que $\mathbf{v}$ y $\mathbf{w}$ no son paralelos, y de que $\mathbf{u}$ no está en el plano que contiene a $\mathbf{v}$ y a $\mathbf{w}$. Sabemos que podemos representar a la terna ordenada $(\mathbf{v}, \mathbf{w}, \mathbf{v} \times \mathbf{w})$ con la mano derecha. Ahora, como $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \cos\theta \|\mathbf{u}\| \|\mathbf{v} \times \mathbf{w}\|$, donde θ es el menor ángulo que se forma entre los vectores $\mathbf{u}$ y $\mathbf{v} \times \mathbf{w}$, tenemos que $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ tiene el mismo signo que $\cos\theta$. Observemos que $\theta \in [0, \pi] \backslash \{\frac{\pi}{2}\}$, y que por ello $\cos\theta > 0$ si y sólo si θ es un ángulo agudo.

Equivalentemente, la orientación de la terna ordenada $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ es positiva si y sólo si, situándose un observador en el interior del ángulo sólido determinado por $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ de cara a $\mathbf{0}$, si voltea a ver primero a $\mathbf{u}$, luego a $\mathbf{v}$, y después a $\mathbf{w}$, su mirada se mueve en sentido contrario a las manecillas del reloj.

Por ejemplo, la terna ordenada (e_1, e_2, e_3) tiene orientación positiva.

Antes de volver a las transformaciones lineales, recordemos que, para cualesquiera $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^3$, es de rutina probar que $\mathbf{v}_1 \cdot (\mathbf{v}_2 \times \mathbf{v}_3) = |M_{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3}|$, donde $M_{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3} \in \mathcal{M}_3(\mathbb{R})$ es tal que, para cada $i \in \{1, 2, 3\}$, $(M_{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3})_i = [v_i]^t$. Es decir, si $\mathbf{v}_1 = (x_1, y_1, z_1)$, $\mathbf{v}_2 = (x_2, y_2, z_2)$ y $\mathbf{v}_3 = (x_3, y_3, z_3)$, entonces $\mathbf{v}_1 \cdot (\mathbf{v}_2 \times \mathbf{v}_3) = |M_{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3}| = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$.

Decimos que una función $f : \mathbb{R}^3 \to \mathbb{R}^3$ **preserva orientaciones** si y sólo si ocurre que para cualesquiera $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$, $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ es una terna ordenada de vectores no coplanares con orientación positiva si y sólo si $(f(\mathbf{u}), f(\mathbf{v}), f(\mathbf{w}))$ es una terna ordenada de vectores no coplanares con orientación positiva.

Proposición D.6. *Sea T un operador lineal sobre $\mathbb{R}^3$. Entonces, $|[T]| > 0$ si y sólo si T preserva orientaciones.*

Demostración. Sean $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^3$. Para cada $i \in \{1, 2, 3\}$, tenemos que:

$$\begin{aligned} [T]((M_{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3})^t)^i &= [T]((M_{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3})_i)^t \\ &= [T]([\mathbf{v}_i]^t)^t = [T][\mathbf{v}_i] \\ &= [T\mathbf{v}_i] = ([T\mathbf{v}_i]^t)^t \\ &= ((M_{T\mathbf{v}_1,T\mathbf{v}_2,T\mathbf{v}_3})_i)^t \\ &= ((M_{T\mathbf{v}_1,T\mathbf{v}_2,T\mathbf{v}_3})^t)^i \end{aligned}$$

Por lo tanto, por la manera como se multiplican las matrices, se tiene que

$[T](M_{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3})^t = (M_{T\mathbf{v}_1,T\mathbf{v}_2,T\mathbf{v}_3})^t$. Entonces, ocurre que:

$$\begin{aligned} |[T]|(\mathbf{v}_1 \cdot (\mathbf{v}_2 \times \mathbf{v}_3)) &= |[T]||M_{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3}| \\ &= |[T]||(M_{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3})^t| \\ &= |[T](M_{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3})^t| \\ &= |(M_{T\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3})^t| \\ &= |M_{T\mathbf{v}_1,T\mathbf{v}_2,T\mathbf{v}_3}| \\ &= T\mathbf{v}_1 \cdot (T\mathbf{v}_2 \times T\mathbf{v}_3) \end{aligned}$$

De aquí se desprenden las siguientes observaciones:

(i) Si $|[T]| > 0$, entonces para cualesquiera $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$, $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) > 0$ si y sólo si $T\mathbf{u} \cdot (T\mathbf{v} \times T\mathbf{w}) > 0$ (es decir, T preserva orientaciones).

(ii) Si $|[T]| \leq 0$, entonces $\mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3) > 0$ pero $Te_1 \cdot (Te_2 \times Te_3) \leq 0$, por lo que T no preserva orientaciones.

Combinando (i) y (ii), tenemos el resultado. □

Sea T un operador lineal sobre $\mathbb{R}^3$. Si T preserva orientaciones, entonces por supuesto (Te_1, Te_2, Te_3) es una terna ordenada de vectores no coplanares con orientación positiva. Recíprocamente, si $Te_1 \cdot (Te_2 \times Te_3)$ es una terna ordenada de vectores no coplanares con orientación positiva, entonces $|[T]| = |(M_{Te_1,Te_2,Te_3})^t| = |M_{Te_1,Te_2,Te_3}| = Te_1 \cdot (Te_2 \times Te_3) > 0$. Así, basta verificar que T preserve la orientación de la base canónica para tener que preserva orientaciones.

Ya probamos que si $T \in O_3$, entonces $|[T]| = \pm 1$. A la luz de esto, tenemos que $SO_3 = \{T \in O_3 \mid |[T]| > 0\} = \{T \in O_3 \mid T$ preserva orientaciones$\}$. Como $O_3 = G_3 \cap GL_3$, tenemos la siguiente afirmación:

Proposición D.7. *$SO_3 = \{T \in GL_3 \mid T$ preserva distancias y orientaciones$\}$.*

Se puede ver una demostración algo diferente de este último resultado en [E2, Cap. III, Núm. 103].[4]

Concluyamos este apéndice.

Teorema D.8. *$SO_3 = \{\rho \mid \rho$ es una rotación de $\mathbb{R}^3$ cuyo eje pasa por el origen$\}$.*

[4]Lo que ahí se demuestra es que $T \in SO_3$ si y sólo si T es un operador lineal sobre $\mathbb{R}^3$ que transforma la base canónica en una terna ordenada de vectores de norma 1, perpendiculares dos a dos y con orientación positiva. Con lo que aquí probamos, eso basta.

Demostración. Es evidente que toda rotación de $\mathbb{R}^3$ cuyo eje pase por el origen es un operador lineal no singular sobre $\mathbb{R}^3$ que preserva distancias y orientaciones, y entonces, por el resultado anterior, es un elemento de SO_3. El hecho de que toda isometría de $\mathbb{R}^3$ que sea también un operador lineal sobre $\mathbb{R}^3$ y que preserve orientaciones sea una rotación de $\mathbb{R}^3$ cuyo eje pase por el origen es razonablemente claro. Una demostración formal, que utiliza herramientas de la geometría del espacio, se puede encontrar en [M, Teo. 16.18].[5] □

En general, para $n \in \mathbb{Z}^+$, por analogía con $\mathbb{R}^2$ y con $\mathbb{R}^3$, los elementos de SO_n son a veces llamados rotaciones, y a SO_n se le puede llamar el grupo de rotaciones de la esfera $\mathbb{S}^{n-1}$.

[5]Observando que todo operador lineal de $\mathbb{R}^3$ deja fijo a $\mathbf{0}$.

Apéndice E

Cardinalidad sin Elección

El objetivo de este apéndice es recuperar algunos resultados (básicos) de la teoría de la cardinalidad sin recurrir al Axioma de Elección. Estamos trabajando con la convención de que un conjunto A es **contable** si y sólo si existe $f : A \to \mathbb{N}$ inyectiva.[1] Esto, por supuesto, no implica que tal función se pueda dar explícitamente. Por ejemplo, el Axioma de Eleccion garantiza que la unión de una cantidad contable de conjuntos contables es contable, pero no da la regla de correspondencia de una función inyectiva que atestigüe esa afirmación. (Comúnmente decimos que el Axioma de Elección dice que tal función existe, pero no dice cuál es.)

Lema E.1. *Sea I un conjunto, y sea $\{A_i\}_{i\in I}$ una colección de conjuntos. Si se puede dar $f : I \to \mathbb{N}$ inyectiva, y para cada $i \in I$, se puede dar $f_i : A_i \to \mathbb{N}$ inyectiva, entonces se puede dar $g : \bigcup_{i\in I} A_i \to \mathbb{N}$ inyectiva.*

Demostración. Sea $m : \bigcup_{i\in I} A_i \to \mathbb{N}$ tal que $m : x \mapsto \text{mín}\{f(i) \mid x \in A_i\}$. Definimos ahora $h : \bigcup_{i\in I} A_i \to \mathbb{N} \times \mathbb{N}$ tal que $h : x \mapsto (m(x), f_{m(x)}(x))$. Se observa que h es inyectiva. Sea $t : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ tal que $t : (m, n) \mapsto 2^m 3^n$. Por el Teorema Fundamental de la Aritmética, t es inyectiva. Sea $g = t \circ h$. Notemos que $g : \bigcup_{i\in I} A_i \to \mathbb{N}$ y que g es una función inyectiva por ser una composición de funciones inyectivas. □

[1] No es difícil hacer ver, sin apelar al Axioma de Elección, que un conjuto es contable si y sólo si es finito o numerable (donde numerable significa biyectable con $\mathbb{N}$).

El lema anterior da condiciones suficientes para que, sin utilizar el Axioma de Elección, podamos demostrar que una unión contable de conjuntos contables es contable. Ahora apliquémoslo.

Si S es un conjunto de símbolos, llamemos $\mathbb{E}_S$ al conjunto de todas las S-palabras, es decir, sucesiones finitas de símbolos de S.

Proposición E.2. *Sea S un conjunto de símbolos. Si se puede dar $f : S \to \mathbb{N}$ inyectiva, entonces se puede dar $g : \mathbb{E}_S \to \mathbb{N}$ inyectiva.*

Demostración. Para cada $n \in \mathbb{N}$, definimos $W_n = \{\alpha \in \mathbb{E}_S \mid \lg(\alpha) = n\}$. Observemos que $W_0 = \{\Lambda\}$, donde Λ denota la palabra vacía. Numeramos a los números primos con $\mathbb{Z}^+$ de manera creciente, es decir, $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, etc. Ahora, sea $f_0 : W_0 \to \mathbb{N}$ tal que $f_0 : \Lambda \mapsto 1$. Trivialmente, f_0 es inyectiva. Para cada $n \in \mathbb{Z}^+$, sea $f_n : W_n \to \mathbb{N}$ tal que $f_n : \rho_1\rho_2 \dots \rho_n \mapsto 2^{f(\rho_1)}3^{f(\rho_2)} \cdots p_n^{f(\rho_n)}$. En virtud del Teorema Fundamental de la Aritmética, f_n es inyectiva para cada $n \in \mathbb{Z}^+$. Ahora, notemos que $\mathbb{E}_S = \bigcup_{n\in\mathbb{N}} W_n$. Entonces, por E.1, se puede dar $g : \mathbb{E}_S \to \mathbb{N}$ inyectiva. □

Corolario E.3. *Sea M un conjunto libre de símbolos. Si se puede dar $f : M \to \mathbb{N}$ inyectiva, entonces se puede dar $g : F_M \to \mathbb{N}$ inyectiva.*

Demostración. La función $r : M \to \{\rho^{-1} \mid \rho \in M\}$ tal que $h : \rho \mapsto \rho^{-1}$ es, claramente, biyectiva. Sea $S = M \cup \{\rho^{-1} \mid \rho \in M\}$. Definimos $h : S \to (\mathbb{N} \times \{0\}) \cup (\mathbb{N} \times \{1\})$ tal que $h : \sigma \to \begin{cases} (f(\sigma), 0) & \text{si } \sigma \in M \\ ((f \circ h^{-1})(\sigma), 1) & \text{si } \sigma \notin M \end{cases}$. Se observa que h es inyectiva. Sea $s : (\mathbb{N} \times \{0\}) \cup (\mathbb{N} \times \{1\}) \to \mathbb{N}$ tal que $s : (n, 0) \mapsto 2n$ y $s : (n, 1) \mapsto 2n + 1$. También se observa que s es inyectiva (de hecho, es biyectiva). Entonces, $s \circ h : S \to \mathbb{N}$ es inyectiva. Por E.2, se puede dar $t : \mathbb{E}_S \to \mathbb{N}$ inyectiva. La inclusión $i : F_M \to \mathbb{E}_S$ es, trivialmente, inyectiva. Por lo tanto, $t \circ i : F_M \to \mathbb{N}$ es inyectiva. □

Decimos que un número complejo (o real) es **algebraico** si y sólo si hay un polinomio no cero en una indeterminada con coeficientes en $\mathbb{Z}$ del cual es raíz. Decimos que un conjunto X es **numerable** si y sólo si hay $f : X \to \mathbb{N}$ biyectiva. No tenemos que apelar al Axioma de Elección para escribir $|X| = \aleph_0$ si y sólo si X es numerable.[2]

Lema E.4. *Hay una cantidad numerable de números complejos algebraicos.*

[2]De entrada, la afirmación $|X| = \aleph_0$ equivale a que X sea bien ordenable y biyectable con $\mathbb{N}$, pero si X es biyectable con $\mathbb{N}$, podemos ordenar a X como $\mathbb{N}$, y ese orden es, por supuesto, un buen orden.

Demostración. Sea $\mathbb{Z}[x]$ el anillo de polinomios en una indeterminada con coeficientes en $\mathbb{Z}$. Demos $g : \mathbb{Z}[x] \to \mathbb{N}$ inyectiva. Primero, sea $r : \mathbb{Z} \to \mathbb{N}$ tal que $r : n \mapsto \begin{cases} 2n & \text{si } n \geq 0 \\ -2n-1 & \text{si } n < 0 \end{cases}$. Se observa que r es inyectiva (de hecho, es biyectiva). Para cada $n \in \mathbb{N}$, definimos $P_n = \{a_0 + a_1x + \cdots + a_nx^n, a_0, \ldots, a_n \in \mathbb{Z}\}$. Numeramos a los números primos con $\mathbb{N}$ de manera creciente, es decir, $p_0 = 2$, $p_1 = 3$, $p_2 = 5$, etc. Para cada $n \in \mathbb{N}$, sea $f_n : P_n \to \mathbb{N}$ tal que $f_n : a_0 + a_1x + \cdots + a_nx^n \mapsto 2^{r(a_0)}3^{r(a_1)} \cdots p_n^{r(a_n)}$. En virtud del Teorema Fundamental de la Aritmética, f_n es inyectiva para cada $n \in \mathbb{N}$. Ahora, $\mathbb{Z}[x] = \bigcup_{n\in\mathbb{N}} P_n$. Entonces, por E.1, se puede dar $g : \mathbb{Z}[x] \to \mathbb{N}$ inyectiva. (Como, por supuesto, la inclusión $i : \mathbb{N} \to \mathbb{Z}[x]$ es inyectiva, el Teorema de Schröder-Bernstein (que es una consecuencia de la prueba de 2.8) da que $|\mathbb{Z}[x]| = \aleph_0$.)

Ahora, sea $\mathbb{A} = \{z \in \mathbb{C} \mid z \text{ es algebraico}\}$. Sea $f \in \mathbb{Z}[x]\backslash\{0\}$. Definimos $R_f = \{z \in \mathbb{C} \mid f(z) = 0\}$. Como $f \neq 0$, f tiene una cantidad finita de raíces complejas. Como podemos ordenar a $\mathbb{C}$ lexicográficamente (ver antes de 3.2), y este orden es estricto y total, se puede dar $g_f : R_f \to \mathbb{N}$ inyectiva. En vista de que $\mathbb{A} = \bigcup_{f\in\mathbb{Z}[x]\backslash\{0\}} R_f$ (y de que como g es inyectiva, $g \restriction_{\mathbb{Z}[x]\backslash\{0\}}$ -la restricción- también lo es), E.1 da $h : \mathbb{A} \to \mathbb{N}$ inyectiva. Otra vez, la inclusión $i : \mathbb{N} \to \mathbb{A}$ es inyectiva, así que el Teorema de Schröder-Bernstein garantiza que $|\mathbb{A}| = \aleph_0$. □

De manera muy similiar, se puede probar que hay una cantidad numerable de números reales algebraicos (la única diferencia es que el orden usual de $\mathbb{R}$, que es estricto y total, se puede usar).

Para hacer concisa la demostración del siguiente resultado, es conveniente adoptar un poco de notación. Para X, Y conjuntos, escribiremos $X \preccurlyeq Y$ si y sólo si hay una función inyectiva de X en Y. (Habitualmente, $X \preccurlyeq Y$ se lee "X es *dominado* por Y".)

Lema E.5. *Sean A, B conjuntos tales que A es contable y B es biyectable con $\mathbb{R}$. Entonces, hay $x \in B\backslash A$.*

Demostración. Supongamos que no. Entonces, $B \subseteq A$. Tenemos que $A \preccurlyeq \mathbb{N} \preccurlyeq \mathbb{R} \preccurlyeq B \preccurlyeq A$, y por el Teorema de Schröder-Bernstein (que es una consecuencia de la prueba de 2.8), que $\mathbb{N}$ es biyectable con $\mathbb{R}$, cosa que contradice un famoso teorema de Cantor. □

Apéndice F

La Paradoja de Sierpiński-Mazurkiewicz

En este apéndice daremos una construcción paradójica que no depende del Axioma de Elección. Estamos trabajando con la convención de que, para $n \in \mathbb{Z}^+$, $\mathbb{S}^{n-1}$ es la esfera unitaria centrada en el origen de $\mathbb{R}^n$. Decimos que un conjunto X es **numerable** si y sólo si hay $f : X \to \mathbb{N}$ biyectiva y en tal caso, a la luz de lo explicado antes de E.4, podemos escribir $|X| = \aleph_0$. Decimos que un número complejo es **algebraico** si y sólo si hay un polinomio no cero en una indeterminada con coeficientes en $\mathbb{Z}$ del cual es raíz. En E.4 se prueba, sin usar el Axioma de Elección, que hay una cantidad numerable de números complejos algebraicos. Decimos que un número complejo es **trascendente** si y sólo si no es algebraico.

Teorema F.1 (Paradoja de Sierpiński-Mazurkiewicz)**.** *Existe un subconjunto de $\mathbb{R}^2$ que es no vacío y G_2-paradójico.*

Demostración. Identifiquemos a los puntos de $\mathbb{R}^2$ con los números complejos. Por supuesto, $\mathbb{S}^1 = \{z \in \mathbb{C} \mid |z| = 1\}$. Como en $\mathbb{S}^1$ hay una cantidad contable de números algebraicos (pues $\mathbb{A} \cap \mathbb{S}^1 \subseteq \mathbb{A}$) y $\mathbb{S}^1$ es biyectable con $\mathbb{R}$, por E.5 hay $c \in \mathbb{S}^1$ trascendente (de hecho, si apelamos al Teorema de Hermite-Lindemann, tenemos que e^i es trascendente). Sea $E = \{z \in \mathbb{C} \mid z = a_0 + a_1 c + \cdots + a_n c^n,\ n \in \mathbb{N},\ a_0, \ldots, a_n \in \mathbb{N}\}$. Cada elemento de E tiene una escritura única de la forma descrita, ya que si hubiera $z \in E$ con $a_0 + a_1 c + \cdots + a_n c^n = z = b_0 + b_1 c + \cdots + b_m c^m$, suponiendo sin pérdida de generalidad que $m \leq n$ y conveniendo en que, para cada i tal que $m < i \leq n$, $b_i = 0$, tendríamos $(a_0 - b_0) + (a_1 - b_1)c + \cdots + (a_n - b_n)c^n = 0$, cosa que,

por ser c trascendente, obliga a que $a_i = b_i$ para cada $i \in \{1, \ldots, n\}$. Sean $\tau, \rho : \mathbb{C} \to \mathbb{C}$ tales que $\tau : z \mapsto z + 1$ y $\rho : z \mapsto cz$. Como τ es una traslación y ρ es una rotación (pues $|c| = 1$), tenemos que $\tau, \rho \in G_2$. Ahora, sean $A = \{z \in E \mid z = a_0 + a_1 c + \cdots + a_n c^n,\ n \in \mathbb{N},\ a_0 = 0,\ a_1, \ldots, a_n \in \mathbb{N}\}$ y $B = \{z \in E \mid z = a_0 + a_1 c + \cdots + a_n c^n,\ n \in \mathbb{N},\ a_0 \in \mathbb{Z}^+,\ a_1, \ldots, a_n \in \mathbb{N}\}$. Por supuesto, $A, B \subseteq E$. En virtud de la unicidad de la representación de los elementos de E, tenemos que $A \cap B = \emptyset$. Como $\rho E = A$ y $\tau E = B$, tenemos que $\rho^{-1} A = E = \tau^{-1} B$. Por lo tanto, E es G_2-paradójico. □

Probemos ahora, también sin recurrir al Axioma de Elección, que el conjunto E construido en la demostración anterior es numerable. Usaremos una técnica que ya antes hemos aplicado.

Para cada $n \in \mathbb{N}$, definimos $E_n = \{a_0 + a_1 c + \cdots + a_n c^n,\ a_0, \ldots, a_n \in \mathbb{N}\}$. Numeramos a los números primos con $\mathbb{N}$ de manera creciente, es decir, $p_0 = 2$, $p_1 = 3$, $p_2 = 5$, etc. Para cada $n \in \mathbb{N}$, sea $f_n : E_n \to \mathbb{N}$ tal que $f_n : a_0 + a_1 c + \cdots + a_n c^n \mapsto 2^{a_0} 3^{a_1} \cdots p_n^{a_n}$. En virtud del Teorema Fundamental de la Aritmética, y de la unicidad de la representación de los elementos de E, f_n es inyectiva para cada $n \in \mathbb{N}$. Ahora, $E = \bigcup_{n \in \mathbb{N}} E_n$. Entonces, por E.1, hay $g : E \to \mathbb{N}$ inyectiva. Claramente, la inclusión $i : \mathbb{N} \to E$ es inyectiva, por lo que, por el Teorema de Schröder-Bernstein, $|E| = \aleph_0$.

La existencia de este conjunto no es un hecho abiertamente contradictorio. Finalmente, en términos de medida de Lebesgue, los conjuntos numerables tienen medida cero, así que la G_2-paradojicidad de E sólo implica que $2 \cdot 0 = 0$. Sin embargo, en dimensiones superiores, las cosas cambian.

Apéndice G

Una Consecuencia Importante

Probaremos ahora que, como consecuencia inmediata de la Paradoja de Banach-Tarski, $\mathbb{R}^3$ no admite un cierto tipo de medida.

Definición G.1. *Sea X un conjunto. Una* ***medida finitamente aditiva*** *sobre X es una función μ que cumple que*

(i) $\mu : \mathscr{P}(X) \to \mathbb{R}^+ \cup \{0, +\infty\}$, *y*

(ii) para cualesquiera $n \in \mathbb{Z}^+$ y $A_1, \ldots, A_n \subseteq X$ ajenos dos a dos, ocurre que $\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i)$.

Si μ es una medida finitamente aditiva sobre un conjunto X y $A \subseteq X$, decimos que μ **normaliza** a A si y sólo si $\mu(A) = 1$.

Si μ es una medida finitamente aditiva sobre un conjunto X y un grupo G actúa sobre X, decimos que μ es ***G*-invariante** si y sólo si para cualesquiera $g \in G$ y $A \subseteq \mathbb{X}$, ocurre que $\mu(gA) = \mu(A)$.

Llamemos $\mathscr{U}$ al cubo unitario en $\mathbb{R}^3$ (es decir, $\mathscr{U} = [0,1] \times [0,1] \times [0,1]$).

Corolario G.2 (AE)**.** *No existe una medida finitamente aditiva sobre $\mathbb{R}^3$ que sea G_3-invariante y que normalice a $\mathscr{U}$.*

Demostración. Supongamos que μ es una medida finitamente aditiva G_3-invariante sobre $\mathbb{R}^3$ que normaliza a $\mathscr{U}$. Sea τ una traslación de $\mathbb{R}^3$ tal que $\mathscr{U} \cap \tau\mathscr{U} = \emptyset$. En virtud de 4.6, $\mathscr{U} \sim_{G_3} \mathscr{U} \cup \tau\mathscr{U}$. Entonces, tenemos, para

algún $n \in \mathbb{Z}^+$ una partición de $\mathscr{U}$, $\{A_i\}_{i=1}^n$, una partición de $\mathscr{U} \cup \tau\mathscr{U}$, $\{B_i\}_{i=1}^n$, y $\{\alpha_i\}_{i=1}^n \subseteq G_3$ tales que, para cada $i \in \{1, \ldots, n\}$, $B_i = \alpha_i A_i$. Ocurre lo siguiente:

$$\begin{aligned}
1 &= \mu(\mathscr{U}) \\
&= \mu(\bigcup_{i=1}^{n} A_i) \\
&= \sum_{i=1}^{n} \mu(A_i) \\
&= \sum_{i=1}^{n} \mu(\alpha_i A_i) \\
&= \sum_{i=1}^{n} \mu(B_i) \\
&= \mu(\bigcup_{i=1}^{n} B_i) \\
&= \mu(\mathscr{U} \cup \tau\mathscr{U}) \\
&= \mu(\mathscr{U}) + \mu(\tau\mathscr{U}) \\
&= \mu(\mathscr{U}) + \mu(\mathscr{U}) \\
&= 1 + 1 \\
&= 2
\end{aligned}$$

Eso es una contradicción. Por lo tanto, tal μ no puede existir. □

En particular, y a la luz de que la medida de Lebesgue, definida en el álgebra de subconjuntos Lebesgue-medibles de $\mathbb{R}^3$, es finitamente aditiva, G_3-invariante y normaliza a $\mathscr{U}$, $\mathbb{R}^3$ no admite una medida finitamente aditiva y G_3-invariante que extienda a la medida de Lebesgue.

Se sigue de la prueba de G.2 (y del hecho de que en el paso de la forma básica a la forma fuerte de la Paradoja de Banach-Tarski no se utilizó que estuviéramos trabajando en tres dimensiones) que si $n \in \mathbb{Z}^+$, la existencia de un resultado análogo a la Paradoja de Banach-Tarski en $\mathbb{R}^n$ es incompatible con la existencia de una medida finitamente aditiva y G_n-invariante sobre $\mathbb{R}^n$ que normalice a $[0, 1]^n$.

Apéndice H

El Problema de los Prisioneros

Se incluirá aquí la demostración de otra consecuencia contraintuitiva del Axioma de Elección. Se trata de una extensión de un acertijo clásico.

Revisemos primero el problema original. Supongamos que 100 prisioneros se forman en fila, mirando todos en la misma dirección, de manera que cada uno puede ver a todos los que están delante de él en la fila. Un guardia coloca entonces un gorro, que puede ser blanco o negro, sobre la cabeza de cada prisionero. Ahora, empezando desde atrás, el guardia le pregunta a cada prisionero de qué color es su propio gorro (es decir, le pregunta primero al que ve a los otros 99). Los que acierten quedan libres. Cada uno escucha las respuestas de los anteriores, y se entera de si fueron correctas o no. Si los prisioneros tienen la oportunidad de ponerse de acuerdo de antemano, ¿cuál sería la mejor estrategia?

Antes de resolver el caso finito, plantearemos el problema que nos interesa. Supongamos ahora que la fila es de tamaño $\aleph_0$, y que los prisioneros están formados como $\mathbb{N}$ y numerados de la forma $P_0, P_1, P_2, \ldots$, mirando todos en la dirección positiva (es decir, cada uno puede ver una infinidad de prisioneros). Como antes, se coloca un gorro sobre la cabeza de cada uno, y luego se le pregunta a cada uno, comenzando por P_0, de qué color es su gorro. Sin embargo, ahora cada prisionero *no* puede escuchar qué han dicho los anteriores *ni* saber si han acertado o no. (Lo que sí sabe cada prisionero es qué lugar ocupa él en la fila.) En esta situación, ¿cuál sería la mejor estrategia? Observemos que, de entrada, como cada prisionero no recibe información alguna acerca del color de su gorro, parece que no se

puede planear una estrategia (pues parece que cada uno está adivinando a ciegas). Sorprendentemente, aceptando el Axioma de Elección, hay una manera de garantizar que todos salvo una cantidad finita queden libres.

Revisemos ahora la solución del problema clásico. El primer prisionero al que le preguntan puede decir el color del gorro del segundo. El segundo prisionero acertaría, pero entoces el tercero estaría en la misma situación que el primero. Repetir esta idea 50 veces garantiza la libertad de 50 prisioneros y da una probabilidad de 75 % de que un prisionero cualquiera quede libre. Pero esa estrategia no es óptima. Describiremos una que garantiza la libertad de 99 prisioneros. El primer prisionero al que le preguntan tiene que resignarse al hecho de que no es su día de suerte. Él tendrá una probabilidad de 50 % de acertar. Lo que hace es contar los gorros blancos que puede ver delante de él. Si es un número impar, dice "blanco", y si es un número par, dice "negro". Ahora, el segundo prisionero cuenta los gorros blancos que él puede ver, y sabe que tiene un gorro blanco sobre la cabeza si y sólo si la paridad de ese número difiere de la paridad del número de gorros blancos que el primero contó. El tercer prisionero conoce la paridad del número de gorros blancos que el primero tenía delante, y sabe también si el segundo tenía o no un gorro blanco, así que puede saber cuántos gorros blancos veia el segundo prisionero, y compara la paridad de este número con la de la cantidad de gorros blancos que él ve. Así puede saber si su propio gorro es blanco. Este argumento se repite para todos los prisioneros siguientes, así que resulta que todos menos el primero aciertan.

En el caso finito, si tenemos una cantidad finita arbitraria de colores, y si llamamos H al conjunto de colores, los prisioneros le pueden dar a H estructura de grupo abeliano. Entonces, el primer prisionero hace la suma de todos los colores que ve. El segundo prisionero sustrae la suma de todos los colores que puede ver del color que el primero dijo y obtiene el color de su propio gorro. Otra vez, el argumento se repite, y al final resulta que todos menos el primer prisionero salen libres. Por supuesto, en el caso de los gorros blancos y negros, tenemos que $H = \{\text{negro, blanco}\} \cong \mathbb{Z}_2$, donde el color negro está comportándose como 0 y el blanco como 1.

Estudiemos ahora el caso infinito. En primer lugar, los prisioneros convienen en identificar al color negro con el 0 y al blanco con el 1. Para hablar del conjunto de sucesiones de 0's y 1's, es usual la notación siguiente: ${}^{\mathbb{N}}2 = \{f \mid f : \mathbb{N} \to 2\}$, donde $2 = \{0, 1\}$. Los prisioneros definen entonces la relación $\sim$ sobre ${}^{\mathbb{N}}2$ tal que para cualesquiera $\alpha, \beta \in {}^{\mathbb{N}}2$, $\alpha \sim \beta$ si y sólo si α y β difieren en una cantidad finita de lugares. Es fácil verificar que $\sim$ es de equivalencia, cosa que nos permite hablar de clases de equivalencia. Los prisioneros invocan el Axioma de Elección, y obtienen un conjunto M

de representantes de las clases de equivalencia. Seguidamente, todos memorizan a M. Supongamos que los prisioneros ya están formados y tienen todos un gorro. Digamos que la manera como los gorros están colocados es $s = (s_0, s_1, s_2, \ldots) \in {}^{\mathbb{N}}2$. Llamemos $x = (x_0, x_1, x_2, \ldots) \in M$ al representante de la clase de equivalencia de s. Para cada $n \in \mathbb{N}$, P_n ve algo de la forma $(\underbrace{?, \ldots, ?}_{n+1 \text{ entradas}}, s_{n+1}, s_{n+2}, s_{n+3}, \ldots)$ (las ?'s son porque él no ve el color de su propio gorro, ni el de los gorros de los prisioneros anteriores). Mentalmente, él da a las ?'s valores arbitrarios. La sucesión que obtiene está $\sim$-relacionada con s (por diferir de s en a lo más $n+1$ entradas) y por tanto está en la clase de eqivalencia cuyo representante es x, así que puede saber cuánto vale x_n, y dice ese color. De esta forma, para cada $n \in \mathbb{N}$, P_n dice x_n. Como $x \sim s$, todos los prisioneros salvo una cantidad finita acertarán (o equivalentemente, habrá un prisionero a partir del cual todos acertarán).

Lo verdaderamente asombroso es que el caso infinito se puede generalizar a una cantidad arbitraria de colores. Por ejemplo, supongamos que hay una cantidad infinita no numerable, κ, de colores. Entonces, en vez de ${}^{\mathbb{N}}2$, estamos trabajando en ${}^{\mathbb{N}}\kappa$, pero la estrategia y el resultado son los mismos. En esta situación, la probabilidad de que un prisionero adivine al azar el color de su gorro es 0. El guardia estará entonces sorprendido, no sin razón, cuando llegue a ese primer prisionero que conteste correctamente. Sin embargo, eso no será nada comparado con el hecho de que, a partir de un cierto prisionero, *todos* acierten.

Comentarios

A pesar de su nombre, la Paradoja de Banach-Tarski (PBT) es un teorema. No es una falacia en cuya prueba haya algún error. Ahora, en general, las paradojas surgen cuando lo demostrable choca con la intuición natural. Cuando esto ocurre, típicamente, la explicación es que la intuición estaba equivocada en primer lugar. Reconocer este hecho y afinar la intuición es lo que puede conducir a un avance epistémico.

Resolvamos ahora la PBT. Las transformaciones rígidas de $\mathbb{R}^3$ preservan el volumen de los subconjuntos acotados de $\mathbb{R}^3$. Entonces, uno pensaría que no puede haber un subconjunto acotado de $\mathbb{R}^3$ que sea G_3-paradójico. Sin embargo, si pensamos en la descomposición paradójica descrita en 4.4, no hay un volumen que preservar; los pedazos que intervienen en tal descomposición no son Lebesgue-medibles (atendiendo a la prueba de G.2, no pueden serlo, ya que la medida de Lebesgue es finitamente aditiva, isometría-invariante, y asigna a cualquier bola cerrada en $\mathbb{R}^3$ un número real positivo). Entonces, no tiene sentido hablar del volumen de estos pedazos ni, por supuesto, esperar que se preserve (pues los movimientos rígidos preservan el volumen de los conjuntos *que tienen volumen*). Desde luego, estos pedazos carecen de realidad física, por lo que la descomposición no se puede efectuar en el mundo real.

Lo que la PBT muestra inequívocamente es la naturaleza imaginaria de la idea irrestricta de un subconjunto de $\mathbb{R}^3$. En el Apéndice G quedó establecido que, aceptando el Axioma de Elección (AE), no puede existir una medida finitamente aditiva e isometría-invariante definida sobre $\mathbb{R}^3$ que normalice al cubo unitario. La PBT, en fin, hace ver la necesidad de adoptar construcciones más finas como la medida de Lebesgue.

Lejos de ser un problema terminal, la PBT tiene repercusiones importantes en varias ramas de las matemáticas. De hecho, la PBT pertenece, junto con el llamado Último Teorema de Fermat, a la categoría de problemas matemáticos cuyo estudio ha motivado la creación de nuevas áreas del

conocimiento. La idea de las descomposiciones paradójicas es el fundamento de una teoría de medidas finitamente aditivas, en la que se intersecan el análisis (teoría de la medida y funcionales lineales), el álgebra (teoría combinatoria de grupos), la geometría (grupos de isometrías) y la topología (grupos topológicos localmente compactos).

A manera de ejemplo, sigamos un razonamiento tomado de la teoría de los grupos paradójicos. Si G es un grupo, decimos que G es **manejable** (en inglés, *amenable*) si y sólo si hay una medida finitamente aditiva μ sobre G que normaliza a G, y que es traslación-invariante (es decir, que para cualesquiera $g \in G, X \subseteq G$, $\mu(gX) = \mu(X)$). Los grupos finitos son claramente manejables (pues si $|G| = n$, basta definir, para $A \subseteq G$, $\mu(A) = \frac{|A|}{n}$) y, aceptando AE, los grupos solubles también lo son.[1] Atendiendo a las definiciones, es directo que un grupo manejable no puede ser paradójico, ya que si un grupo G fuera manejable y paradójico, tendríamos lo siguiente, siguiendo la notación de 1.1:

$$\begin{aligned}
1 &= \mu(G) \\
&\geq \mu(\bigcup_{i=1}^{m} A_i \cup \bigcup_{i=1}^{n} B_i) \\
&= \sum_{i=1}^{m} \mu(A_i) + \sum_{i=1}^{n} \mu(B_i) \\
&= \sum_{i=1}^{m} \mu(g_i A_i) + \sum_{i=1}^{n} \mu(h_i B_i) \\
&\geq \mu(\bigcup_{i=1}^{m} g_i A_i) + \mu(\bigcup_{i=1}^{n} h_i B_i) \\
&= \mu(G) + \mu(G) \\
&= 1 + 1 \\
&= 2
\end{aligned}$$

En este trabajo probamos (1.8) que todo grupo que tenga un subgrupo libre no abeliano es paradójico. (Ahora se sabe que, en general, la afirmación recíproca no es cierta.) Un razón por la que no es posible reproducir la demostración de la PBT presentada en este trabajo para $\mathbb{R}$ o para $\mathbb{R}^2$ es que G_1 y G_2 son solubles, cosa que no es difícil demostrar.[2] Como son

[1] Véase [W, Teo. 10.2].
[2] Véase [W, Ap. A].

solubles, son manejables, y por tanto no paradójicos. Entonces, ninguno tiene subgrupos libres no abelianos, en particular de rango 2.

Por supuesto, hay una razón más simple. Claramente, $SO_1 = \{id_{\mathbb{R}}\}$ y se vio en D.4 que SO_2 es también abeliano. Entonces, ninguno de los dos admite un subgrupo libre de rango 2, que sería no abeliano.

Banach construyó,[3] apoyándose en la manejabilidad de G_1 y de G_2, medidas finitamente aditivas e isometría-invariantes sobre $\mathbb{R}$ y $\mathbb{R}^2$ que extienden a la medida de Lebesgue (y que por lo tanto normalizan a $[0,1]$ y a $[0,1] \times [0,1]$, respectivamente). Con ello, y en vista de lo discutido al final del Apéndice G, queda establecido que no hay PBT en $\mathbb{R}$ ni en $\mathbb{R}^2$.

Si $n \geq 3$, no es difícil (pero tampoco es inmediato) hacer ver que hay PBT en $\mathbb{R}^n$.[4] Para obtener esta generalización basta modificar la prueba de 3.1, con lo que la demostración de 1.4, que es el punto en el que AE fue utilizado, queda intacta. Entonces, si $n \geq 3$, $\mathbb{R}^n$ no admite una medida finitamente aditiva y G_n-invariante que normalice a $[0,1]^n$.

De manera interesante, la construcción de Banach utiliza AE. Así, en $\mathbb{R}^n$, AE construye paradojas si $n \geq 3$ y las destruye si $n \leq 2$.

Sin embargo, no todas las consecuencias de la PBT son negativas. Tarski la usó para probar un resultado sobre medidas finitamente aditivas definidas en el álgebra de subconjuntos Lebesgue-medibles de $\mathbb{R}^3$, y este resultado fue utilizado recientemente para demostrar la unicidad de la medida de Lebesgue.[5]

Ahora, un resultado, conocido con el nombre de Teorema de Tarski,[6] afirma que, aceptando AE, si un grupo G actúa sobre un conjunto X y $E \subseteq X$, E no es G-paradójico si y sólo si hay una medida finitamente aditiva y G-invariante sobre X que normaliza a E. (La implicación hacia atrás se prueba fácilmente, de manera similar a la demostración recién vista de que un grupo manejable no puede ser paradójico.) Esto, además de vincular la teoría de las descomposiciones paradójicas con la de las medidas finitamente aditivas, proporciona que, como todo grupo actúa sobre sí mismo por traslación, los grupos manejables son precisamente aquellos que no son paradójicos.

La demostración de la PBT, como todas las pruebas de la existencia de conjuntos no Lebesgue-medibles (ver abajo), no es constructiva, desde el momento en que apela al Axioma de Elección (AE). Con todo y lo ya explicado, se ha argumentado que el resultado es tan contraintuitivo, tan patentemente falso en el mundo real, que alguna de las suposiciones subya-

[3] Véase [W, Cor. 10.9].
[4] Véase [W, Cap. 5].
[5] Véase [W, Lema 9.7 y los comentarios después de Teo. 11.11].
[6] Véase [W, Cor. 9.2].

centes debe ser incorrecta; AE es comúnmente señalado como el culpable. Este argumento se discute con profundidad en [W, Cap. 13], donde el papel que AE desempeña en los fundamentos de la teoría de la medida es analizado con detalle (y donde se pueden encontrar referencias para los resultados que mencionaremos en lo que resta de esta sección). Denotemos como ZF al conjunto de los axiomas de Zermelo-Fraenkel de la teoría de conjuntos (que son los axiomas estándar). Recordemos que Gödel dio un modelo de ZF, el universo construible, en el cual AE es verdadero, cosa que demuestra la consistencia relativa a ZF de AE (es decir, que si ZF es consistente, también lo es $ZF \cup \{AE\}$). Así, independientemente de su valor de verdad en la visión de cualquier individuo acerca del universo de la teoría de conjuntos, la PBT es, por lo menos, consistente.

Ahora, en este trabajo se hizo ver que AE es suficiente para obtener la PBT. Cabe la pregunta de si es necesario. En 1964, usando la recién descubierta técnica de forcing, Solovay dio un modelo de ZF en el cual todos los subconjuntos de $\mathbb{R}$ son Lebesgue-medibles,[7] y por lo tanto[8] todos los subconjuntos de $\mathbb{R}^n$, para cualquier $n \in \mathbb{Z}^+$, son Lebesgue-medibles. Como la PBT implica la existencia de subconconjuntos de $\mathbb{R}^3$ no Lebesgue-medibles, en ese modelo no hay PBT. Con ello, aceptando la consistencia de ZF, la PBT no se puede probar a partir de ZF (se dice que no es un *teorema* de ZF).

Decimos que un enunciado (es decir, una afirmación) σ es un **debilitamiento** de otro enunciado τ si y sólo si τ implica σ pero σ no implica τ. Hay un debilitamiento de AE llamado Axioma de Elecciones Dependientes (ED), comúnmente aceptado en la teoría de la medida, que básicamente garantiza la existencia de conjuntos numerables de representantes $\{a_n\}_{n\in\mathbb{N}}$ tales que, para todo $n \in \mathbb{N}$, la elección de a_{n+1} depende de a_n. Solovay también probó que, aceptando la consistencia de ZF, la PBT no es un teorema de $ZF \cup ED$.

No es razonable esperar que la PBT sea equivalente a AE, ya que para obtener la PBT basta tener un buen orden sobre $\mathbb{R}$ (revísese la prueba de 1.4 y la aplicación de este resultado en 3.2), mientras que AE es equivalente a la afirmación de que *todo* conjunto es bien ordenable. Por supuesto, ED no puede proporcionar la existencia de un buen orden sobre $\mathbb{R}$.

Por último, no está de más hacer énfasis en el hecho de que la demostración aquí expuesta de la PBT, con su aplicación de ingeniosos métodos geométricos y algebraicos para hacer construcciones paradójicas, pone

[7] Esto equivale a la consistencia relativa a ZF de la afirmación de que todos los subconjuntos de $\mathbb{R}$ son Lebesgue-medibles.

[8] Véase la prueba de [W, Teo. 7.9].

de manifiesto la estrecha, pero en general difícil de apreciar, relación que hay entre áreas de la matemática tan aparentemente disímiles como el álgebra, el análisis y la teoria de conjuntos.

Notas Históricas

Las observaciones de Galileo Galilei sobre conjuntos infinitos se remontan al siglo XVII (1638).

La construcción clásica de un conjunto no Lebesgue-medible fue hecha por Giuseppe Vitali, en 1905.

En 1914, Wacław Sierpiński planteó la posibilidad de la existencia de un subconjunto no vacío de $\mathbb{R}^2$ que fuera G_2-paradójico. La pregunta fue respondida en sentido positivo por su alumno Stefan Mazurkiewicz, cuya prueba fue simplificada por Sierpiński. Como resultado, se dio a conocer la paradoja que hoy lleva el nombre de ambos.

También en 1914, Felix Hausdorff construyó, en $\mathbb{S}^2$, la descomposición paradójica que luego tomó su nombre. A él se debe el descubrimiento de que los grupos libres encierran paradojas a las que se puede dar una interpretación geométrica. Hausdorff estaba motivado por la búsqueda de la prueba de la no existencia de ciertas medidas sobre los espacios euclidianos. Su técnica para probar 3.1 fue distinta de la que aparece en este trabajo; fue Stanisław Świerczkowski quien, en 1958, desarrolló la demostración que se presenta aquí.

La PBT fue demostrada por Stefan Banach y Alfred Tarski en 1924, y muchos artículos que la discuten o simplifican han aparecido desde entonces. La PBT surgió como una sofisticación de la Paradoja de Hausdorff. Banach y Tarski estaban motivados, de manera parecida a Hausdorff, por la pregunta de si $\mathbb{R}^3$ admite una medida finitamente aditiva e isometría-invariante que normalice al cubo unitario. Fueron también Banach y Tarski los que hablaron primero de G-equidescomponibilidad. Banach demostró 2.8 y generalizó así el Teorema de Schröder-Bernstein.

Kurt Gödel probó, en 1938, la consistencia relativa a ZF de AE.

Sierpiński obtuvo 4.1 en 1948, y lo aplicó para obtener la Paradoja de Banach-Tarski a partir de la de Hausdorff. La derivación original difiere, pero no mucho.

Paul Cohen desarrolló la técnica conocida como forcing. Poco tiempo después (en 1964), y en parte siguiendo sugerencias del propio Cohen, Robert Solovay la empleó para obtener sus resultados sobre consistencias relativas.

Reconocimientos

La estructura general de este trabajo sigue a [W], que es un excelente (y elegante) compendio de lo que actualmente se sabe, y de lo que no se sabe, en torno a la Paradoja de Banach-Tarski.

[S] es una "tesis menor" (*Minor Thesis*) preparada para obtener un Ph.D. en la Universidad de Harvard. A pesar de que, en general, sigue a [W], diferen en varios puntos técnicos. Por ejemplo, las rotaciones específicas que figuran en 3.1 son las usadas en [S].

Se tomaron también algunas ideas de [I], que presenta una prueba distinta, pero sustentada en los mismos principios, de la Paradoja de Banach-Tarski. Su enfoque es más gráfico, y su prueba de la Paradoja de Hausdorff se acerca más a la original.

Las ideas acerca del papel que desempeñan las paradojas en la epistemología son del director de este trabajo, el Dr. José Alfredo Amor y Montaño, y se pueden encontrar en [A].

El crédito por E.1 corresponde a Osvaldo Guzmán González, que además de desarrollar el resultado, me hizo ver su importancia.

La Paradoja de Sierpiński-Mazurkiewicz fue publicada en [M-S].

El problema de los prisioneros (Apéndice H) fue tomado de una plática de Mike O'Connor.

Bibliografía

[A] Amor, J.A., "Paradojas, Lógica y Epistemología", artículo sometido en 2008 a las memorias del XIV Congreso Internacional de Filosofía, realizado en noviembre de 2007.

[E1] Efimov, N., *Curso Breve de Geometría Analítica*, Moscú: MIR, 1969.

[E2] Efimov, N., *Matrices y Formas Cuadráticas*, Moscú: MIR, 1969.

[H] Hoffman, K., Kunze, R., *Linear Algebra*, Second Edition, New Jersey: Prentice-Hall, 1971.

[I] Ivorra, C., "La Paradoja de Banach-Tarski", www.uv.es/=ivorra

[M] Martin, G., *Transformation Geometry - An Introduction to Symmetry*, New York: Springer-Verlag, 1982.

[M-S] Mazurkiewicz, S., y Sierpiński W., "Sur un ensemble superposable avec chacune de ses deux parties", *Comptes Rendus Acad. Sci. Paris*, **158**(1914), pp. 618-619.

[S] Su, F.E., "The Banach-Tarski Paradox", tesis doctoral, Univ. de Harvard, 1990.

[W] Wagon, S., *The Banach-Tarski Paradox*, New York: Cambridge Univ. Press, 1985.