

UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA

FACULTAD DE CIENCIAS

LA PARADOJA DE BANACH-TARSKI

*TRABAJO FIN DE MÁSTER EN MATEMÁTICAS AVANZADAS
ESPECIALIDAD DE ANÁLISIS MATEMÁTICO*



Autor *JOSÉ LUIS DEL OLMO CLAUDIO*

Tutor *JORGE LÓPEZ ABAD*

JUNIO 2021

Este trabajo es presentado por José Luis Del Olmo Claudio para la obtención del Título de Máster en Matemáticas Avanzadas por la UNED en la especialidad de Análisis Matemático. En consonancia con la normativa aprobada en Consejo de Gobierno de 10 de octubre de 2017, sobre la regulación de los trabajos fin de Máster en las enseñanzas conducentes al título oficial de Máster de la UNED, yo, José Luis Del Olmo Claudio, pongo el presente trabajo a disposición del tutor, el Dr. Jorge López Abad, para su revisión y ,en su caso, visto bueno para su posterior defensa ante tribunal.

Fdo. José Luis Del Olmo Claudio

Madrid, 21 de Junio de 2021

Fdo. Jorge López Abad

Madrid, de Junio de 2021

A mis padres

Agradecimientos

A mis padres les agradezco todo, a ellos les debo, entre otras muchas cosas, mi existencia y el hecho de encontrarme en un mundo que siempre despertó en mi la curiosidad por conocer su funcionamiento. Esto último ha sido uno de los principales motivos por el que decidí estudiar este Máster.

Al profesor tutor Jorge López Abad por su dirección y tutela. Ha sido atento, implicado, comunicativo y siempre ha estado inmediatamente disponible para cualquier tipo de consulta. Sus revisiones y consejos han sido fundamentales para la elaboración del presente trabajo.

A la UNED por la oportunidad de haberme brindado la posibilidad de satisfacer, ya de adulto, mi inquietud por profundizar en el estudio de las matemáticas. Deseo y vaticino que esa inquietud nunca desaparecerá.

RESUMEN

El presente trabajo tiene como tema central la paradoja formulada por Banach y Tarski en 1924. En su forma original enuncia que dados dos subconjuntos acotados A, B de \mathbb{R}^3 con interior no vacío, existen dos particiones de los mismos, $\{A_i : 1 \leq i \leq n\}$ y $\{B_i : 1 \leq i \leq n\}$ de tal manera que A_i y B_i son congruentes para todo i . En sentido intuitivo, nos dice que si cogemos una superficie esférica de cierto tamaño, podemos dividirla en un número finito de pedazos de tal manera que mediante sólo rotaciones y traslaciones podemos formar una superficie esférica de tamaño superior o varias superficies esféricas del mismo tamaño. El desarrollo del trabajo se realiza a lo largo de 5 capítulos. En el primer capítulo se realiza una exposición del desarrollo histórico de los conceptos relacionados con la paradoja. En el segundo capítulo se presenta una serie de conocimientos preliminares sobre diversas áreas de la matemática para posibilitar una total comprensión de la misma. El capítulo 3 está dedicado a la exposición de la paradoja en sí, siendo uno de los capítulos más importantes, mientras que el capítulo 4 está dedicado al estudio de la paradoja en dimensiones inferiores y superiores a 3. En el quinto capítulo analizaremos la relación que existe entre la posible existencia de una medida y la existencia de la paradoja e introduciremos la noción de grupo amenable y en el sexto capítulo se expondrá unas breves conclusiones finales. Finalmente, debido al importante papel que juega el Axioma de Elección en la demostración de la paradoja se ha incorporado un apéndice final en el que se habla del mismo.

ABSTRACT

The present work has as its central subject the paradox formulated by Banach and Tarski in 1924. In its original form it states that given two bounded subsets A, B of $\mathbb{R}^n, n \geq 3$ with non-empty interior, there are two partitions, $\{A_i : 1 \leq i \leq n\}$ and $\{B_i : 1 \leq i \leq n\}$ in such a way that A_i y B_i are congruent for every i . From an intuitive point of view, it tells us that if we take a spherical surface of a certain size, we can divide it into a finite number of pieces in such a way that by means of rotations and translations alone we can form a spherical surface of larger size or several spherical surfaces of the same size as the original one. The development of the work is carried out in x chapters. In the first chapter an exposition of the historical development of the concepts related to to the paradox is presented. In the second chapter a series of preliminary knowledge about different areas of mathematics is showed in order to make possible a complete understanding. Chapter 3 is dedicated to the exposition of the paradox itself, being one of the most important chapters, while chapter 4 will be dedicated to the study of the paradox in dimensions lower and higher than 3. In the fifth chapter we will analyze the relationship between the possible existence of a measure and the existence of the paradox and we will introduce the notion of amenable group and in

the sixth chapter we will present some brief final conclusions. Finally, due to the important role played by the Axiom of Election in the demonstration of the paradox, a final appendix has been included in which it is discussed.

Un segmento y un segmento parcial arbitrariamente pequeño, un kilómetro y un milímetro, el globo solar y una gota de agua tienen en este sentido el mismo número de puntos

Felix Hausdorff en Grundzüge der Mengenlehre. Leipzig, 1914.

Índice general

1. Introducción	9
2. Conceptos preliminares	15
2.1. Sobre números y conjuntos	15
2.2. Unas nociones de teoría de grupos	19
2.3. Una breve introducción a la teoría de la medida	27
3. la Paradoja de Banach Tarski	36
3.1. Conjuntos paradójicos	36
3.2. La paradoja de Hausdorff	40
La paradoja de Hausdorff original	44
3.3. Congruencia y equidescomponibilidad	46
El teorema de Banach–Schröder–Bernstein	49
3.4. La paradoja de Banach Tarski en \mathbb{R}^3	50
3.5. Mínimo número de piezas en una descomposición paradójica	52
Análisis para la esfera \mathbb{S}^2	52
Análisis para la bola tridimensional B	56
4. Tratamiento en otras dimensiones	58
4.1. La paradoja de Banach Tarski en dimensión mayor que 3	58
4.2. La paradoja de Banach-Tarski en dimensión menor que 3	63
Descomposición paradójica en \mathbb{R}^2	64
Descomposición paradójica en \mathbb{R}	67
5. Medida y descomposiciones paradójicas	71
5.1. Teorema de Tarski	71
5.2. Grupos amenables	75
Grupos amenables en \mathbb{R} y \mathbb{R}^2	79

6. Comentarios finales	80
A. El papel del Axioma de Elección	82

Introducción

Desde el principio de los tiempos, el infinito ha sido una entidad abstracta que ha generado inquietud en la mayoría de los seres humanos. Cuando estaba en enseñanza media, una de las cosas que me fascinaba y no lograba entender eran hechos como que un intervalo finito (por ejemplo en la recta real) tuviera infinitos puntos o cómo una serie de infinitos términos podría tener una suma no infinita. Estoy seguro que este tipo de preguntas o inquietudes ayudaron a despertar muchas vocaciones tempranas en muchos de los que ya de adultos decidimos estudiar matemáticas.

La paradoja de Banach Tarski, que es sobre la que versa este trabajo, está en la raíz de lo que mencionaba antes (la relación entre lo finito y lo infinito). Se puede enunciar del siguiente modo:

Dada una esfera en el espacio tridimensional existe una descomposición de ésta en un número finito de piezas disjuntas de tal manera que solamente mediante rotaciones y traslaciones de cada pieza pueden juntarse todas las piezas de nuevo generando dos copias idénticas de la esfera original.

El resultado es contraintuitivo (como ocurre en toda paradoja). La paradoja fue planteada por Stefan Banach y Alfred Tarski en 1924 en el artículo *Sur la decomposition des ensembles de points en parties respectivement congruentes*.

Evidentemente no era la primera vez que se abordaba de manera indirecta el concepto de infinito (sólo si se consideran infinitos puntos, se puede “entender” aunque no sin dificultad que puedas duplicar una figura en dos utilizando sólo rotaciones) sino que desde muy antiguo ya habían sido planteadas paradojas en torno a este concepto. A todo el mundo le viene a la memoria la paradoja clásica de Zenón de Aquiles y la tortuga ya planteada por los filósofos griegos. Desde entonces hasta el siglo XVII, el infinito es un concepto escabroso que es tratado de esquivar por la mayoría de los filósofos y matemáticos. Es en ese siglo cuando Galileo aborda el concepto esta vez intentando relacionar conjuntos infinitos de números enteros, por ejemplo el conjunto de los números naturales y el conjunto de los números naturales que son cuadrados perfectos (esto es, que son otro número natural al cuadrado). En su célebre paradoja, menciona

Cantor decidió dar a todos los conjuntos numerables una designación. Les llamó conjuntos de potencia \aleph_0 de tal manera que todos los conjuntos infinitos numerables tenían el mismo número de elementos que \mathbb{N} . La aportación de Cantor no terminó aquí sino que también demostró que hay conjuntos infinitos que son tan “densos” que no pueden ser numerados (contados), siendo uno de los ejemplos más representativos el conjunto de los números reales. Este conjunto, como el conjunto de cualquier intervalo finito dentro de \mathbb{N} o como el conjunto de los puntos en un círculo, son continuos. Su infinito es de una potencia más alta que \aleph_0 , 2^{\aleph_0} .

Pero fue más allá demostrando que había conjuntos infinitos de una potencia más alta que 2^{\aleph_0} . Lo hizo demostrando que el conjunto de todos los subconjuntos de un conjunto dado, conjunto potencia, siempre tiene más elementos que el propio conjunto, de hecho si el conjunto original tiene n elementos, el número de elementos de su conjunto potencia es 2^n . Trasladando estas ideas a conjuntos infinitos, a las potencias \aleph_0 , 2^{\aleph_0} , $2^{2^{\aleph_0}}$ les llamó números cardinales transfinitos creando una aritmética singular entre ellos (basta mencionar por ejemplo que $\aleph_0 + \aleph_0 = \aleph_0$).

Por sus ideas revolucionarias, Cantor tuvo muchos detractores. Uno de ellos especialmente importante fue su profesor Leopold Kronecker quien diría al respecto del infinito:

...futilidad perniciosa heredada de filosofías anticuadas y teologías confusas, pudiendo llegar sin él tan lejos como se quiera...

Sin embargo David Hilbert diría años más tarde: *Nadie sería capaz de expulsarnos del paraíso que Cantor creó para nosotros* y de hecho el primer problema de su famosa lista consistiría en averiguar si habría un conjunto cuya cardinalidad fuera mayor que la cardinalidad de \mathbb{Z} pero menor que la de \mathbb{R} .

De hecho fue también Cantor con su nuevo planteamiento el que inconscientemente estableció las bases para llegar a una fundamentación de una nueva e importante área de la lógica matemática que es la teoría de conjuntos. Cantor decidió aplicar sus ideas a la teoría de conjuntos comenzando a hacerse preguntas como cuál sería el conjunto más grande, el conjunto de todos los conjuntos. Tal conjunto si existiera, tendría el cardinal más alto posible. Pero por otro lado como Cantor ya había apuntado, de cualquier conjunto con un cardinal dado, uno puede crear un conjunto con un cardinal mayor, por ejemplo el conjunto potencia del conjunto original resultando una de las primeras paradojas en teoría de conjuntos.

Posteriormente, el filósofo y matemático Bertrand Russell recoge el testigo de Cantor en ese sentido y descubre otra paradoja haciéndose el mismo tipo de pregunta. Un ejemplo de paradoja de Bertrand Russell (se puede expresar de múltiples formas) sería la siguiente: Consideremos el conjunto S de todos los conjuntos que no pertenecen a sí mismos y estudiemos la pertenencia del propio conjunto S a sí mismo. Si S pertenece a sí mismo entonces eso contradice la definición de S . Si en cambio no pertenece a sí mismo, entonces debe pertenecer a sí mismo pues cumple con su propia definición y eso también es una contradicción.

Este tipo de paradojas condujo a muchos matemáticos a pensar que era necesaria una redefinición de las bases de la teoría de conjuntos. Esa reformulación de los fundamentos de la teoría de conjuntos vino de la mano de un matemático también alemán, Ernst Zermelo, el cual en 1908 formuló una serie de nueve axiomas que definen el concepto de conjunto y regulan su utilización. Más específicamente, los nueve axiomas establecen las relaciones básicas entre conjuntos (unión, igualdad, intersección, inclusión) y reafirman la existencia de conjuntos nulos, infinitos y conjuntos que pertenecen a sí mismos. Catorce años más tarde Fraenkel mejoraría la definición de los axiomas, motivo por el que se conocen como el sistema axiomático de Zermelo-Fraenkel.

Entre los nueve axiomas hay uno especialmente importante que todavía en la actualidad genera controversia: el Axioma de Elección. Tiene un papel importante en la demostración de la paradoja de Banach Tarski y sus precursoras (paradoja de Hausdorff) ya que para su demostración (y para muchas otras) se presupone el poder utilizarlo. Se enuncia de la manera siguiente:

”Sea S un conjunto de conjuntos no vacíos y de intersección nula. Entonces es posible crear un nuevo conjunto eligiendo un único elemento de cada uno de los conjuntos que forman S .”

Si hablamos de conjuntos finitos lo anterior es trivial. Sin embargo, si nos referimos a conjuntos infinitos inmediatamente surge una cuestión: ¿cuál será la forma en que haremos la elección?. Puede parecer evidente pero no lo es y ya sabemos que la intuición muchas veces falla cuando hablamos de este tipo de conjuntos.

El Axioma de Elección es un recurso que ha sido utilizado para demostrar importantes y famosos teoremas en diversos campos de la Matemática (por ejemplo el teorema de Hahn-Banach en Análisis Funcional), no obstante, genera resultados a veces poco intuitivos. Entre ellos se encuentra la existencia de conjuntos que no son medibles (Lebesgue).

En 1902, Lebesgue defiende en su tesis doctoral la propuesta de una medida μ en \mathbb{R}^n que cumpla varias propiedades, entre ellas, que sea:

- Invariante por traslaciones, es decir, si A es un conjunto medible $\mu(A) = \mu(A + x) \forall x \in \mathbb{R}^n$
- Numerablemente aditiva: Si $\{A_i\}$ es una colección infinito-numerable de conjuntos disjuntos, entonces:

$$\mu\left(\bigcup_{i=1}^n A_i\right) = \mu(A_1) + \mu(A_2) + \cdots + \mu(A_n)$$

Esta medida extendía la posibilidad de dotar de medida a otro tipo de conjuntos como los números racionales, \mathbb{Q} , resultando en un conjunto de medida Lebesgue nula. Sin embargo, en 1905 Vitali encuentra un conjunto que no es medible Lebesgue. En la construcción de ese conjunto se utiliza el Axioma de Elección. Unos diez años después Hausdorff haciendo uso de este tipo de conjuntos enuncia la paradoja que lleva su nombre:

"Hay una descomposición de la esfera unitaria \mathbb{S}^1 en el espacio euclídeo \mathbb{R}^3 en cuatro conjuntos disjuntos A, B, C, D de tal manera que A, B, C y $B \cup C$ son congruentes entre sí y D es numerable"

Por lo tanto, según Hausdorff la esfera es "casi divisible" en 3 pedazos: A, B y C congruentes entre sí y con $B \cup C$. Esto último representa lo extraño y es lo que nos permite duplicar casi toda la esfera original en otras dos esferas. El problema estriba en que Hausdorff construye esos conjuntos utilizando el Axioma de Elección.

Claramente este teorema es el preámbulo que anticipa la redacción por parte de Stefan Banach y Alfred Tarski en 1924 del artículo *Sur la decomposition des ensembles de points en parties respectivement congruentes*. El Lema número 27 de dicho artículo se enuncia a continuación:

"Toda superficie S de una esfera puede ser descompuesta en dos subconjuntos disjuntos A_1 y A_2 tales que S es equidescomponible con A_1 y S es equidescomponible con A_2 ".

El concepto de equidescomponibilidad entre dos conjuntos (A, B) viene a decir que cada conjunto está formado por un conjunto (finito o infinito numerable) de pedazos de tal manera que hay una relación de congruencia entre un pedazo del primer conjunto y otro pedazo del segundo conjunto. Formalmente, si $A = \bigcup_{i=1}^n A_i, B = \bigcup_{i=1}^n B_i$, tales que $A_i \cap A_j = \emptyset = B_i \cap B_j, i < j \leq n$, entonces existen $g_1, \dots, g_n \in G$ tal que para cada $i \leq n, g_i(A_i) = B_i$.

Lógicamente los conjuntos A_1 y A_2 no son μ -medibles pues si μ es una medida invariante se debe cumplir:

$$\mu(S) = \mu(A_1) + \mu(A_2) = \mu(S) + \mu(S) = \mu(2S), \text{ y esto es imposible.}$$

En el mismo artículo generalizan los resultados de la esfera a la bola y a conjuntos arbitrarios acotados en \mathbb{R}^n con $n > 3$, lo que da resultado la forma fuerte de la paradoja:

"Si A y B son dos subconjuntos acotados cualesquiera de \mathbb{R}^3 , cada uno con interior no vacío, entonces A y B son equidescomponibles."

En su momento, debido a la facilidad con que era entendido su enunciado, la paradoja causó estupor tanto en el público no especializado como en la comunidad matemática aunque diversos representantes en ésta última (entre ellos Borel) ya sabían que el problema estribaba de nuevo en la utilización del Axioma de Elección. Concretamente éste era utilizado para la construcción de los conjuntos (infinitos) no medibles utilizados en su demostración.

La polémica con el Axioma de Elección (AE) y su utilización en Matemáticas continúa hasta el día de hoy. De hecho, desde que fue postulado y sobre todo a raíz de detectar los resultados contraintuitivos que generaba su uso surgió un enfrentamiento entre los partidarios de no utilizarlo (corriente constructivista) y aquellos que sí veían necesaria su utilización (corriente intuicionista). En 1935 Zorn formuló el lema que

lleva su nombre subrayando que era equivalente a AE. Poco después, Kurt Gödel demostró que el Axioma de Elección era compatible con los otros axiomas contenidos en el sistema axiomático de Zermelo-Fraenkel y en 1963 Paul Cohen demostró que no se podía deducir del resto de axiomas.

Es cierto que el desarrollo de muchos resultados importantes en áreas muy diversas de las Matemáticas en los últimos cien años no hubiera sido posible sin utilizarlo. Por mencionar algunos ejemplos, el Axioma de elección interviene en teoremas como el teorema de Hahn-Banach en Análisis Funcional, el teorema de Tychonoff en topología, el teorema de König en teoría de conjuntos,...

Por último, merece la pena mencionar que el estudio de las descomposiciones paradójicas ha sido ampliamente tratado a lo largo del siglo anterior y no se ha circunscrito sólo a los conjuntos acotados en \mathbb{R}^n bajo el grupo de transformaciones SO_n con $n \geq 2$. Por citar algunos de muchos ejemplos, se han estudiado descomposiciones paradójicas entre conjuntos acotados en el plano bajo grupos de transformaciones más restringidos (Von Neumann), descomposiciones paradójicas entre conjuntos de primera categoría (Baire), etc.

ESQUEMA DEL TRABAJO

Para el desarrollo del trabajo hemos utilizado principalmente como libro de referencia el libro *La paradoja de Banach Tarski* escrito por Stan Wagon [11].

Cabe resaltar que para una total comprensión de la paradoja y todas las demostraciones que tienen relación con ella es necesario tener conocimientos básicos sobre ciertas áreas de la matemática. En este sentido se ha pretendido que este trabajo sea autocontenido y que cualquier lector sea capaz de entender lo que en él se expone partiendo de cero, sin tener que acudir a ningún libro de consulta. Para cumplir con este objetivo en el primer capítulo expondremos un compendio bastante variado de conocimientos, entre ellos, nociones básicas de teoría de conjuntos, teoría de grupos y teoría de la medida. Únicamente se presupondrá al lector unos conocimientos mínimos de álgebra de conjuntos que generalmente se suelen adquirir en enseñanza primaria o básica.

En el segundo capítulo introduciremos el concepto de descomposición paradójica y ciertos teoremas relacionados. Más adelante postularemos el Teorema de Hausdorff e introduciremos el concepto de G -equidescomponibilidad y enunciaremos y demostraremos la Paradoja en sí en su versión fuerte y débil. Finalizaremos este capítulo con un análisis del número mínimo de piezas para que se verifique la paradoja con la esfera y la bola tridimensionales. En el cuarto capítulo estudiaremos la paradoja en dimensiones 1 y 2 y en dimensiones superiores a 3. En el quinto capítulo analizaremos la relación que existe entre la posible existencia de una medida y la existencia de la paradoja y el sexto capítulo contendrá unas breves conclusiones finales.

Conceptos preliminares

En este capítulo trataremos de exponer al lector una serie de conceptos matemáticos básicos los cuales aparecerán en el desarrollo de los capítulos posteriores. Principalmente sera una exposición, en la cual se obviarán mucha parte de las demostraciones, dejando el esfuerzo demostrativo para los teoremas más relacionados con la demostración de la paradoja en sí desarrollados más adelante.

2.1. Sobre números y conjuntos

De una manera informal podemos definir un conjunto como cierta colección de elementos.

Podemos formar un conjunto enumerando los objetos que pertenecen a él o en su caso, enunciando la propiedad que comparten. Por ejemplo, $A=\{1, 2, 3, 5\}$, $B=\{\text{conjunto de números pares entre 0 y 100}\}$, o $C=[3,7]$ siendo éste último un intervalo de la recta real que comprende todos los números reales entre 3 y 7, ambos inclusive.

Podemos construir conjuntos de muchas maneras diferentes. El Axioma de Elección, que más tarde mencionaremos, nos posibilita el poder construir conjuntos a partir de una familia de conjuntos dada.

A continuación haremos referencia a los distintos conjuntos de números que habitualmente son utilizados en Matemáticas:

- Números naturales: son aquéllos que se utilizan para contar elementos. El conjunto de los números naturales se designa con la letra \mathbb{N} , $\mathbb{N}=\{1, 2, 3, \dots\}$ (obsérvese que no está incluido el cero).
- Números enteros: es el conjunto de los números naturales, sus opuestos y el cero. A este conjunto se le hace referencia con la letra \mathbb{Z} .
- Números racionales: aquellos que pueden ser representados como el cociente de dos números enteros. A este conjunto se le denomina con la letra \mathbb{Q} .

- Números irracionales: aquéllos cuyo valor no puede ser representado como el cociente de dos números enteros. Algunos representantes muy conocidos de este tipo de números son $\sqrt{2}$, π , e ,... Hay dos tipos de números irracionales: los números algebraicos son aquellos irracionales que son solución de alguna ecuación del tipo:

$$a_0 + a_1x^1 + a_2x^2 + \dots + a_nx^n=0, a_i \in \mathbb{Z} \setminus \{0\}, i = 1, \dots, n$$

y los números trascendentes que no son solución de ninguna ecuación algebraica con coeficientes enteros no nulos. De hecho, si introducimos la definición de números algebraicamente independientes sobre un cuerpo K como aquel conjunto tal que no satisface una ecuación polinómica no trivial, esto es una ecuación del tipo:

$$a_0 + a_1x^1 + a_2x^2 + \dots + a_nx^n=0, a_i \in K$$

podemos redefinir el conjunto de números trascendentes como el conjunto de números algebraicamente independientes sobre $\mathbb{Z} \setminus \{0\}$. No existe una letra específica para designar el conjunto de los irracionales, de tal manera que siempre que nos refiramos a ellos lo haremos como el complemento de los racionales ($\mathbb{R} \setminus \mathbb{Q}$).

- Números reales: es el conjunto unión de los racionales e irracionales (\mathbb{R})
- Números complejos: Incluye a las combinaciones lineales de números reales e imaginarios puros (raíces pares de números reales negativos). Es el conjunto más amplio y es designado como \mathbb{C} .

De lo anterior es claro que:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

Cardinalidad de un conjunto

Una vez introducidos los conjuntos anteriores, surge de manera natural una pregunta: ¿Podemos cuantificar su contenido?. En este sentido aparece la siguiente definición:

Definición 2.1.1. El cardinal de un conjunto finito A es el número de elementos que tiene dicho conjunto.

Es intuitivo entender que no hay problema a la hora de encontrar el cardinal de un conjunto finito. Pero como vimos en la introducción no es tan obvio cuando el conjunto es infinito.

Al cardinal de un conjunto infinito le llamaremos número transfinito utilizando la terminología de Cantor. De este modo, no todos los conjuntos infinitos tienen el mismo cardinal. Al cardinal de \mathbb{N} se le denomina \aleph_0 .

Teorema de Cantor-Schröder-Bernstein. Sean dos conjuntos A, B . Si existen dos funciones inyectivas $f_1: A \rightarrow B$, $f_2: B \rightarrow A$ entre los conjuntos A y B , entonces existe una función biyectiva g entre A y B .

Más concretamente el Teorema de Schröder-Bernstein establece que para dos conjuntos A, B si $\text{card}(A) \leq \text{card}(B)$ y $\text{card}(B) \leq \text{card}(A)$, entonces $\text{card}(A) = \text{card}(B)$.

En la siguiente proposición se establece cómo asignar cardinales a otros conjuntos infinitos con cierta similitud a \mathbb{N} .

Proposición 2.1.2. *Dos conjuntos A y B tienen la misma cardinalidad, ambos son equipotentes, si y sólo si existe una aplicación $f : A \rightarrow B$ que es biyectiva (esto es, una correspondencia uno a uno entre los elementos de A y los de B).*

De esta manera conjuntos “contenidos” en \mathbb{N} como pudieran ser el conjunto de los pares o impares tendrían el mismo cardinal ya que entre \mathbb{N} y los números pares se puede establecer la siguiente biyección:

$$\begin{aligned} f : \mathbb{N} &\rightarrow \mathbb{P} \\ n &\mapsto 2n \end{aligned}$$

siendo \mathbb{P} el conjunto de los números pares, de igual modo que para los impares (\mathbb{I}):

$$\begin{aligned} f : \mathbb{N} &\rightarrow \mathbb{I} \\ n &\mapsto 2n - 1 \end{aligned}$$

También se daría el mismo hecho para algunos conjuntos aparentemente más extensos que \mathbb{N} , como \mathbb{Z} definiendo $f : \mathbb{N} \rightarrow \mathbb{Z}$ como:

$$f(x) = \begin{cases} \frac{x-2}{2} & \text{si } x \in \mathbb{P}, \\ \frac{-x-1}{2} & \text{si } x \in \mathbb{I} \end{cases}$$

O de una manera algo más complicada para la relación entre los naturales y los racionales. Definamos los conjuntos $A_j = \{i/j, i \in \mathbb{Z}\}$, es decir A_j es el conjunto de todos los números racionales no irreducibles con denominador j . Entonces es trivial comprobar que:

$$\mathbb{Q} = \bigcup_{i=1}^{\infty} A_i$$

de manera que la unión infinito-numerable de conjuntos finitos es infinito-numerable. Otra forma de demostración del mismo hecho ya se vió en la introducción del presente trabajo.

Definición 2.1.3. Sea un conjunto X . Se define el conjunto potencia de X ($\mathcal{P}(X)$) como el conjunto de todos los posibles subconjuntos de X incluyendo el conjunto vacío.

Es totalmente trivial ver que el conjunto potencia de un conjunto infinito es también infinito puesto que X está contenido en $\mathcal{P}(X)$.

Definición 2.1.4. Una partición de X es una familia de conjuntos $\{A_i\}_{i=1}^n$ que cumple las siguientes condiciones:

- $\emptyset \not\subseteq A_i$
- $\bigcup_{i=1}^n A_i = X$
- $\bigcap_{i=1}^n A_i = \emptyset$

Teorema de Cantor:

El conjunto potencia de cualquier conjunto X ($\mathcal{P}(X)$) tiene una cardinalidad estrictamente mayor que la cardinalidad de X . Concretamente como el número de subconjuntos de k elementos de un conjunto formado por N elementos viene dado por el número combinatorio $\binom{N}{k}$. Entonces es fácil comprobar que:

$$\text{Card}(\mathcal{P}(X)) = \sum_{k=1}^N \binom{N}{k} = 2^N$$

La siguiente hipótesis fue formulada por Cantor:

Hipótesis del continuo El cardinal del conjunto de los números reales es el inmediatamente superior al cardinal de los números naturales:

$$2^{\aleph_0} = \aleph_1$$

Esta hipótesis fue la que animó a David Hilbert a plantear como primer problema en su famosa lista la propuesta de encontrar una respuesta a la pregunta sobre si había algún conjunto infinito con una cardinalidad mayor que la de los naturales pero menor que la de los reales.

En 1940 Kurt Gödel demostró que la hipótesis del continuo era consistente con los principios axiomáticos de Zermelo-Fraenkel y que estos axiomas no pueden ni confirmar ni desaprobar la hipótesis del continuo incluso si el Axioma de Elección era tenido en cuenta.

Unos veinte años después Paul Cohen demostró la independencia de la hipótesis del continuo respecto a ZF+AE (Zermelo-Fraenkel y Axioma de Elección).

Para terminar este apartado nos parece interesante recordar el concepto de relación de equivalencia y conjunto cociente ya que aparecen habitualmente en la demostración de diversos resultados relacionados con la paradoja.

Definición 2.1.5. Una relación de equivalencia \sim es una relación binaria definida en un conjunto X que satisface las siguientes propiedades:

- Reflexiva: $\forall x \in X : x \sim x$
- Simétrica: $\forall x, y \in X, x \neq y: x \sim y \implies y \sim x.$
- Transitiva: $\forall x, y, z \in X, x \neq y \neq z: x \sim y, y \sim z \implies x \sim z.$

Proposición 2.1.6. *Sea el conjunto X . Entonces se cumple:*

- *Si \sim es una relación de equivalencia en X , entonces el conjunto de clases de equivalencia de X bajo \sim es una partición de X .*
- *Si $P = \{A_i\}_{i=1}^n$ es una partición de X entonces la relación en X definida por $x \sim y \iff x$ e y están en el mismo subconjunto A_i es una relación de equivalencia en X .*

2.2. Unas nociones de teoría de grupos

Definición 2.2.1. Un grupo es una estructura algebraica formada por un conjunto A no vacío y una ley de composición interna (\cdot) . La estructura se denota por (A, \cdot) . Para que el par (A, \cdot) sea un grupo, se deben cumplir las siguientes condiciones:

1. La ley u operación (\cdot) es interna, esto es, para cada par de elementos x e y de A , la composición $x \cdot y$ debe ser un elemento de A . Se conoce también como propiedad de clausura.
2. Asociatividad para la ley (\cdot) , esto es, para cualquier terna x, y y z debe cumplirse que $x \cdot (y \cdot z) = (x \cdot y) \cdot z$
3. Existe un elemento neutro e para \cdot ; esto es, para cualquier x de A , se cumple $x \cdot e = x$ y $e \cdot x = x$ siendo e el elemento neutro. Este elemento neutro es único.
4. Todo elemento x de A tiene simétrico y ; esto es, para todo elemento x de A existe un elemento y tal que $x \cdot y = e = y \cdot x$. El elemento simétrico de x es único.

Si en la definición anterior consideramos un conjunto X con una operación binaria (\cdot) que cumpla 1) y 2) diremos que (X, \cdot) es un semigrupo y si cumple 1), 2) y 3) se le definirá como un monoide.

Por ejemplo, es sencillo ver que el conjunto de los números naturales $(\mathbb{N}, +)$ es un semigrupo, pero no es un monoide ya que carece de elemento neutro.

Diremos además que el grupo es conmutativo o abeliano si además para todo $x, y \in G$ se cumple $x \cdot y = y \cdot x$. Veámos en primer lugar que un grupo ante todo es un conjunto, pero se va más allá en cuanto que se define una operación entre los elementos de ese conjunto y en la definición de ese conjunto interviene el cumplir ciertas propiedades respecto a la operación definida.

Algunos ejemplos de grupos son $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $GL(n, \mathbb{R})$ (grupo de matrices cuadradas $n \times n$ con determinante no nulo),...

A un subconjunto G' de un grupo G se le denomina subgrupo si mantiene la estructura de grupo. Si además es invariante por conjugación, esto es, si $\forall x \in G, y \in G', x \neq y, x \cdot y \cdot x^{-1} \in G'$ se le denomina subgrupo normal y se escribe como $G' \trianglelefteq G$. Es trivial comprobar que si G' es subgrupo normal, es abeliano.

Definición 2.2.2. (Grupo resoluble) Un grupo se dice resoluble si existe una cadena de grupos $\{G_i\}_{i=1}^n$ tal que:

$$\{\mathbb{I}_G\} = G_0 \subseteq G_1 \subseteq \cdots \subseteq G_n = G$$

tal que para cada $i = 0, 1, \dots, n - 1$:

1. G_i es subgrupo normal en G_{i+1} .
2. G_{i+1}/G_i es abeliano.

Por ser un conjunto de transformaciones que utilizaremos más adelante, pondremos especial hincapié en el subgrupo $SO_3(\mathbb{R})$ que es el conjunto de las matrices reales 3×3 ortogonales con determinante 1. Por la condición de ortogonalidad, $\forall A \in SO_3$, se verifica:

$$A^T = A^{-1}$$

Este subgrupo de simetrías representan el conjunto de transformaciones en \mathbb{R}^3 llamadas isometrías ya que conservan la distancia. Las isometrías más comunes son traslaciones, reflexiones y rotaciones y cualquiera de sus combinaciones.

Definición 2.2.3. Dados los grupos (A, \cdot) y (B, \circ) se denomina homomorfismo de grupos a la aplicación $\phi : A \rightarrow B$, si verifica que $\forall a \in A, b \in B, a \neq b$:

$$\phi(a \cdot b) = \phi(a) \circ \phi(b)$$

Concretamente si el conjunto imagen coincide con el de partida, es decir si $\phi : (G, \circ) \rightarrow (G, \circ)$, entonces ϕ es denominado automorfismo.

Acciones en grupos

Definición 2.2.4. (Definición de acción por la izquierda)

Supongamos que (G, \cdot) es un grupo y X un conjunto. Una acción por la izquierda del grupo G sobre X es una regla para combinar elementos $g \in G$ y elementos $x \in X$, denotada por $g \cdot x$. Requerimos además los siguientes 3 axiomas:

- (1) $g \cdot x \in X \quad \forall g \in G, x \in X$
- (2) $e \cdot x = x, \forall x \in X$
- (3) $g_1 \cdot (g_2 \cdot x) = (g_1 \cdot g_2) \cdot x \quad \forall x \in X \text{ y } g_1, g_2 \in G.$

Ejemplos

Ejemplo 2.2.1. Sea G el grupo de matrices 2×2 de la forma:

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}, \theta \in \mathbb{R}$$

bajo la ley de composición dictada por las reglas del producto matricial. Veamos que el grupo de rotaciones en $\mathbb{R}^2(SO_2(\mathbb{R}))$ es generado por matrices de este tipo. De tal manera que una matriz del grupo puede actuar por la izquierda sobre un vector columna y el resultado de la acción es otro vector columna rotado un ángulo θ respecto al primero en el sentido de las agujas del reloj.

Es trivial comprobar que es realmente una acción de grupo. De hecho la matriz identidad está en G (corresponde a $\theta=0$) y su acción se puede interpretar como una rotación de 0 grados la cual deja invariante el vector. Por otro lado, es claro que si $R_1, R_2 \in SO_2(\mathbb{R})$ se cumple: $R_1 \cdot (R_2 \cdot x) = (R_1 \cdot R_2) \cdot x$.

Ejemplo 2.2.2. Cada grupo actúa sobre sí mismo por multiplicación y por conjugación. En el primer caso al ser G un grupo automáticamente cumple las 3 condiciones establecidas en la definición 2.2.3. En el segundo caso, sea la acción definida por la conjugación, esto es, $g \cdot x = gxg^{-1}$. Logicamente la primera condición se cumple, la segunda también pues $e \cdot x = exe^{-1} = x$. Y la tercera también es satisfecha ya que:

$$g_1 \cdot (g_2 \cdot x) = g_1 \cdot (g_2 x g_2^{-1}) = g_1 (g_2 x g_2^{-1}) g_1^{-1} = (g_1 g_2) x (g_1 g_2)^{-1}$$

Definición 2.2.5. Sea una acción A de un grupo G sobre un conjunto X . Se dice que A es libre si $\forall x \in X$, $g \in G$ se verifica $g \cdot x = x \iff g = e$ (siendo e el elemento neutro del mismo grupo).

Ejemplo 2.2.3. Si consideramos el grupo de las rotaciones actuando sobre la esfera de inmediato, se comprueba que cualquier rotación de un ángulo θ respecto a un eje no es libre pues existen dos puntos fijos llamados polos pertenecientes al eje de la rotación. Si suponemos una esfera de radio 1 centrada en $(0, 0, 0)$ y consideramos la rotación de un ángulo θ alrededor del eje Z :

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \pm 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \pm 1 \end{pmatrix}$$

Definición 2.2.6. Sea G un grupo y un conjunto X sobre el que actúa por la izquierda. Dado $x \in X$ se define la órbita de x como el siguiente conjunto:

$$\mathcal{O}_x = \{g \cdot x \mid g \in G\}$$

Es trivial ver que $x \in \mathcal{O}_x$.

Definición 2.2.7. Supongamos que G es un grupo actuando por la izquierda sobre un conjunto X . Entonces si $x \in X$, se define el estabilizador de x , $Estab(G(x))$, como:

$$Estab(G(x)) = \{g \in G \mid g \cdot x = x\}$$

Trivialmente si la acción es libre $Estab(G(x)) = e$.

Proposición 2.2.8. *Supongamos que un grupo G actúa por la izquierda sobre un conjunto X . Entonces el conjunto de las distintas órbitas \mathcal{O}_x de G sobre X forman una partición de X .*

Demostración

Claramente $x \in \mathcal{O}_x$, pues $x = 1 \cdot x, \forall x \in X$. Luego $X = \bigcup_{x \in X} \mathcal{O}_x$. A continuación veremos que para $x, y \in X, x \neq y \implies \mathcal{O}_x \cap \mathcal{O}_y = \emptyset$. Sea $z \in \mathcal{O}_x \cap \mathcal{O}_y \neq \emptyset$. Entonces por pertenecer z a \mathcal{O}_x , existe $g \in G$ tal que $z = gx$. Por el mismo argumento, por pertenecer a \mathcal{O}_y existe $g' \in G$ tal que $z = g'y$, luego:

$$gx = g'y \implies x = (g)^{-1}g'y, y = (g')^{-1}gx$$

Luego $x \in \mathcal{O}_y, y \in \mathcal{O}_x \implies \mathcal{O}_x = \mathcal{O}_y$

Grupos libres

Definición 2.2.9. Un grupo F se dice libre si hay un subconjunto H de F , tal que todo elemento de F puede escribirse de manera única como producto de finitos elementos de H y sus inversos.

Utilizaremos el conjunto, finito o infinito, $H = a_1, a_2, \dots, a_n$ como la base de nuestro grupo libre F . A este conjunto a veces se le denomina alfabeto y a los elementos se les denomina generadores.

Una palabra en F es una secuencia finita de elementos que se puede escribir como:

$$\{a_1^{\epsilon_1} a_2^{\epsilon_2} \dots a_n^{\epsilon_n} \mid a_i \in F, \epsilon_i \in \mathbb{Z} \setminus \{0\}\}$$

Sea w una palabra de la forma definida anteriormente. A n se le denomina la longitud de la palabra y se designa por $|w|$.

Una palabra se dirá reducida si no contiene una subpalabra de los tipos $\{a_i a_i^{-1}, a_i^{-1} a_i\}$, es decir si a_i y a_i^{-1} nunca son adyacentes.

Si $X \subseteq G$ entonces cada palabra del grupo $w = a_1^{\epsilon_1} a_2^{\epsilon_2} \dots a_n^{\epsilon_n}$ determina un único elemento de G que es igual al producto $a_1^{\epsilon_1} a_2^{\epsilon_2} \dots a_n^{\epsilon_n}$ de los elementos $a_i^{\epsilon_i}$.

A continuación comentaremos la operación de reducción la cual permite obtener una palabra reducida a partir de una palabra arbitraria.

Una reducción elemental de una palabra w consiste en eliminar una subpalabra del tipo $a_i a_i^{-1} (a_i^{-1} a_i)$ de w .

Una operación de reducción sobre una palabra w consiste en la aplicación de una sucesión de operaciones de reducción elemental a w :

$$w_1 \rightarrow w_2 \dots \rightarrow w_n$$

En general, puede haber diferentes reducciones posibles de una palabra w . No obstante todas las posibles reducciones finalizan con la misma palabra reducida.

Ejemplo 2.2.4. Sea la palabra $w = a_1 a_2 a_2^{-1} a_3 a_4 a_4^{-1} a_5$, y sea la reducción λ_i que reduce un par $a_i a_i^{-1}$. Entonces:

$$w \xrightarrow{\lambda_2} a_1 a_3 a_4 a_4^{-1} a_5 \xrightarrow{\lambda_4} a_1 a_3 a_5$$

$$w \xrightarrow{\lambda_4} a_1 a_2 a_2^{-1} a_3 a_5 \xrightarrow{\lambda_2} a_1 a_3 a_5$$

Sobre el conjunto de palabras reducidas en F se puede definir la acción por la izquierda de una palabra $w = a_1^{\epsilon_1} a_2^{\epsilon_2} \dots a_n^{\epsilon_n}$ sobre otra $w' = a_1^{\delta_1} a_2^{\delta_2} \dots a_m^{\delta_m}$ de la siguiente forma:

$$ww' = a_1^{\epsilon_1} a_2^{\epsilon_2} \dots a_n^{\epsilon_n} a_1^{\delta_1} a_2^{\delta_2} \dots a_m^{\delta_m}$$

Es importante tener en cuenta que se puede dar el caso de que siendo w y w' reducidas ww' no lo sea. El caso genérico es cuando $w = w_1 w_2$ y $w' = w_2^{-1} w_3$. En este caso, el resultado de la acción queda definido como:

$$ww' = w_1 w_3$$

Es importante destacar que en un grupo libre existe un conjunto generador X de G tal que toda palabra reducida no vacía compuesta por elementos en X define un elemento no trivial de G .

En el contexto anterior, a X se le denomina base libre de G . De la definición se deduce que todo elemento de un grupo libre generado por X puede ser definido por una palabra reducida. Es más, diferentes palabras reducidas definen elementos diferentes de G .

Teorema 2.2.10. Existe un isomorfismo entre grupos libres si el cardinal de sus bases es el mismo. A ese cardinal se le denomina rango del grupo libre.

Un grupo especial. Grupo de las rotaciones

Aunque ya las introdujimos anteriormente, vamos a hablar en este apartado un poco más extensamente del subgrupo de rotaciones $SO_3(\mathbb{R})$ perteneciente al grupo de isometrías del espacio tridimensional euclídeo ($G_3(\mathbb{R})$).

Definición 2.2.11. Una transformación ortogonal f en un espacio euclídeo U es un endomorfismo que preserva el producto escalar:

$$f(\vec{x}) \cdot f(\vec{y}) = \vec{x} \cdot \vec{y} \quad \forall \vec{x}, \vec{y} \in U$$

Es inmediato ver que las transformaciones ortogonales conservan también la distancia entre dos puntos pues la distancia euclídea se define en función del producto escalar habitual. A este tipo de transformaciones se les conoce también como isometrías o movimientos rígidos y se denota como $G(n)$. Si nos circunscribimos al espacio euclídeo en 3 dimensiones (\mathbb{R}^3), una isometría del espacio es cualquier traslación, reflexión y rotación y sus posibles combinaciones. Las rotaciones además son un tipo de isometría que conserva la orientación.

La representación de una isometría en \mathbb{R}^3 generalmente suele ser una matriz ortogonal de dimensión 3×3 cuya inversa es igual a su traspuesta, esto es:

$$A^T = A^{-1}$$

De esto se deduce que $\det(A) = \pm 1$ puesto que, por las propiedades de los determinantes, tenemos:

$$\det(A^T) = \det(A) ; \det(A^{-1}) = 1/\det(A)$$

luego:

$$\det(A) = 1/\det(A) \implies \det(A) = \pm 1$$

Proposición 2.2.12. *El conjunto de las isometrías del espacio euclídeo \mathbb{R}^3 (G_3) tiene la estructura de grupo.*

Sean A, B subconjuntos de \mathbb{R}^3 . Se dice que A y B son congruentes si existe una isometría \mathcal{I} tal que $\mathcal{I}(A) = B$.

A partir de este momento nos centraremos en el estudio específico de las rotaciones y mencionaremos algunas de sus propiedades.

Definición 2.2.13. Una rotación R en un espacio euclídeo es una isometría que deja al menos un punto fijo.

Concretamente, si consideramos \mathbb{R}^3 , el conjunto de puntos fijos determina una recta (eje de la rotación). Las rotaciones conservan la distancia entre los puntos y su orientación. Esto último se deduce de que si R es una rotación, $\det(R) = +1$ (lo cual le convierte en una transformación directa).

Teorema 2.2.14. *Una matriz 3×3 describe una rotación alrededor de algún eje que pasa por el origen si y sólo si es ortogonal y tiene determinante 1.*

Demostración

\implies)

Primero demostraremos que cualquier matriz de rotación R es ortogonal con determinante $|R|=1$. Sea l el correspondiente eje de rotación y θ el ángulo de rotación alrededor del eje. Sabemos que una rotación al ser isometría, conserva el producto escalar de dos vectores, luego si \vec{v}_1, \vec{v}_2 son dos vectores de \mathbb{R}^3 se debe cumplir:

$$(R\vec{v}_1)^T \cdot (R\vec{v}_2) = \vec{v}_1^T \cdot \vec{v}_2$$

Luego:

$$\vec{v}_1^T (R^T R - I_{3 \times 3}) \vec{v}_2$$

Como \vec{v}_1, \vec{v}_2 han sido elegidos arbitrariamente, se tiene:

$$R^T R - I_3 = 0$$

Luego $R^T R = I_3$ y por lo tanto R es ortogonal. Además:

$$\det(R^T R) = \det(R^T) \det(R) = (\det(R))^2 = 1$$

por lo que $\det(R) = \pm 1$. Pero como una rotación conserva la orientación $\det(R) \neq -1$, luego $\det(R) = 1$.

\Longleftarrow)

Supongamos que R es una matriz ortogonal con determinante 1. Queremos demostrar que representa una rotación. Primero demostraremos que existe un vector \vec{u} tal que $R\vec{u} = \vec{u}$. Este vector estará en el eje de rotación. Bastará con demostrar que la matriz R tiene un autovalor 1, lo cual es verdad si y sólo si $\det(R - I_3) = 0$. Por lo tanto:

$$\det(R - I_3) = \det(R - I_3) = \det(R^T R - R^T) = \det(I_3 - R^T) = \det(I_3 - R)^T = \det(I_3 - R) = -\det(R - I_3)$$

Luego $\det(R - I_3) = 0$. A continuación mostraremos que el plano que contiene el origen y es ortogonal a \vec{u} se transforma en sí mismo bajo la rotación. Consideremos un vector arbitrario en este plano \vec{l} , es decir, $\vec{l}^T \cdot \vec{u} = 0$. El vector $R\vec{l}$ es ortogonal a \vec{u} ya que:

$$(R\vec{l})^T \vec{u} = \vec{l}^T R^T R \vec{u} = \vec{l}^T I \vec{u} = \vec{l}^T \vec{u} = 0$$

Luego finalmente el producto escalar de dos vectores es invariante bajo la transformación porque R es ortogonal y por lo tanto las longitudes de \vec{v}_1 y \vec{v}_2 y el ángulo entre ellos son conservados al aplicar R .

Representación básica de una matriz de rotación

Las siguientes matrices de rotación realizan rotaciones de vectores de ángulo θ en el sentido horario alrededor de los ejes X , Y , o Z , en \mathbb{R}^3 :

$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\operatorname{sen} \theta \\ 0 & \operatorname{sen} \theta & \cos \theta \end{pmatrix}$$

$$R_y(\theta) = \begin{pmatrix} \cos \theta & 0 & -\operatorname{sen} \theta \\ 0 & 1 & 0 \\ \operatorname{sen} \theta & 0 & \cos \theta \end{pmatrix}$$

$$R_z(\theta) = \begin{pmatrix} \cos \theta & -\operatorname{sen} \theta & 0 \\ \operatorname{sen} \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Aunque las representaciones anteriores son las más conocidas, sólo representan un caso bastante particular que es aquél en que los ejes de rotación coinciden con los ejes de la base ortonormal utilizada para describir el propio espacio. Se puede demostrar que la fórmula de la matriz de rotación para una rotación genérica de ángulo θ alrededor de un eje $\vec{n}(n_1, n_2, n_3)$ viene dada por la conocida fórmula de Rodrigues:

$$R_{ij}(\vec{n}, \theta) = \cos \theta \delta_{ij} + (1 - \cos \theta) n_i n_j - \sin \theta \epsilon_{ijk} n_k$$

donde δ_{ij} y ϵ_{ijk} son dos conocidos objetos tensoriales: la delta de Kronecker y el símbolo de Levi-Civita respectivamente.

A continuación estudiaremos los autovectores y autovalores de una matriz de rotación genérica en \mathbb{R}^3 los cuales nos darán información sobre la misma rotación. Trabajaremos con la matriz de rotación $R_z(\theta)$. Podemos hacer esta elección dado que es demostrable que el conjunto de autovalores de cualquier matriz es invariante frente a cualquier cambio de base. Realizando un sencillo cálculo, obtenemos el polinomio característico de $R_z(\theta)$:

$$(1 - \lambda)[(\cos \theta - \lambda)^2 + \sin^2 \theta] = 0$$

La resolución de la ecuación anterior nos proporciona:

$$\lambda_1 = 1, \lambda_2 = e^{i\theta}, \lambda_3 = e^{-i\theta}, 0 \leq \theta \leq \pi$$

El valor $\lambda_1=1$ corresponde al subespacio invariante correspondiente al eje de rotación y los valores complejos representan el subespacio de dimensión 2 correspondiente al plano que pasa por el centro de la rotación y es perpendicular al eje de la misma.

Podemos obtener un resultado importante sobre las rotaciones gracias al estudio de los anteriores autovalores.

Proposición 2.2.15. *Dos rotaciones R_1, R_2 conmutan si y sólo si representan rotaciones alrededor del mismo eje.*

Demostración

Consideremos dos rotaciones en el mismo eje $R_1=R(\vec{n}, \theta_1), R_2=R(\vec{n}, \theta_2)$. Al tener el eje de rotación en común, el autovector correspondiente al autovalor $\lambda=1$ es común. Es decir:

$$R_1\vec{n} = \vec{n}$$

$$R_2\vec{n} = \vec{n}$$

Veamos que ambas rotaciones conmutan puesto que:

$$R_1R_2\vec{n} = R_1\vec{n} = \vec{n}$$

$$R_2R_1\vec{n} = R_2\vec{n} = \vec{n}$$

Es sencillo deducir que el caso general es que en \mathbb{R}^3 los movimientos de rotación no suelen conmutar. Este hecho es importante porque permite la formación de grupos libres de rotaciones con más de un generador, lo cual jugará un papel esencial en el desarrollo de la demostración de la paradoja que se verá más adelante.

2.3. Una breve introducción a la teoría de la medida

Introducción histórica

La idea de medida es un concepto cuyos orígenes históricos se remontan a hace más de 4000 años con la necesidad del cálculo de longitudes, áreas y volúmenes en campos diversos. En el denominado Papiro de Moscú, que data del año 1890 a. d. C. y es el documento más importante sobre matemáticas del Antiguo Egipto ya aparece el planteamiento y la resolución del problema consistente en calcular el volumen de un tronco de una pirámide cuadrangular. Por supuesto en el más famoso libro de geometría de la antigüedad, los Elementos de Euclides aparecen también teoremas sobre áreas y volúmenes y es donde por primera vez se utiliza la palabra medir.

Distintas aportaciones se van sucediendo posteriormente hasta que en la década de 1880 George Cantor propone la primera definición de medida de un conjunto acotado en \mathbb{R}^n . También Harnack proporciona una definición similar pero en \mathbb{R} .

Según estas definiciones una de las propiedades que caracterizaba a la medida era su aditividad, esto es, que si A, B son tales que $A \cap B = \emptyset$ entonces $\mu(A \cup B) = \mu(A) + \mu(B)$. Sin embargo de las mismas definiciones surge la equivalencia entre la medida de un conjunto y la de su clausura. Esto originaba un problema con cierto tipo de conjuntos. Si se aplica esta última equivalencia a los conjuntos $A = \mathbb{Q} \cap [0, 1]$ y $B = (\mathbb{R} \setminus \mathbb{Q}) \cap [0, 1]$ (teniendo en cuenta que $\overline{A} = \overline{B} = \mathbb{R}$) tenemos:

$$\mu(A \cup B) = \mu(A) + \mu(B) = \mu(\overline{A}) + \mu(\overline{B}) = 2\mu(\mathbb{R} \cap [0, 1])$$

Pero por otro lado como $A \cup B = [0, 1] \implies \mu(A \cup B) = 1$ lo cual es una contradicción.

El desarrollo de las aportaciones que van surgiendo a partir de este momento se centran en intentar ampliar la noción de conjunto medible a una familia cada vez mayor de conjuntos.

En 1892 Camille Jordan, basándose en ideas ya propuestas unos años antes por Giuseppe Peano, propone dos tipos de medida de un conjunto, medida exterior e interior, definiendo conjunto medible como aquellos conjuntos para los cuales ambas medidas coinciden. Sin embargo esta medida continúa no ofreciendo una solución para la medida de conjuntos como los racionales.

Unos años más tarde, Émile Borel extiende la idea de conjunto medible a cualquier conjunto que se pueda formar con conjuntos abiertos mediante ciertas operaciones (unión e intersección finita o infinito-numerable, complementariedad,..). Dentro de este nuevo contexto demuestra que $\mathbb{Q} \cap [0, 1]$ se puede expresar como una unión infinito numerable de conjuntos abiertos cuya medida individual tiende a cero.

De esta manera la medida introducida por Borel también verifica que es numerablemente aditiva, es decir:

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

Esta extensión introducida por Borel permitió a Lebesgue introducir la medida que lleva su nombre y que es la medida más ampliamente aceptada y utilizada desde entonces. Parte de la noción de volumen de un rectángulo n -dimensional y amplía este concepto por medio del teorema de Hahn-Kolmogorov.

A pesar de esta evolución en el desarrollo de la idea de medida para tratar abarcar cada vez más conjuntos, cabe destacar la existencia de conjuntos que continúan sin ser medibles, no medibles Lebesgue, como el conjunto de Vitali del cual hablaremos al final de este apartado.

Conceptos relacionados

Definición 2.3.1. Sea X un conjunto arbitrario. Una familia \mathcal{A} de subconjuntos de X es un álgebra en X si:

1. $X \in \mathcal{A}$
2. Para cada conjunto A que pertenece a \mathcal{A} , el conjunto A^c pertenece a \mathcal{A} .

3. Para cada secuencia finita A_1, \dots, A_n de conjuntos pertenecientes a \mathcal{A} , el conjunto $\bigcup_{i=1}^n A_i$ pertenece a \mathcal{A} .

Si trasladamos la misma definición a la combinación de conjuntos infinitos, obtenemos la definición de σ -álgebra:

Definición 2.3.2. Sea X un conjunto arbitrario. Una familia \mathcal{A} de subconjuntos de X es una σ -álgebra en X si:

1. $X \in \mathcal{A}$
2. Para cada conjunto A que pertenece a \mathcal{A} , el conjunto $A^c (X \setminus A)$ pertenece a \mathcal{A} .
3. Para cada secuencia finita A_1, \dots, A_n de conjuntos pertenecientes a \mathcal{A} , el conjunto $\bigcup_{i=1}^{\infty} A_i$ pertenece a \mathcal{A} .

Veamos que si se tienen en cuenta las leyes de Morgan para $A_1, \dots, A_n \in X$, se verifica:

$$X \setminus \left(\bigcup_{i=1}^{\infty} A_i \right) = \bigcap_{i=1}^{\infty} X \setminus A_i$$

Luego de los apartados b) y c) de la Definición 2.3.2 se deduce que para $A_1, \dots, A_{\infty} \in \mathcal{A}$, σ -álgebra:

$$\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$$

De igual manera podemos decir que si \mathcal{A} es un álgebra y $A_1, \dots, A_n \in \mathcal{A}$, entonces:

$$\bigcap_{i=1}^n A_i \in \mathcal{A}$$

Por otro lado, es trivial ver que toda σ -álgebra es un álgebra.

Ejemplos

Ejemplo 2.3.1. Sea un conjunto X . El conjunto potencia de X ($\mathcal{P}(X)$) es una σ -álgebra.

Ejemplo 2.3.2. Sea un conjunto X y sea \mathcal{A} la familia de todos los subconjuntos A de X tal que A o A^c es numerable. Entonces \mathcal{A} es una σ -álgebra.

Ejemplo 2.3.3. Sea \mathcal{A} la familia de todos los subconjuntos de \mathbb{R} que son uniones finitas de intervalos de la forma $(a, b]$, $(a, +\infty)$ o $(-\infty, b]$. En la definición del propio conjunto ya se puede ver que cumple la condición de que las uniones finitas de conjuntos pertenecientes a \mathcal{A} pertenecen a \mathcal{A} . De igual modo es muy sencillo ver que las intersecciones finitas de conjuntos de \mathcal{A} también están en \mathcal{A} y el conjunto vacío está en \mathcal{A} ya que es la unión de una colección vacía (por tanto finita) de intervalos. Por todo lo anterior \mathcal{A} es un álgebra sobre \mathbb{R} . Sin embargo veamos que \mathcal{A} no es una σ -álgebra ya que los subintervalos abiertos acotados de \mathbb{R} , aunque pueden ser descritos por uniones finitas de intervalos pertenecientes a \mathcal{A} , ellos mismos no pertenecen a \mathcal{A} .

Ahora utilizaremos las definiciones anteriores para introducir una familia importante de σ -álgebras: la σ -álgebra de Borel en \mathbb{R}^n generada por la familia de subconjuntos abiertos de \mathbb{R}^n y la denotaremos por $\mathcal{B}(\mathbb{R}^n)$.

Proposición 2.3.3. *La σ -álgebra $\mathcal{B}(\mathbb{R})$ de los subconjuntos Borel de \mathbb{R} es generada por la siguiente familia de conjuntos:*

- a) *La familia de todos los subconjuntos cerrados de \mathbb{R} .*
- b) *La familia de todos los subintervalos de \mathbb{R} de la forma $(\infty, b]$.*
- c) *La familia de todos los subintervalos de la forma $(a, b]$.*

Demostración

Sea $\mathcal{B}_1, \mathcal{B}_2$ y \mathcal{B}_3 las σ -álgebras generadas por los conjuntos a los que se hace referencia en los apartados a), b) y c) respectivamente. Demostraremos que $\mathcal{B}_3 \subseteq \mathcal{B}_2 \subseteq \mathcal{B}_1 \subseteq \mathcal{B}(\mathbb{R})$, y por lo tanto $\mathcal{B}_3 \subseteq \mathcal{B}(\mathbb{R})$. Ya que $\mathcal{B}(\mathbb{R})$ contiene la familia de los subconjuntos abiertos de \mathbb{R} por ser σ -álgebra y por ser los subconjuntos cerrados complementarios de los abiertos, también contiene a éstos últimos, esto es, \mathcal{B}_1 . Los conjuntos de la forma $(\infty, b]$ también son cerrados y por tanto también pertenecen a \mathcal{B}_1 ($\mathcal{B}_2 \subseteq \mathcal{B}_1$). Y como $(a, b] = (\infty, b] \cap (\mathbb{R} \setminus (-\infty, a])$ los conjuntos de esta forma pertenecen a \mathcal{B}_2 . Luego $\mathcal{B}_3 \subseteq \mathcal{B}_2$. Finalmente, observamos que todo intervalo abierto de \mathbb{R} es la unión de conjuntos de la forma $(a, b]$ y que la unión de conjuntos abiertos es un conjunto abierto. Luego cada abierto de \mathbb{R} pertenece a \mathcal{B}_3 y por tanto $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{B}_3$.

Ampliaremos la definición anterior a $\mathcal{B}(\mathbb{R}^n)$.

Proposición 2.3.4. *La σ -álgebra $\mathcal{B}(\mathbb{R}^n)$ de los subconjuntos Borel de \mathbb{R}^n es generada por la siguiente familia de conjuntos:*

- a) *La familia de todos los subconjuntos cerrados de \mathbb{R}^n .*
- b) *La familia de todos los subintervalos semicerrados en \mathbb{R}^n que tienen la forma $\{(x_1, \dots, x_n) : x_i \leq b_i, i = 1, \dots, n\}$ para algún índice i y $b_i \in \mathbb{R}$.*
- c) *La familia de todos los rectángulos semicerrados de la forma $\{(x_1, \dots, x_n) : a_i < x_i \leq b_i, i = 1, \dots, n\}$*

La demostración es similar a la utilizada anteriormente para \mathbb{R} . De hecho la σ -álgebra generada por los conjuntos del apartado c) está incluida en la σ -álgebra generada por los conjuntos del apartado b) considerando diferencias entre los conjuntos de esta última σ -álgebra. De igual manera cada uno de los rectángulos del conjunto del apartado c) puede obtenerse a través de intersecciones de conjuntos de la familia del apartado b).

Definición 2.3.5. Sea X un conjunto y \mathcal{A} una σ -álgebra en X . Una medida en \mathcal{A} es una función $\mu: \mathcal{A} \rightarrow [0, +\infty)$ que satisface:

a) $\mu(\emptyset)=0$.

b) $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ para toda familia infinita $\{A_i\}$ de conjuntos disjuntos dos a dos pertenecientes a \mathcal{A} .

La segunda propiedad indica que μ es numerablemente aditiva.

De igual manera podemos definir una medida finitamente aditiva sobre el álgebra \mathcal{A} (no tiene por qué ser una σ -álgebra) como una función μ con dominio en \mathcal{A} y recorrido en $[0, +\infty)$ que satisface $\mu(\emptyset)=0$ y es finitamente aditiva, esto es:

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

para toda familia finita de conjuntos disjuntos pertenecientes a \mathcal{A} . Claramente toda medida numerablemente aditiva es finitamente aditiva.

Si existe un conjunto X , siendo \mathcal{A} una σ -álgebra definida sobre el mismo y μ una medida en \mathcal{A} , entonces a (X, \mathcal{A}, μ) se le denomina espacio de medida.

Ejemplos

Ejemplo 2.3.4. Sea X un conjunto y \mathcal{A} una σ -álgebra sobre X . Sea la función:

$$\mu(A) = \begin{cases} n, & \text{si } A \text{ es un conjunto finito con cardinal } n \\ \infty & \text{si } A \text{ es un conjunto infinito} \end{cases}$$

con $A \in \mathcal{A}$

La función así definida es una medida sobre \mathcal{A} .

Ejemplo 2.3.5. Sea X un conjunto que tiene al menos dos miembros y \mathcal{A} la σ -álgebra que consta de todos los subconjuntos de X . Definamos la función:

$$\mu(A) = \begin{cases} 0 & \text{si } A = \emptyset \\ 1 & \text{si } A \neq \emptyset \end{cases}$$

con $A \in \mathcal{A}$.

En este caso μ no es una medida, e incluso no es finitamente aditiva. Veamos que si $A_1, A_2 \in \mathcal{A}$, $\mu(A_1 \cup A_2) = 1$ y sin embargo $\mu(A_1) + \mu(A_2) = 2$.

A continuación veremos ciertas técnicas para construir medidas lo cual nos servirá para construir la medida de Lebesgue en \mathbb{R}^n .

Definición de medida exterior

Sea X un conjunto y sea $\mathcal{P}(X)$ la familia de todos los subconjuntos de X . Una medida exterior en X es una función $\mu^*: \mathcal{A} \rightarrow [0, +\infty)$ tal que:

- a) $\mu^*(\emptyset) = 0$.
- b) Si $A \subseteq B \subseteq X$, entonces $\mu^*(A) \leq \mu^*(B)$.
- c) Si $\{A_n\}$ es una sucesión numerable de subconjuntos de X , entonces $\mu^*(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu^*(A_i)$ (medida numerablemente subaditiva).

Aunque una medida puede no ser una medida exterior se puede demostrar que para toda medida exterior μ^* en X hay una σ -álgebra $\mathcal{M}_\mu(X)$ tal que la restricción de μ^* a $\mathcal{M}_\mu(X)$ es numerablemente aditiva y por lo tanto una medida.

Ejemplos

Ejemplo 2.3.6. Sea X un conjunto y $\mathcal{P}(X)$ su conjunto potencia. Sea la función:

$$\mu^*(A) = \begin{cases} 0 & \text{si } A = \emptyset \\ 1 & \text{si } A \neq \emptyset \end{cases}$$

En este caso μ^* es una medida exterior

Ejemplo 2.3.7. Sea X un conjunto y $\mathcal{P}(X)$ su conjunto potencia. Sea la función:

$$\mu^*(A) = \begin{cases} 0 & \text{si } A \text{ es un conjunto finito} \\ 1 & \text{si } A \text{ es un conjunto infinito} \end{cases}$$

Obviamente μ^* no es una medida exterior pues no es numerablemente aditiva.

Definición de intervalo n -dimensional

Un intervalo n -dimensional es un subconjunto de \mathbb{R}^n de la forma $I_1 \times \cdots \times I_n$ donde I_1, \dots, I_n son intervalos de \mathbb{R} y $I_1 \times \cdots \times I_n$ viene dado por:

$$I_1 \times \cdots \times I_n = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_i \in I_i, i = 1, \dots, n\}$$

Es importante comentar que los intervalos $I_1 \times \cdots \times I_n$ (y por lo tanto el intervalo n -dimensional $I_1 \times \cdots \times I_n$) pueden ser abiertos, cerrados o ni abiertos ni cerrados. El volumen del intervalo n -dimensional $I_1 \times \cdots \times I_n$ es el producto de las longitudes de los intervalos I_1, \dots, I_n y se denota por $\text{vol}(I_1 \times \cdots \times I_n)$.

Definición de medida exterior de Lebesgue

Sea un subconjunto A de \mathbb{R}^n y sea $\mathcal{C}_A(X)$ el conjunto de todas las secuencias numerables tales que $A \subseteq \bigcup_i X_i$, se define medida exterior de Lebesgue de A , y se denota por $\lambda^*(A)$, como el ínfimo del conjunto:

$$\left\{ \sum_i \text{vol}(X_i) : \{X_i\} \in \mathcal{C}_A(X) \right\}$$

Introduciremos ahora la definición de conjunto medible Lebesgue.

Definición de conjunto medible Lebesgue (en el sentido de Carathéodory)

Sea X un conjunto y sea λ^* la medida exterior de Lebesgue definida sobre X . Un subconjunto B de X es λ^* -medible (medible Lebesgue) si:

$$\lambda^*(A) = \lambda^*(A \cap B) + \lambda^*(A \cap B^c)$$

para todo subconjunto A de X . Luego dado un conjunto medible A , $\lambda(A) = \lambda^*(A)$ es la medida Lebesgue de A .

Se puede demostrar que todo subconjunto Borel de \mathbb{R}^n es medible Lebesgue, de hecho a la restricción de la medida exterior de Lebesgue a $\{\mathcal{B}(\mathbb{R}^n)\}$ se le denomina medida Lebesgue. Además la medida de Lebesgue es la única en $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ que asigna a cada intervalo n -dimensional del modo anteriormente descrito su volumen.

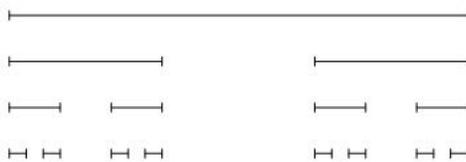
Proposición 2.3.6. *La medida de Lebesgue en \mathbb{R}^n es invariante por traslaciones, esto es, si $x \in \mathbb{R}^n$ y $A \subseteq \mathbb{R}^n$, entonces:*

$$\lambda^*(A) = \lambda^*(A+x)$$

Pudiera parecer que la medida Lebesgue es la medida más completa y que la familia de los conjuntos medibles Lebesgue pudiera abarcar todos los conjuntos de \mathbb{R}^n pero no es así. A continuación describiremos dos conjuntos importantes, el primero porque proporciona un resultado hasta cierto punto paradójico y el segundo porque es un ejemplo de un conjunto no medible Lebesgue.

Conjunto de Cantor

Seguidamente describiremos los pasos en la construcción de este conjunto. Llamemos A_0 al intervalo de \mathbb{R} $[0,1]$. El intervalo A_1 se obtendrá de A_0 eliminando el intervalo $(\frac{1}{3}, \frac{2}{3})$ obteniendo $A_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. Para obtener A_2 de igual manera dividimos en 3 cada uno de los intervalos en que se divide A_1 y eliminando su intervalo central, proporcionando $A_3 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$. Si continuamos este procedimiento, observamos que A_n es la unión de 2^n intervalos disjuntos de igual longitud, $(\frac{1}{3})^n$ cada uno. Luego finalmente el conjunto de Cantor es $A = \bigcap_{n=1}^{\infty} A_n$.



Intuitivamente podemos observar que el conjunto de Cantor finalmente contendrá todos los puntos frontera de los intervalos que se han ido generando en los infinitos pasos de su construcción:

$$0, \frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{2}{9}, \frac{7}{9}, \frac{8}{9}, \frac{1}{27}, \frac{2}{27}, \frac{7}{27}, \frac{8}{27}, \frac{1}{81}, \frac{2}{81}, \frac{7}{81}, \frac{8}{81}, \dots$$

Si $\{d_i : i = 1, \dots, \infty\}$ es una secuencia infinita de ceros y unos, entre el conjunto de todas las posibles secuencias a_i (llamémoslo \mathcal{D}) y el conjunto de Cantor se puede construir la biyección $f: \mathcal{D} \rightarrow A$ definida por:

$$\sum_{i=1}^{\infty} \frac{2d_i}{3^i}$$

Como la cardinalidad del conjunto \mathcal{D} es la del continuo, entonces la cardinalidad del conjunto de Cantor es la misma.

Proposición 2.3.7. *El conjunto de Cantor es un conjunto con la cardinalidad del continuo y con medida Lebesgue cero.*

Demostración

El primer hecho ya se ha demostrado antes. Por otro lado veamos que para cada A_n , $\lambda(A_n) = (\frac{2}{3})^n$. Luego $\lambda(A_n) \leq (\frac{2}{3})^n \forall n$ y por lo tanto: $\lambda(A) = \lambda(\bigcap_{n=1}^{\infty} A_n) = 0$

Conjunto de Vitali(AE)

Definamos una relación \sim para x, y en \mathbb{R} de tal manera que $x \sim y$ si $x - y$ es racional. Es sencillo comprobar que cumple las tres propiedades de las relaciones de equivalencia (reflexiva, simétrica y transitiva). Cada clase de equivalencia tendrá la forma $\mathbb{Q} + x$ para algún x y será densa en \mathbb{R} .

Como las clases de equivalencia son disjuntas y cada una tiene intersección no vacía con el intervalo $(0, 1)$, utilizando el Axioma de Elección formaremos un conjunto R que contenga un elemento (representante) de cada clase de equivalencia. Demostraremos que ese conjunto es no medible Lebesgue.

Sea r_n el conjunto de los números racionales en $(-1, 1)$ y para cada n sea $R_n = R + r_n$. Demostraremos que:

a) los conjuntos R_n son disjuntos

b) $\bigcup_{i=1}^n R_n \subseteq (-1, 2)$

$$c) (0, 1) \subseteq \bigcup_{i=1}^n R_n$$

Para demostrar a) veamos que si $R_i \cap R_j \neq \emptyset$ entonces hay elementos r, r' tales que $r + r_i = r' + r_j$ luego $r \sim r'$ lo cual implica $r = r'$. b) se deduce de que $R \subseteq (0, 1)$ y de la definición de r_n . Para demostrar c), sea x un miembro de $(0, 1)$ y sea r el miembro de R tal que $r \sim x$. Es claro que $r - x$ es racional y pertenece a $(0, 1)$. Supongamos que R es medible Lebesgue. Así como cada R_n es medible y por ser la medida Lebesgue numerablemente aditiva se tiene:

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

La invariancia traslacional de la medida Lebesgue también implica por otro lado $\lambda(R_n) = \lambda(R)$. Luego si $\lambda(R) = 0$, entonces $\mu\left(\bigcup_{i=1}^n R_i\right) = 0$, en contradicción con c), pero si $\lambda(R) \neq 0$ entonces $\mu\left(\bigcup_{i=1}^n R_i\right) = +\infty$ hecho que contradice el apartado b).

Deducimos finalmente que R no es medible Lebesgue.

Como hemos mencionado, en la demostración de la existencia del anterior conjunto se ha utilizado la posibilidad de construirlo utilizando el Axioma de Elección.

En este sentido, en los años 70 el matemático americano Robert M. Solovay elaboró un modelo en el que utilizando todos los principios axiomáticos del sistema Zermelo-Fraenkel, a excepción del Axioma de Elección, demostró que todos los conjuntos de \mathbb{R}^n son medibles Lebesgue.

la Paradoja de Banach Tarski

3.1. Conjuntos paradójicos

Comenzaremos con una definición formal que nos propone la idea de duplicar un conjunto usando ciertas transformaciones.

Definición 3.1.1. Sea G un grupo que actúa sobre un conjunto X y supongamos $E \subseteq X$ un subconjunto no vacío de X . Diremos que E es paradójico respecto a G si existen m, n enteros positivos tales que para subconjuntos disjuntos dos a dos de E , $A_1, \dots, A_m, B_1, \dots, B_n$, y $g_1, \dots, g_m, h_1, \dots, h_n \in G$ tenemos que:

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

Informalmente hablando el conjunto E consta de dos subconjuntos disjuntos A_i, B_j cada uno de los cuales pueden separarse y reordenarse mediante G generando todo E . Veamos que los conjuntos que pueden hacer que E sea G -paradójico ($g_i(A_i), h_j(B_j)$ y A_i, B_j) pueden ser elegidos de manera que los conjuntos $\{g_i(A_i), h_j(B_j)\}$ y $\{A_i \cup B_j\}$ sean cada uno particiones de E . Para el primer caso sólo hay que sustituir A_i, B_j por conjuntos más pequeños para asegurar la disyunción por pares de $g_i(A_i), h_j(B_j)$. Sin embargo la demostración del segundo caso es algo más complicada y la veremos más adelante.

Podemos ampliar la definición anterior permitiendo la utilización de conjuntos numerables de piezas. Esto es, se dirá que E es numerablemente paradójico si existen dos familias de subconjuntos disjuntos de $E, \{A_i\}_{i=1}^{\infty}, \{B_i\}_{i=1}^{\infty}$, y $\{g_i\}_{i=1}^{\infty}, \{h_i\}_{i=1}^{\infty} \in G$ tal que se verifica:

$$E = \bigcup_{i=1}^{\infty} g_i(A_i), E = \bigcup_{i=1}^{\infty} h_i(B_i)$$

Es claro que todo conjunto finitamente paradójico es numerablemente paradójico. Si tenemos que E es finitamente paradójico respecto a $\{A_i\}_{i=1}^m, \{B_i\}_{i=1}^n$, y $\{g_i\}_1^m, \{h_i\}_1^n \in G$, basta elegir:

$$\{A'_i\}_1^\infty = \left\{ \bigcup_{i=1}^m \{A_i\} \right\} \cup \left\{ \bigcup_{i=m+1}^\infty \{\emptyset\} \right\}$$

$$\{B'_i\}_1^\infty = \left\{ \bigcup_{i=1}^n \{B_i\} \right\} \cup \left\{ \bigcup_{i=n+1}^\infty \{\emptyset\} \right\}$$

entonces E es numerablemente paradójico:

$$E = \bigcup_{i=1}^\infty g_i(A'_i), E = \bigcup_{i=1}^\infty h_j(B'_j)$$

Observando estas definiciones y particularizando a acciones de grupo que son isometrías, es sencillo comprobar dónde está lo paradójico ya que, por un lado, como $E \supseteq \left(\bigcup_{i=1}^m A_i \right) \cup \left(\bigcup_{i=1}^n B_i \right)$:

$$\mu(E) \geq \sum_{i=1}^m \mu(A_i) + \sum_{j=1}^n \mu(B_j)$$

Pero:

$$\mu(E) = \sum_{i=1}^m \mu(g_i(A_i)) = \sum_{j=1}^n \mu(h_j(B_j)) = \sum_{i=1}^m \mu(A_i) = \sum_{i=1}^n \mu(B_i)$$

Como ya mencionamos en la introducción del presente trabajo en la demostración de la paradoja de Banach-Tarski y en la demostración de que otros conjuntos son paradójicos suele ser habitual emplear el Axioma de Elección. A partir de este momento cada vez que se exponga una demostración en la que se utilice este Axioma, lo indicaremos con el acrónimo AE.

Veremos a continuación ejemplos de conjuntos paradójicos.

Ejemplos

Ejemplo 3.1.1. *El conjunto de los números naturales \mathbb{N} es G -paradójico siendo G el grupo formado por las biyecciones $g: \mathbb{N} \rightarrow \mathbb{N}$*

Ejemplo 3.1.2. *Un grupo libre F de rango 2 es F -paradójico, donde F actúa sobre sí mismo mediante la multiplicación por la izquierda.*

Demostración

Supongamos que σ, τ son generadores libres de F . Si ρ es igual a $\sigma^{\pm 1}, \tau^{\pm 1}$, sea $W(\rho)$ el conjunto de elementos de F cuya representación es una palabra formada por $\sigma^{\pm 1}, \tau^{\pm 1}$ y que comienza por la izquierda con ρ .

Entonces $F = \{e\} \cup W(\sigma) \cup W(\sigma^{-1}) \cup W(\tau) \cup W(\tau^{-1})$ y estos subconjuntos son disjuntos por pares. Veamos que si $h \in F \setminus W(\sigma)$, entonces $\sigma^{-1}h \in W(\sigma^{-1})$ y por lo tanto $h = \sigma(\sigma^{-1}h) \in \sigma W(\sigma^{-1})$. El mismo razonamiento podemos utilizar para $g \in F \setminus W(\tau)$. Luego finalmente:

$$W(\sigma) \cup \sigma W(\sigma^{-1}) = F, \text{ y}$$

$$W(\tau) \cup \tau W(\tau^{-1}) = F$$

Es conveniente observar que la demostración anterior utiliza cinco piezas. Más adelante (ver Teorema 3.5.2) se encontrará una partición del mismo grupo libre en cuatro piezas.

Ejemplo 3.1.3. *Un semigrupo libre $S \subseteq F$, con generadores libres σ y τ , contiene dos conjuntos disjuntos A y B tales que $\sigma S = A$ y $\tau S = B$, luego es F -paradójico.*

Demostración

Sea A el conjunto de palabras en σ, τ (sin inversos) cuyo término más a la izquierda es σ y B el conjunto cuyo primer término es τ . Entonces $\sigma S = A$ y $\tau S = B$. Si S pertenece al grupo F , entonces S es un subconjunto paradójico del grupo porque $\sigma^{-1}(A) = S = \tau^{-1}(B)$ siendo $\sigma^{-1}, \tau^{-1} \in S$.

A continuación mencionaremos algunos ejemplos de paradojas geométricas:

Ejemplo 3.1.4. *Hay un subconjunto paradójico no vacío en el plano \mathbb{R}^2 .*

La demostración se realizará en dos pasos (Teorema 3.1.2 y Proposición 3.1.3):

Teorema 3.1.2. *Existen dos isometrías, ρ, τ , de \mathbb{R}^2 que generan un subsemigrupo libre de G_2 . Además, ρ y τ pueden elegirse de modo que para dos palabras cualesquiera ω_1 y ω_2 en ρ, τ , que tienen términos más a la izquierda ρ, τ , respectivamente, $\omega_1(0, 0) \neq \omega_2(0, 0)$, siendo $(0, 0)$ el origen en \mathbb{R}^2 .*

Demostración

Elijamos θ de forma que $e^{i\theta}$ sea trascendente. Se puede demostrar que hay una cantidad infinita numerable de números algebraicos en el círculo unidad, luego tal θ existe.

Entonces sea ρ la rotación de ángulo θ ($\rho z = e^{i\theta} z$) y que τ sea la traslación por $(1, 0)$ ($z = z + 1$). Sólo necesitamos demostrar que ρ y τ satisfacen la segunda afirmación, esto es, $\omega_1(0, 0) \neq \omega_2(0, 0)$. En el caso $\omega_1 = \omega_2$, siendo ω_1 y ω_2 palabras distintas del semigrupo y si una de ellas es (la identidad o) un segmento inicial de la otra, entonces la cancelación por la izquierda genera $\omega' = e$ (elemento neutro) para una palabra no trivial ω' . Pero entonces $\omega' \tau(0) = \tau(0)$, y $\omega' \rho(0) = \rho(0)$, contradiciendo la segunda afirmación en cualquier caso. En el caso de que ninguna de las dos sea un segmento inicial de la otra, entonces la cancelación por la izquierda da lugar a ω_1 y ω_2 , que son iguales en G_2 pero tienen diferentes términos izquierdos.

Luego, supongamos que $\omega_1 = \tau^{j_1} \rho^{j_2} \dots \tau^{j_m}$ y $\omega_2 = \tau^{k_1} \rho^{k_2} \dots \tau^{k_l}$, donde $m, l \geq 1$ y cada exponente es un número entero positivo. Como $\rho(0) = 0$ (ya que partimos de que al menos una de las palabras no actúa trivialmente), es correcto asumir que ω_1 y ω_2 terminan ambas en una potencia de τ a menos que ω_2 sea simplemente ρ^{k_1} . Luego:

- $\omega_1(0) = j_1 + j_3 u^{j_2} + j_5 u^{j_2+j_4} + \dots + j_{2m-1} u^{j_2+j_4+\dots+j_{m-1}}$
- $\omega_2(0) = k_2 u^{k_1} + k_4 u^{k_1+k_3} + \dots + k_l u^{k_1+k_3+\dots+k_{l-1}} = 0$ (si $\omega_2 = \rho^{k_1}$)

Si $\omega_1(0) = \omega_2(0)$, estas dos expresiones pueden restarse y se obtiene un polinomio no constante con coeficientes enteros que se hace cero para el valor $e^{i\theta}$, pero esto contradice la elección de θ .

Proposición 3.1.3. *Sea un grupo G que actúa sobre X y que contiene a ρ y τ y sean dos palabras cualesquiera ω_1, ω_2 en τ y ρ comenzando por τ y ρ respectivamente. Si se verifica:*

$$\omega_1 x = \omega_2 x$$

Entonces hay un subconjunto G -paradójico no vacío de X .

Demostración

Sea S el semigrupo de G generado por ρ y τ y sea E la S -órbita de X (es decir el conjunto $\{g \cdot x \mid g \in S\}$). Entonces $\tau(E) \subseteq E$ y $\rho(E) \subseteq E$ y como, por hipótesis, $\tau \omega'_1 x = \rho \omega'_2 x$ siendo ω'_1 y ω'_2 dos palabras cualesquiera pertenecientes a S , entonces $\tau(E) \cap \rho(E) = \emptyset$. Como $\rho^{-1}(\rho)(E) = E = \tau^{-1}(\tau)(E)$, esto demuestra que es G -paradójico.

Ejemplo 3.1.5. (AE) *La esfera 2-dimensional \mathbb{S}^1 es numerablemente SO_2 -paradójica. Sea la relación de equivalencia que establece que un punto de \mathbb{S}^1 es equivalente a otro punto de \mathbb{S}^1 si existe una rotación que lleve el primero al segundo con ángulo múltiplo racional de 2π . Formemos el conjunto M formado por un representante de cada clase de equivalencia conforme a la relación de equivalencia anterior. Como los números racionales son numerables, el conjunto $\{\rho_i\}_{i=1}^{\infty}$ formado por las rotaciones antes mencionadas son numerables también. Es sencillo comprobar que $\{M_i\}_{i=1}^{\infty} = \{\rho_i(M)\}_{i=1}^{\infty}$ es una partición en \mathbb{S}^1 y además si tenemos $M_i, M_j, i \neq j$ existe una rotación de ángulo múltiplo racional de 2π que lleva de M_i a M_j , luego son congruentes. Ya hemos visto en el capítulo anterior que existe una biyección de los números pares en los naturales, luego podemos rotar todos los $\{M_i\}$ con índice i par hasta cubrir todos los $\{M_i\}, i \in \mathbb{N}$. q e d .*

La proposición que se expone a continuación y lo expuesto en el Ejemplo 3.1.2 serán claves en la demostración de la paradoja de Banach-Tarski sobre \mathbb{S}^2 . Más adelante como primer paso hacia la demostración, encontraremos un grupo libre basado en rotaciones SO_3 que es paradójico respecto a sí mismo y cuya acción es libre sobre toda la esfera menos un conjunto numerable de puntos ($\mathbb{S}^2 \setminus D$). De esa manera podremos deducir aplicando la proposición siguiente que $\mathbb{S}^2 \setminus D$ es SO_3 -paradójico.

Proposición 3.1.4. (AE)

Si G es paradójico y actúa sobre X sin puntos fijos no triviales, entonces X es G -paradójico. Por lo tanto, X es F -paradójico siempre que F , un grupo libre de rango 2, actúe sobre X sin puntos fijos no triviales.

Demostración

Supongamos que con los siguientes conjuntos (en este caso grupos), $A_i, B_j \subseteq G$ y las siguientes transformaciones $g_i, h_j \subseteq G$, G es paradójico. Por el Axioma de Elección, podemos formar un conjunto M que contiene exactamente un elemento de cada G -órbita en X :

$$M = \{m_i\}_{i \in I}$$

Entonces $G^* = \{g(M) : g \in G\}$ es una partición de G . Veamos que es así porque la acción de G es libre y esto nos asegura que $\forall k, l$ los conjuntos $G^k = \{g(m_k) : g \in G\}$, $G^l = \{g(m_l) : g \in G\}$ son conjuntos disjuntos, esto es, $G^k \cap G^l = \emptyset$. Si no fuera así $gm_r = hm_l$ para ciertos r, l . Entonces $h^{-1}gm_r = m_l$

Por tanto $m_r \sim m_l$ y esto quiere decir (por la elección de M) que $R = L$ siendo R y L las clases de equivalencia de m_r y m_l respectivamente. Es decir que $h^{-1}gm_r = m_r$. Como la acción es libre $h^{-1}g = e$ y por lo tanto $g = h$.

Sea $A_i^* = \{g(M) : g \in A_i\}$ y $B_j^* = \{g(M) : g \in B_j\}$ Entonces $\{A_i^*\} \cup \{B_j^*\}$ son familias de subconjuntos disjuntos dos a dos de X (ya que $\{A_i\} \cup \{B_j\}$ lo son). Y como $G = \bigcup g_i(A_i) = \bigcup h_j(B_j)$, entonces $X = \bigcup g_i(A_i^*) = \bigcup h_j(B_j^*)$ q.e.d.

Proposición 3.1.5. (AE) *Un grupo, conteniendo un subgrupo paradójico, es paradójico*

Por lo tanto cualquier grupo con un subgrupo libre de rango 2 (paradójico) es paradójico.

Demostración

Sea $G' \in G$. Veamos que la acción del subgrupo G' sobre el grupo G no tiene puntos fijos, ya que si $g' \in G', g \in G$, tenemos $g'g = g$, entonces $g' = e$.

3.2. La paradoja de Hausdorff

Ya se ha demostrado en el capítulo anterior que se pueden utilizar grupos libres no abelianos para generar conjuntos paradójicos. El hecho de que G_1 (grupo de isometrías en \mathbb{R}) y G_2 (grupo de isometrías en \mathbb{R}^2) sean resolubles (y por tanto no contengan ningún subgrupo libre no abeliano) implica que la primera oportunidad de que aparezca un grupo libre de este tipo entre los grupos de isometría euclídea sea en \mathbb{R}^3 . A partir de ahora, un subconjunto S de un grupo G se llamará independiente si es un conjunto libre de generadores de G' , el subgrupo de G generado por S . Por lo tanto G' será un grupo libre de rango $|S|$. Hay muchas maneras de generar un par de rotaciones independientes de \mathbb{S}^2 , la bola unidad en \mathbb{R}^3 . A continuación mostramos dos encontradas por K. Satô :

$$\sigma = \frac{1}{7} \begin{pmatrix} 6 & 2 & 3 \\ 2 & 3 & -6 \\ -3 & 6 & 2 \end{pmatrix}, \tau = \frac{1}{7} \begin{pmatrix} 2 & -6 & 3 \\ 6 & 3 & 2 \\ -3 & 2 & 6 \end{pmatrix}$$

las cuales llamamos las rotaciones de Satô.

Usaremos F_n para referirnos al grupo libre de rango n .

Teorema 3.2.1. *Las dos rotaciones de Satô son independientes y representan rotaciones alrededor de ejes que pasan por el origen en \mathbb{R}^3 . Por lo tanto si $n \geq 3$, SO_n tiene un subgrupo libre de rango 2.*

Demostración

Queremos demostrar que ninguna palabra reducida no trivial en $\sigma^{\pm 1}, \tau^{\pm 1}$ es igual a la identidad. Supongamos que ω es tal palabra y que ω es igual a la identidad. Como la identidad es invariante por conjugación y es igual a su inversa, realizando esas operaciones las veces que fuera necesario, conseguiríamos una palabra equivalente ω' que termina en σ . Definamos las siguientes cuatro matrices como:

$$M_\sigma = \begin{pmatrix} 6 & 2 & 3 \\ 2 & 3 & -6 \\ -3 & 6 & 2 \end{pmatrix}, M_\tau = \begin{pmatrix} 2 & -6 & 3 \\ 6 & 3 & 2 \\ -3 & 2 & 6 \end{pmatrix}$$

$$M_\sigma^- = \begin{pmatrix} 6 & 2 & -3 \\ 2 & 3 & 6 \\ 3 & -6 & 2 \end{pmatrix}, M_\tau^- = \begin{pmatrix} 2 & 6 & -3 \\ -6 & 3 & 2 \\ 3 & 2 & 6 \end{pmatrix}$$

Es trivial comprobar que $M_\sigma^- = M_\sigma^{-1} = M_\sigma^t; M_\tau^- = M_\tau^{-1} = M_\tau^t$. Veamos que es más sencillo trabajar con estas matrices en lugar de las verdaderas matrices correspondientes a las rotaciones; la única diferencia es una potencia entera de 7 que se añadiría al final.

Analizaremos $\omega'(1, 0, 0)$, la primera columna de la matriz $M_\sigma^{k_n} M_\tau^{k_{n-1}} \dots M_\tau^{k_2} M_\sigma^{k_1} M_\sigma$, donde $k_1 \geq 0, k_n \in \mathbb{Z}$, y las demás k_s son distintas de cero. Una potencia negativa se referirá a una potencia de $M_{\tau, \sigma}^-$. Trabajaremos en módulo 7 en todo momento, que es suficiente porque un entero que no es divisible por 7 es distinto de cero. Definamos los siguientes cuatro conjuntos de vectores:

- $V_\sigma = \{(3, 1, 2), (5, 4, 1), (6, 2, 4)\}$,
- $V_\sigma^- = \{(3, 2, 6), (5, 1, 3), (6, 4, 5)\}$,
- $V_\tau = \{(3, 5, 1), (5, 6, 4), (6, 3, 2)\}$,
- $V_\tau^- = \{(1, 5, 4), (2, 3, 1), (4, 6, 2)\}$

A continuación mostramos las siguientes propiedades de la acción de los generadores sobre los anteriores conjuntos (se pueden comprobar por medio de sencillos cálculos matriciales):

1. Para todo $v \in V_\sigma \cup V_\tau \cup V_\tau^-, \sigma v \in V_\sigma$.

2. Para todo $v \in V_\sigma^- \cup V_\tau \cup V_\tau^-$, $\sigma^- v \in V_\sigma^-$.

3. Para todo $v \in V_\tau \cup V_\sigma \cup V_\sigma^-$, $\tau v \in V_\tau$.

4. Para todo $v \in V_\tau^- \cup V_\sigma \cup V_\sigma^-$, $\tau^- v \in V_\tau^-$.

Veamos que en el primer paso en la aplicación de ω' sobre $(1, 0, 0)$ es $M_\sigma(1,0,0)$ que resulta $(6, 2, -3) = (6, 2, 4)(\text{mod } 7)$, siendo un vector $v \in V_\sigma$. Como posteriormente tenemos la acción de $M_\sigma^{k_1}, k_1 \geq 0$ sobre v , aplica el apartado 1 y el resultado pertenece a V_σ . Por 3 ó 4 (dependiendo del signo de k_2) veamos que la aplicación de $M_\tau^{k_2}$ sobre el resultado anterior deja el vector en $V_\tau \cup V_\tau^-$. La aplicación de $M_\sigma^{k_3}$, teniendo en cuenta 1 ó 2 (dependiendo del signo de k_3) convierte a un vector en $V_\sigma \cup V_\sigma^-$. Esta alternancia continúa conforme vamos evolucionando en la aplicación de las sucesivos componentes de la palabra en sentido de derecha a izquierda, de tal manera que el vector final pertenecerá a uno de los conjuntos $V_\sigma, V_\tau, V_\sigma^-, V_\tau^-$.

La propuesta anterior es sólo una de las muchas construcciones de un sistema libre no abeliano de rotaciones en \mathbb{R}^3 . Hausdorff dio la primera construcción de este tipo en 1914 ; demostró que si ϕ y ρ son rotaciones de 180° y 120° , respectivamente, sobre ejes que contienen el origen, y si $\cos 2\theta$ es trascendente donde θ es el ángulo entre los ejes, entonces ϕ y ρ son generadores libres de $\mathbb{Z}_2 * \mathbb{Z}_3$. Al final de esta sección dedicaremos un apartado para hablar de la paradoja tal y como la planteó originalmente su autor. Se puede ver que hay más casos que los anteriores, de hecho se puede demostrar que si se dota a $SO_3(\mathbb{R}) \times SO_3(\mathbb{R})$ de la topología del producto, entonces $\{(\phi, \rho) \in SO_3(\mathbb{R}) \times SO_3(\mathbb{R}) : \phi \text{ y } \rho \text{ son independientes}\}$ es comagro y por tanto denso.

Cada elemento del grupo libre de rotaciones (llamémoslo F) construido anteriormente fija todos los puntos en alguna línea de \mathbb{R}^3 , por lo que la Proposición 3.1.4 no puede todavía aplicarse. Podemos utilizar una aproximación ingenua para salvar esta dificultad inicial. Cada rotación no igual a la identidad en F tiene dos puntos fijos en \mathbb{S}^2 , la esfera unitaria (la intersección del eje de la rotación con la esfera). Si llamamos a D la colección de todos esos puntos, como F es numerable también lo es D .

Ahora bien, si $p \in \mathbb{S}^2 \setminus D$ y $g \in F$, entonces $g(p)$ está en $\mathbb{S}^2 \setminus D$ también: Si h fija $g(p)$, entonces p sería un punto fijo de $g^{-1}hg$ y esto es una contradicción pues supondría que $p \in D$. Por lo tanto, F actúa en $\mathbb{S}^2 \setminus D$ sin puntos fijos no triviales, y la Proposición 3.1.4 puede aplicarse a esta acción para obtener el siguiente resultado.

Teorema 3.2.2. *(Paradoja de Hausdorff) (AE).*

Existe un subconjunto numerable D de \mathbb{S}^2 tal que $\mathbb{S}^2 \setminus D$ es $SO_3(\mathbb{R})$ -paradójico.

Veremos en el próximo capítulo cómo la pequeñez del conjunto D permite eliminarlo por completo, dando lugar a la Paradoja de Banach–Tarski: \mathbb{S}^2 es $SO_3(\mathbb{R})$ -paradójico.

A continuación exploraremos la relación que existe entre la imposibilidad de la existencia de medidas y la existencia de descomposiciones paradójicas.

Supongamos que un grupo G actúa sobre un conjunto X , y $E \subseteq X$.

Definición 3.2.3. E se llama G -despreciable si $\mu(E) = 0$ para una medida μ finitamente aditiva e invariante de G sobre $\mathcal{P}(X)$ y $\mu(E) < \infty$.

Proposición 3.2.4. Si E es un conjunto G -paradójico, entonces E es G -despreciable.

Demostración

Supongamos que μ es una medida finitamente aditiva e invariante de G en $\mathcal{P}(X)$ y $\mu(E) < \infty$. Sean A_i , B_j y g_i, h_j las particiones de E y las transformaciones con las cuales se verifica que E es G -paradójico. Entonces $\mu(E) \geq \sum \mu(A_i) + \sum \mu(B_j) = \sum \mu(g_i A_i) + \sum \mu(h_j B_j) = \mu(\cup g_i A_i) + \mu(\cup h_j B_j) = \mu(E) + \mu(E) = 2\mu(E)$. Como $\mu(E) < \infty$, esto significa $\mu(E) = 0$.

El siguiente teorema deduce la despreciableidad de $SO_3(\mathbb{R})$ de \mathbb{S}^2 a partir de la Paradoja de Hausdorff demostrando que los conjuntos numerables son despreciables con respecto a las medidas finitas sobre $\mathcal{P}(\mathbb{S}^2)$.

Teorema 3.2.5. *(AE)* La esfera \mathbb{S}^2 es $SO_3(\mathbb{R})$ -despreciable. Por lo tanto, no hay ninguna medida finita aditiva e invariante por rotaciones sobre $\mathcal{P}(\mathbb{S}^2)$ que tenga medida total uno. Además, para cualquier $n \geq 3$, no existe ninguna medida finitamente aditiva e isométricamente invariante en $\mathcal{P}(\mathbb{R}^n)$ que normalice el cubo unidad.

Demostración

Supongamos que μ es una medida finitamente aditiva y $SO_3(\mathbb{R})$ -invariante- sobre $\mathcal{P}(\mathbb{S}^2)$ con $\mu(\mathbb{S}^2) < \infty$. Si D es el conjunto numerable de la Paradoja de Hausdorff, entonces, por la Proposición 3.2.4, $\mu(\mathbb{S}^2 \setminus D) = 0$. Por tanto, basta con demostrar que $\mu(D) = 0$.

Sea una recta l que pasa por el origen siendo el conjunto D y esa recta disjuntos. Para cada punto $p \in D$, sea $A(p)$ el conjunto de ángulos θ tales que la rotación de p alrededor de l de θ^j radianes, para cualquier número entero positivo j , lleva p a otro punto de D . La numerabilidad de D y el conjunto de posibles j implica que $A(p)$ es numerable y, por tanto, que $A = \bigcup_{p \in D} A(p)$ es numerable. Si elegimos ρ de tal manera que sea una rotación alrededor de uno de los innumerables ángulos que no están en A , entonces ρ tiene la propiedad de que, para cualquier j , $\rho^j(D)$ y D son disjuntos. Se deduce que $D \cup \rho(D) \cup \rho^2(D) \cup \dots$ es una unión disjunta dos a dos. Supongamos ahora que $\mu(D) > 0$. Entonces podremos escoger un entero k de tal manera que $k\mu(D) > l$ siendo l cualquier número real positivo. Entonces, como partimos de la hipótesis de que μ es invariante- $SO_3(\mathbb{R})$ y como suponemos $\mu(\mathbb{S}^2) < \infty$ tenemos:

$$\mu(D) + \mu(\rho(D)) + \dots + \mu(\rho^k(D)) = k\mu(D) > \mu(\mathbb{S}^2)$$

lo cual es una contradicción por lo que $\mu(D)$ debe ser cero, y por tanto $\mu(\mathbb{S}^2) = 0$ por lo que \mathbb{S}^2 es $SO_3(\mathbb{R})$ -despreciable.

Para demostrar la afirmación sobre \mathbb{R}^n , basta con considerar $n = 3$, porque una medida en una dimensión superior induce una en \mathbb{R}^3 . Ahora bien, si μ es una medida isométrica-invariante en \mathbb{R}^3 que normaliza el cubo unidad, entonces μ debe desaparecer en los conjuntos unidad. Esto se debe a que dos conjuntos unidad cualesquiera son congruentes y por lo tanto reciben la misma medida; por lo tanto, si la medida de un conjunto unidad fuera positiva, entonces la medida del cubo unidad sería infinita. Además, la invariancia de traslación implica que cualquier cubo tiene una medida finita, no nula, y se deduce que $0 < \mu(B) < \infty$, donde B denota la bola de radio unidad. Si definimos una medida ν en $P(\mathbb{S}^2)$ por la adición de radios, es decir, $\nu(A) = \mu\{\alpha P : P \in A, 0 < \alpha \leq 1\}$. Como $\mu(\{0\}) = 0$, $\nu(\mathbb{S}^2) = \mu(B)$. Además, ν es finitamente aditiva y $SO_3(\mathbb{R})$ -invariante. porque μ lo es. Esto contradice la $SO_3(\mathbb{R})$ -despreciabilidad de \mathbb{S}^2 . \square

La paradoja de Hausdorff original

Dedicamos este apartado a exponer la paradoja de Hausdorff tal y como la planteó originalmente en su libro *Grundzuge der Mengenlehre*[4]. Su propuesta fue construir un grupo libre con las dos siguientes rotaciones sobre \mathbb{S}^2 :

$$\psi = \begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \phi = \begin{pmatrix} -\cos \theta & 0 & \sin \theta \\ 0 & -1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix}$$

La primera representa una rotación de $\frac{2\pi}{3}$ respecto al eje Z . La segunda es una rotación de ángulo π respecto a un eje que se halla en el plano XZ y que forma un ángulo de θ con el eje Z siendo θ un ángulo cuyo coseno es un número trascendente. Trivialmente se cumple $\psi^3 = \mathbb{I}$ y $\phi^2 = \mathbb{I}$. Debido a estas relaciones se podran formar palabras del siguiente tipo:

- $\alpha = \phi\psi^{m_1}\phi\psi^{m_2}\dots\phi\psi^{m_n}$
- $\beta = \psi^{m_1}\phi\psi^{m_2}\phi\dots\psi^{m_n}\phi$
- $\gamma = \phi\psi^{m_1}\phi\psi^{m_2}\dots\phi\psi^{m_n}\phi$
- $\delta = \psi^{m_1}\phi\psi^{m_2}\dots\phi\psi^{m_n}$

donde n es un número natural y $m_1, m_2, \dots, m_n = 1$ ó 2 ($m_1, m_2, \dots, m_n = 1$ ó -1). Veamos que si $\alpha \neq 1 \implies \beta \neq 1$ ya que $\phi\beta\phi = \alpha$. Por otro lado se verifica $\gamma \neq 1 \implies \delta \neq 1$ pues $\phi\delta\phi = \gamma$. Finalmente $\psi^{-m_1}\delta = \alpha$ luego $\alpha \neq 1 \implies \delta \neq 1$. Luego es sencillo ver que no puede existir una palabra reducida, de alguna de las formas anteriores, no identitaria si alguna de las restantes es la identidad. Todo se reduce por tanto a encontrar una rotación de ángulo θ para ψ tal que α siempre sea distinta de la identidad. Y esta condición demuestra Hausdorff que se alcanza siendo θ un ángulo cuyo coseno sea trascendente. No demostraremos esta aseveración, aunque explicaremos sin entrar en muchos detalles por qué se obtiene esta conclusión. Como α consta de términos en $\phi\psi$ o $\phi\psi^2$ y estos tienen la siguiente forma:

$$\phi\psi = \begin{pmatrix} \frac{1}{2}\cos\theta & \frac{\sqrt{3}}{2}\cos\theta & \text{sen}\theta \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2}\text{sen}\theta & -\frac{\sqrt{3}}{2}\text{sen}\theta & \cos\theta \end{pmatrix}, \phi\psi^2 = \begin{pmatrix} \frac{1}{2}\cos\theta & -\frac{\sqrt{3}}{2}\cos\theta & \text{sen}\theta \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2}\text{sen}\theta & \frac{\sqrt{3}}{2}\text{sen}\theta & \cos\theta \end{pmatrix}$$

Entonces si llamamos $v_i = \phi\psi, \phi\psi^2$, es decir v_i toma una de las dos formas, entonces α tendrá la forma:

$$\alpha = v_k \cdots v_2 v_1 \text{ para algún } k \in \mathbb{N}$$

Para el caso $k = 1$, α es $\phi\psi$ o $\phi\psi^2$. Consideremos el punto

$$p = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Veamos que la acción de una palabra genérica $\alpha = v_k \cdots v_2 v_1$ sobre p nos proporciona otro punto cuya tercera coordenada es un polinomio $P(\cos\theta)$ de orden k en $\cos\theta$ el cual tiene coeficientes racionales de tal manera que $\alpha p = p$ implicaría $P(\cos\theta) - 1 = 0$, pero esto no es posible si $\cos\theta$ es trascendente.

Una vez comprobado que forman un sistema de rotaciones independientes, Hausdorff encuentra una partición del grupo libre generado por ψ y ϕ en tres clases A, B y C . Partiendo de una rotación generada inicialmente por $\{\psi, \phi\}$ distribuye el resto de las diferentes rotaciones generadas entre las diferentes clases conforme al siguiente criterio:

- $\rho \in A \iff \rho\phi \in B \cup C$.
- $\rho \in A \iff \rho\psi \in B$.
- $\rho \in A \iff \rho\psi^2 \in C$.

Comenzando por la rotación trivial $\rho = 1$ se expone a continuación un esquema ¹ de las primeras asignaciones. La primera línea designa el grupo A y a la derecha en la misma línea aparecen algunas de las primeras rotaciones que se pueden asignar a esa clase. Lo mismo se puede aplicar para el resto de líneas.

$$\begin{array}{l|l|l|l|l} A & 1 & & \psi\phi & , & \psi^2\phi & , & \phi\psi^2 & & \phi\psi\phi & & \dots \\ B & & \phi & , & \psi & & & & & \phi\psi^2\phi & , & \psi\phi\psi & , & \psi^2\phi\psi & \dots \\ C & & & & \psi^2 & & & & \phi\psi & & \psi\phi\psi^2 & , & \psi^2\phi\psi^2 & \dots \end{array}$$

Por lo tanto $A = \{1, \psi\phi, \psi^2\phi, \phi\psi^2, \phi\psi\phi, \dots\}$, $B = \{\phi, \psi, \phi\psi^2\phi, \psi\phi\psi, \psi^2\phi\psi, \dots\}$ y $C = \{\psi^2, \phi\psi, \psi\phi\psi^2, \psi^2\phi\psi^2, \dots\}$ Además esta partición representa una descomposición paradójica puesto que si llamamos G al grupo libre generado por $\{\psi, \phi\}$ entonces:

$$G = A \cup B \cup C = A \cup A\phi$$

¹Se reproduce el esquema en la misma forma en que Hausdorff lo muestra originalmente en su libro

Por lo tanto con la partición antes elegida el grupo libre de rotaciones G es G -paradójico. Sea D el conjunto de todos los puntos fijos bajo las rotaciones generadas por ψ y ϕ . Sea la partición M sobre $P^* = \mathbb{S}^2 \setminus D$ formada por un representante de cada G -órbita. Entonces:

$$P^* = M + M\phi + M\psi + M\psi^2 + \dots$$

De esta manera se consigue una partición de $P^* = \mathbb{S}^2 \setminus D$ en conjuntos A^*, B^*, C^* de la siguiente manera²:

- $A^* = M \cup M\psi\phi \cup M\psi^2\phi \cup M\phi\psi^2 + \dots$
- $B^* = M\phi \cup M\psi + \dots$
- $C^* = M\psi^2 \cup M\phi\psi + \dots$

Y utilizando la construcción antes mencionada:

$$A^*\phi = B^* \cup C^*, \quad A^*\psi = B^*, \quad A^*\psi^2 = C^*,$$

se deduce trivialmente que los conjuntos A^*, B^*, C^* y $B^* \cup C^*$ congruentes, quedando demostrado que $P^* = A^* \cup B^* \cup C^* = \mathbb{S}^2 \setminus D$ es G -paradójico.

3.3. Congruencia y equidescomponibilidad

La idea de cortar una figura en trozos y reorganizarlos para formar otra figura se remonta al menos a la geometría griega, donde este método se utilizaba para derivar fórmulas de área para paralelogramos. En este contexto surge el concepto de congruencia entre figuras geométricas que definiremos a continuación.

Definición 3.3.1. Dos polígonos en el plano son congruentes por disección si uno de ellos puede descomponerse en un número finito de piezas poligonales que pueden reordenarse mediante isometrías (e ignorando los límites) para formar el otro polígono.

Está claro que los polígonos que son congruentes por disección tienen la misma área. Lo contrario se demostró a principios del siglo XIX, utilizando el hecho de que la congruencia por disección es una relación de equivalencia.

El estudio de la congruencia entre polígonos en dimensiones superiores a 2 es más complicado. En este sentido se puede mencionar el tercer problema de Hilbert en el cual se plantea si un tetraedro regular es congruente por disección en un cubo. Este problema fue resuelto por Max Dehn en 1900 demostrando que no había congruencia posible entre un tetraedro y un cubo.

Introducimos ahora el concepto de equidescomponibilidad que es una generalización del concepto de congruencia.

²En la notación original Hausdorff utiliza el simbolo $+$ para referirse a unión

Definición 3.3.2. Supongamos que G actúa sobre X y que $A, B \subseteq X$. Entonces A y B son G -equidescomponibles (a veces llamados finitamente G -equidescomponibles) si A y B pueden ser particionados cada uno en el mismo número finito de piezas respectivamente G -congruentes. Formalmente,

$$A = \bigcup_{i=1}^n A_i, B = \bigcup_{i=1}^n B_i,$$

$A_i \cap A_j = \emptyset = B_i \cap B_j$ si $i < j \leq n$ y hay g_1, \dots, g_n tales que para cada $i \leq n$:

$$g_i(A_i) = B_i$$

Básicamente se puede entender la equidescomponibilidad entre dos conjuntos de la siguiente manera. Partiendo de los conjuntos A y B , éstos son equidescomponibles, si yo puedo partir A en trozos, aplicar sobre cada uno de los trozos una transformación igual o diferente y componer B con cada uno de los trozos transformados. Podemos extender la relación de equidescomponibilidad considerando la utilización de una cantidad infinito numerable de trozos, es decir A y B serán numerablemente G -equidescomponibles si:

$$A = \bigcup_{i=1}^{\infty} A_i, B = \bigcup_{i=1}^{\infty} B_i,$$

$A_i \cap A_j = \emptyset = B_i \cap B_j$ si $i < j \leq \infty$ y hay g_1, \dots, g_m tales que para cada $i \leq \infty$:

$$g_i(A_i) = B_i, i = 1, \dots, \infty$$

Proposición 3.3.3. *La relación \sim_G es una relación de equivalencia.*

La demostración es bastante sencilla. La propiedad reflexiva se verifica ya que cualquier conjunto es equidescomponible en sí mismo utilizando la transformación identidad. La propiedad simétrica también es fácilmente deducible ya que si $A \sim_G B$, es decir:

$$g_i(A_i) = B_i, A = \bigcup_{i=1}^n A_i, B = \bigcup_{i=1}^n B_i,$$

entonces $B \sim_G A$ puesto que:

$$g_i^{-1}(B_i) = A_i, A = \bigcup_{i=1}^n A_i, B = \bigcup_{i=1}^n B_i,$$

La propiedad transitiva también es satisfecha. Sea $A \sim_G B$ y $B \sim_G C$, esto es:

$$g_i(A_i) = B_i, i = 1, \dots, n \text{ con } A = \bigcup_{i=1}^n A_i, B = \bigcup_{i=1}^n B_i,$$

$$h_j(B_j) = C_j, j = 1, \dots, n, B = \bigcup_{j=1}^n B_j, C = \bigcup_{j=1}^n C_j,$$

Luego finalmente $A \sim_G C$ puesto que:

$$\tau(A_k) = hg(A_k) = C_k, A = \bigcup_{k=1}^n A_i, C = \bigcup_{k=1}^n C_k, k = 1, \dots, n$$

siendo $\tau = hg \in G$.

En contraste, es importante ver que la relación \sim_n no es transitiva. Podemos poner un ejemplo trivial con conjuntos discretos. Sea $A = \{1, 2, 3, 4\}$ y $B = \{1, 2, 7, 8\}$ y $C = \{1, 7, 6, 12\}$. Vemos que $A \sim_2 B$, pues:

$$\begin{aligned} e\{1, 2\} &= \{1, 2\} \\ g\{3, 4\} &= \{7, 8\} \end{aligned}$$

siendo e la identidad y g la traslación de 4 unidades. Por otro lado, $B \sim_2 C$:

$$\begin{aligned} e\{1, 7\} &= \{1, 7\} \\ g\{2, 8\} &= \{6, 12\} \end{aligned}$$

siendo e la identidad y g la traslación de 4 unidades. Sin embargo, se observa que $A \not\sim_2 C$ ($A \sim_4 C$). Generalizaremos este resultado en la siguiente proposición:

Proposición 3.3.4. *Sean 3 conjuntos A, B, C . Si se verifica $A \sim_n B$ y $B \sim_m C$ entonces $A \sim_{nm} C$.*

Demostración

Supongamos $A \sim_n B$ y $B \sim_m C$. Por lo tanto existen particiones $\{A_i\}_{i=1}^n, \{B_j\}_{j=1}^m$ de A y B respectivamente, y transformaciones $\{g_i\}_{i=1}^n \in G$ tales que $g_i(A_i) = B_i, i = 1, \dots, n$. Por otro lado existen también particiones $\{B'_j\}_{j=1}^m$ y $\{C_j\}_{j=1}^m$ y $\{h_j\}_{j=1}^m \in G$ tales que $h_j(B'_j) = C_j, j = 1, \dots, m$. Sean los conjuntos $\{A'_{ij}\} = A_i \cap g_i^{-1}(B'_j)$ y sea $\tau_{ij} = h_j g_i$. Es inmediato que $\tau_{ij} \in G$ y que $\{A'_{ij}, 1 \leq i \leq n, 1 \leq j \leq m\}$ y $\{\tau_{ij} A'_{ij} : 1 \leq i \leq n, 1 \leq j \leq m\}$ particionan A y C respectivamente. Luego $A \sim_{nm} C$.

Veamos que podemos formular una nueva definición de conjunto paradójico usando el concepto de G -equidescomponibilidad: $E \subseteq X$ es G -paradójico si y sólo si existe una familia de subconjuntos disjuntos A y B tales que $A \sim_G E$ y $B \sim_G E$.

Proposición 3.3.5. *Supongamos que G actúa sobre X y que E, E' son subconjuntos G -equidescomponibles de X . Si E es G -paradójico, también lo es E' .*

Demostración

Como E es G -paradójico podemos afirmar que existen conjuntos A, B disjuntos contenidos en E y $A \sim E, B \sim_G E$. Si $E \sim E'$, entonces existen $E_i, E'_i, E = \bigcup_{i=1}^n E_i, E' = \bigcup_{i=1}^n E'_i$, y:

$$g_i(E_i) = E'_i, i = 1, \dots, n, g_i \in G$$

Por otro lado, como $A \subset E$, entonces $\{A \cap E_i\}$ es una partición de A y además como $A \sim E$ se cumple:

$$A \sim \bigcup_{i=1}^n g_i(A \cap E_i) \subset E'$$

Y también se verifica

$$\bigcup_{i=1}^n g_i(A \cap E_i) \sim A \sim E \sim E'$$

De igual manera se podría razonar para la partición $\{B \cap E'_i\}$:

$$\bigcup_{i=1}^n g_i(B \cap E_i) \subset E'.$$

$$\bigcup_{i=1}^n g_i(B \cap E_i) \sim B \sim E \sim E',$$

Aunque a primera vista pudiera parecer que los conceptos de equidescomponibilidad y congruencia por disección son iguales, en parte no es así. La clave está en que no hay restricción para los conjuntos utilizados en demostrar que $A \sim B$, y debido a esto no se puede garantizar que tengan el mismo área (por ejemplo en el caso de esos conjuntos no sean medibles Lebesgue). La conservación de una medida invariante bajo G -equidescomponibilidad está directamente relacionada con la existencia de una extensión finitamente aditiva e invariante de la medida a todos los subconjuntos de X , y esto, a su vez, depende de propiedades abstractas del grupo G .

A continuación introduciremos otra relación de equivalencia necesaria para exponer el Teorema de Banach–Schröder-Bernstein.

Definición 3.3.6. Sean dos conjuntos A, B . Diremos que $A \preceq B$ cuando A es equidescomponible con un subconjunto $C \subset B$.

El teorema de Banach–Schröder-Bernstein

Teorema 3.3.7. *Teorema de Banach–Schröder-Bernstein* Supongamos que G actúa sobre X y $A, B \subseteq X$. Si $A \preceq B$ y $B \preceq A$, entonces $A \sim_G B$. Por lo tanto es un ordenamiento parcial de las \sim_G -clases en $\mathcal{P}(X)$.

Demostración

Se comprueba fácilmente que la relación \sim_G satisface las dos condiciones siguientes:

1. Si $A \sim B$, entonces existe una biyección $g : A \rightarrow B$ tal que $C \sim g(C)$ siempre que $C \subseteq A$.
2. Si $A_1 \cap A_2 = \emptyset = B_1 \cap B_2$, y si $A_1 \sim B_1$, y $A_2 \sim B_2$, entonces $A_1 \cup A_2 \sim B_1 \cup B_2$.

El resto de la demostración sera demostrar que \sim es una relación de equivalencia sobre $\mathcal{P}(X)$.

Sean $f : A \rightarrow B_1$, $g : A_1 \rightarrow B$, donde $B_1 \subseteq B$ y $A_1 \subseteq A$, biyecciones garantizadas por (a). Sea $C_0 = A \setminus A_1$ y, por inducción, definamos C_{n+1} como $g^{-1} f(C_n)$. Sea $C = \bigcup_{i=0}^{\infty} C_n$. Entonces es fácil comprobar que $g(A \setminus C) = B \setminus f(C)$, puesto que;

$$\begin{aligned} g(A \setminus C) &= g(A \setminus \bigcup_{i=1}^{\infty} C_n) = g(A \setminus C_0) \cap g(A \setminus \bigcup_{i=1}^n C_n) = g(A_1) \cap g(A \setminus \bigcup_{i=0}^n g^{-1} f(C_n)) = \\ &= B \setminus f(\bigcup_{i=0}^{\infty} f(C_n)) = B \setminus f(C) \end{aligned}$$

Luego la elección de g implica que $A \setminus C \sim B \setminus f(C)$. Pero, debido a la elección de f , $C \sim f(C)$ y la propiedad (2) da como resultado $(A \setminus C) \cup C \sim (B \setminus f(C)) \cup f(C)$, o lo que es lo mismo: $A \sim B$ como se desea. Se deduce de la demostración anterior que si $A \preceq_n B$ y $B \preceq_m A$ entonces $A \sim_{nm} B$.

Corolario 3.3.8. *Un subconjunto E de X es G -paradójico si y sólo si existen conjuntos disjuntos $A, B \subseteq E$ con $A \cup B = E$ y $A \sim E \sim B$.*

Demostración

Supongamos que un subconjunto E de X es G -paradójico, digamos, A, B son subconjuntos disjuntos de E con $A \sim E \sim B$. Entonces $E \sim B \subseteq E \setminus A \subseteq E$, por lo que el teorema de Banach-Schröder-Bernstein implica que $E \setminus A \sim E$.

3.4. La paradoja de Banach Tarski en \mathbb{R}^3

En este apartado llegamos a la demostración de la paradoja de Banach Tarski para \mathbb{S}^2 y, por medio de una construcción sencilla y considerando otro tipo de isometrías, también ampliamos la paradoja a cualquier bola sólida en \mathbb{R}^3 . Como paso previo para lo primero, exponemos el siguiente teorema:

Teorema 3.4.1. *Si D es un subconjunto numerable de \mathbb{S}^2 , entonces \mathbb{S}^2 y $\mathbb{S}^2 \setminus D$ son $SO_3(\mathbb{R})$ -equidescomponibles (usando dos piezas).*

Demostración

Sea l una recta que pasa por el origen y que no pasa por el conjunto numerable D (recordemos que D está formado por los polos de todas las rotaciones generadas por el grupo libre del Teorema 3.2.1). Sea A el conjunto de ángulos $\theta \in [0, 2\pi)$ tal que para algún $n > 0$ y algún $p \in D$, $\rho(P)$ también está en D , donde ρ es la rotación alrededor de l de $n\theta$ radianes. Como A es numerable, podemos elegir un ángulo θ que no pertenezca a A ; sea la correspondiente rotación alrededor de l . Entonces $\rho^n(D) \cap D = \emptyset$ si $n > 0$, de lo que se deduce que siempre que $0 \leq m < n$, $\rho^m(D) \cap \rho^n(D) = \emptyset$. Por tanto si $D^* = \cup \{\rho^n(D) : n = 0, 1, 2, \dots\}$, entonces:

$$\rho(D^*) = D^* \setminus D.$$

Como $D^* \sim D^* \setminus D$:

$$\mathbb{S}^2 = D^* \cup (\mathbb{S}^2 \setminus D^*) \sim (D^* \setminus D) \cup (\mathbb{S}^2 \setminus D^*) = \mathbb{S}^2 \setminus D.$$

Proposición 3.4.2. (Paradoja de Banach-Tarski)(AE) *La esfera \mathbb{S}^2 es $SO_3(\mathbb{R})$ -paradójica, como lo es cualquier esfera centrada en el origen. Además, cualquier bola sólida en \mathbb{R}^3 es G_3 -paradójica y \mathbb{R}^3 mismo es paradójico.*

Demostración

La Paradoja de Hausdorff (Teorema 3.2.2) afirma que $\mathbb{S}^2 \setminus D$ es $SO_3(\mathbb{R})$ -paradójica para algún conjunto numerable D (de puntos fijos de rotaciones). Combinando esto con el teorema anterior y la Proposición 3.3.5 da como resultado que \mathbb{S}^2 es $SO_3(\mathbb{R})$ -paradójico. Por otro lado ninguno de los resultados anteriores depende del tamaño de las esferas, por lo cual, cualquier esfera sea del tamaño que sea, admite descomposiciones paradójicas. Para demostrar la descomposiciones paradójicas en bolas (esferas sólidas) podemos trabajar inicialmente con bolas centradas en 0 ya que G_3 (el grupo de las isometrías en \mathbb{R}^3) contiene todas las traslaciones. Por simplicidad consideraremos la bola unitaria B , pero la misma demostración funciona para bolas de cualquier tamaño. Veamos que considerando la correspondencia radial $P \rightarrow \{\alpha p : 0 < \alpha \leq 1\}$ la prueba se simplifica ya que si para un α específico la superficie esférica con ese radio admite una descomposición paradójica también lo admite la unión de superficies dentro de la bola. El problema estriba en el origen de la bola que es punto fijo para todas las rotaciones consideradas hasta ahora, esto es las rotaciones alrededor de un eje que pasa por el centro de la bola. Es por ello, que debemos manejar otro tipo de isometrías considerando rotaciones alrededor de un eje que no pase por el origen. Sea $p = (0, 0, \frac{1}{2})$ y sea ρ una rotación de orden infinito alrededor del eje que es la línea horizontal en el plano $x - z$ que contiene a P , esto es, $\rho^n(p) \neq p, \forall n = 1, \dots, \infty$. Es sencillo ver que la actuación de una rotación de tales características sobre p viene dada por $\rho(p) = e^{i\theta} p$ donde $\theta \in \mathbb{R} \setminus \mathbb{Q}$. Luego si $D^* = \cup \{\rho^n(0) : n = 0, 1, 2, \dots\}$ entonces:

$$\rho(D^*) = \cup \{\rho^n(0) : n = 1, 2, \dots\} = D^* \setminus 0$$

Luego D^* y $D^* \setminus 0$ son equidescomponibles y por lo tanto:

$$B \sim B \setminus 0$$

Para demostrar el mismo resultado pero con \mathbb{R}^3 , se puede utilizar la correspondencia radial de \mathbb{S}^2 con todo $\mathbb{R}^3 \setminus \{0\}$ de tal manera que se obtiene una descomposición paradójica de \mathbb{R}^3 utilizando rotaciones. Posteriormente, igual que para la bola, se demuestra que $\mathbb{R}^3 \setminus \{0\} \sim_{G_3} \mathbb{R}^3$, \mathbb{R}^3 es paradójico mediante isometrías.

Veamos que aunque el teorema de Banach Tarski no añade nada nuevo respecto a nuestro conocimiento de las medidas finitamente aditivas incompatibles con la paradoja de Hausdorff, el resultado sin embargo es mucho más sorprendente. En teoría podemos descomponer una bola en muchísimas piezas, rotarlas y volverlas a unir para formar 2,3 o incluso un millón de bolas. Desde luego, dado que el controvertido Axioma de Elección es utilizado esa descomposición sumamente numerosa tendría una forma increíblemente extraña. Por lo tanto, nos encontramos con el resultado paradójico que aunque las rotaciones conservan el volumen, conseguimos duplicarlo, triplicarlo o aumentarlo cuasi-infinitamente, pero una plausible y lógica explicación a esto podría ser de nuevo el considerar que las piezas en las que se descompone la esfera original no son medibles Lebesgue. Aunque el volumen no es una de las propiedades conservadas por la equidescomponibilidad en \mathbb{R}^3 , hay otras propiedades que sí se conservan mediante esa operación como la propiedad de ser un conjunto acotado. Esto viene expresado en la expresión de la paradoja en su forma fuerte:

Teorema 3.4.3. (*Versión fuerte de la Paradoja de Banach-Tarski*) (AE). Si A y B son cualesquiera dos subconjuntos acotados de \mathbb{R}^3 , cada uno con interior no vacío, entonces A y B son equidescomponibles.

Demostración

Basta con demostrar que $A \preceq B$, pues entonces, por el mismo argumento, $B \preceq A$, y el Teorema 3.3.7 da como resultado que $A \sim B$. Elijamos bolas sólidas K y L tales que $A \subseteq K$ y $L \subseteq B$, y que n sea lo suficientemente grande como para que K pueda ser recubierto por n copias (superpuestas) de L . Ahora, si S es un conjunto de n copias disjuntas de L , entonces usando la Paradoja de Banach–Tarski para duplicar repetidamente L , y utilizando traslaciones para mover las copias así obtenidas, se obtiene que $L \succeq S$ y por lo tanto $L \succeq K$. Finalmente $A \subseteq K \preceq S \preceq L \subseteq B$, por lo que $A \preceq B$.

Este notable resultado se suele ver de forma negativa porque ilustra de forma muy contundente el teorema de Hausdorff (Teor. 2.6) de que ciertas medidas no existen.

Para terminar debemos mencionar algunos problemas interesantes de naturaleza topológica en relación con la Paradoja de Banach–Tarski. Como se ha señalado, los trozos en los que se divide la esfera \mathbb{S}^2 son descompuestos para ser duplicados no puede ser medibles Lebesgue y por lo tanto no pueden ser conjuntos de Borel.

En el siguiente apartado estudiaremos, para la esfera \mathbb{S}^2 y la bola si hay un conjunto mínimo de piezas en las que pueden descomponerse de forma paradójica

3.5. Mínimo número de piezas en una descomposición paradójica

Análisis para la esfera \mathbb{S}^2

Empezaremos analizando lo demostrado hasta ahora para deducir el número de piezas que se ha obtenido en la descomposición paradójica obtenida para la esfera. Recordemos que en la demostración de la paradoja de Hausdorff se construyó un grupo libre con dos rotaciones independientes (generadores) que actuaban sobre puntos fijos no triviales. En virtud de lo demostrado en el Ejemplo 3.1.2, observamos que este grupo libre es paradójico utilizando 5 piezas, y por lo tanto según la Proposición 3.1.4 se llega al resultado que $\mathbb{S}^2 \setminus D$ (con D numerable) es $SO_3(\mathbb{R})$ -paradójico utilizando 5 piezas, esto es, existen $A, B \in \mathbb{S}^2$ tal que $A \sim_2 \mathbb{S}^2 \setminus D \sim_3 B$ ³. Por otro lado aplicando el Teorema 3.4.1, $\mathbb{S}^2 \setminus D \sim_2 \mathbb{S}^2$ y en virtud de la Proposición 3.3.4 finalmente $A \sim_4 \mathbb{S}^2 \sim_6 B$, lo que nos indica que en la descomposición paradójica encontrada para \mathbb{S}^2 se han utilizado en total 10 piezas. Cabe preguntarse entonces si se pueden obtener descomposiciones paradójicas de la esfera con un menor número de piezas. El estudio de esta cuestión tiene interés porque veremos que existe una interesante característica del grupo de rotaciones que condiciona cuál es el número mínimo de piezas a utilizar. Adicionalmente, las técnicas usadas para reducir el número de piezas conducen a importantes nuevas ideas para el tratamiento de los puntos fijos de acciones de grupos libres.

³Recuérdese la definición de equidescomponibilidad realizada en 3.3.2

Definición 3.5.1. Supongamos que G actúa sobre X y que $E \subseteq X$. Entonces E es G -paradójico usando r piezas si hay dos subconjuntos disjuntos $A, B \subseteq E$, $E = A \cup B$ tales que $A \sim_m E \sim_n B$ y $m + n = r$.

Ahora veremos que existe una descomposición minimal en 8 piezas para la esfera tridimensional encontrando primero una descomposición paradójica para el grupo libre en dos generadores en cuatro piezas en lugar de en cinco. Por una parte esto se conseguirá eliminando la identidad (e) como pieza y asociándola a una de las restantes piezas.

Teorema 3.5.2. *Sea F el grupo libre generado por σ y τ . Entonces F puede ser particionado en cuatro subconjuntos A_1, A_2, A_3, A_4 tales que $\sigma(A_2) = A_2 \cup A_3 \cup A_4$ y $\tau(A_4) = A_1 \cup A_2 \cup A_4$. Por lo tanto F es F -paradójico usando cuatro piezas.*

En la demostración del teorema se muestra exactamente cuáles son esas cuatro piezas.

Demostración

Veamos que la primera parte del teorema se cumple con la siguiente elección de piezas:

- $A_1 = W(\sigma) \cup \{e, \sigma^{-1}, \sigma^{-2}, \dots\}$
- $A_2 = W(\sigma^{-1}) \setminus \{e, \sigma^{-1}, \sigma^{-2}, \dots\}$
- $A_3 = W(\tau)$
- $A_4 = W(\tau^{-1})$

donde, recordemos, $W(\rho)$ es el conjunto de elementos de F cuya representación es una palabra formada por $\sigma, \sigma^{\pm 1}, \tau, \tau^{\pm 1}$ y que comienza por la izquierda con ρ . Las dos relaciones expresadas anteriormente se pueden expresar de una manera más completa de la siguiente manera:

- $\sigma(A_2) \subseteq A_2 \cup A_3 \cup A_4$
- $\sigma^{-1}(A_2 \cup A_3 \cup A_4) \subseteq A_2$
- $\tau(A_4) \subseteq A_1 \cup A_2 \cup A_4$
- $\tau^{-1}(A_1 \cup A_2 \cup A_4) \subseteq A_4$

Además para cualquier palabra $w \in F$, se puede elegir una partición en la que w esté en la misma pieza que la identidad en F , aunque obviaremos su demostración.

Definición 3.5.3. Una acción de un grupo G sobre X es llamada localmente conmutativa si $Estab(x)$ es conmutativo para todo $x \in X$.

Si la acción es tal que siendo dos elementos g_1 y g_2 de G cualesquiera que tengan un punto fijo común, se verifica que g_1 y g_2 conmutan, entonces la acción es localmente conmutativa. La acción definida por el grupo de rotaciones $SO_3(\mathbb{R})$ de la esfera es localmente conmutativa ya que se tiene la propiedad de que si dos rotaciones comparten un punto fijo, es decir, tienen el mismo eje, entonces conmutan, sin embargo si consideramos la acción del mismo grupo sobre la bola en \mathbb{R}^3 no ocurre lo mismo, ya que dos rotaciones con respecto al eje que pasa por el origen tienen a éste como punto fijo y sin embargo no conmutan.

Teorema 3.5.4. (AE) *Sea F el grupo libre generado por σ y τ . Si la acción de F sobre X es localmente conmutativa, entonces X es F -paradójico usando cuatro piezas.*

Demostración

Veamos que cada órbita de un elemento de X constará enteramente de puntos fijos o no contendrá ninguno. Sean $u, w \in G$ y $x \in X$, entonces si $w(x) = x$ y si $u(x) \in \mathcal{O}_x$, veamos que $u(x)$ es punto fijo de uwu^{-1} :

$$uwu^{-1}(u(x)) = uw(x) = u(x)$$

Entonces, realizaremos la asignación de las piezas en función del tipo de órbita, Para la órbita sin ningún punto fijo, elegiremos un representante m_i de cada órbita y formaremos el conjunto $M = \bigcup m_i, m_i \in \mathcal{O}_{x_i}, x_i \in X$

Entonces definimos los conjuntos $A_i^* = \{g(M) : g \in A_i, i = 1, \dots, 4\}$ siendo los conjuntos $\{A_i\}_1^4$ los utilizados en el Teorema 3.5.2. En el caso de la órbita compuesta de puntos fijos, la elección del representante no puede ser arbitraria. Elijamos en F la palabra no trivial de longitud más corta fijando un elemento de \mathcal{O} y sea x un punto que queda fijo por w . Denotemos por ρ el término más a la izquierda de w , entonces w no termina en ρ^{-1} , ya que si así fuera $\rho^{-1}w\rho$, siendo más corta que w , fijaría también $\rho^{-1}(x)$ que es un elemento de \mathcal{O} (contradiciendo la definición de w). Entonces utilizaremos el Teorema 3.5.2 para particionar F en $A_i, i = 1, \dots, 4$, satisfaciendo las dos ecuaciones anteriores y con la propiedad adicional de que $\{e\}$ y w estén en la misma pieza.

Vamos a ver que todo punto y en \mathcal{O} puede ser escrito como $v(x)$ y v no termina en w ni termina en w^{-1} . Consideremos una palabra v de longitud mínima tal que $y = v(x)$. Trivialmente se ve que v no puede acabar en w (ni en w^{-1}) ya que entonces $t = vw^{-1}(vw)$ sería una palabra más corta y $y = t(x)$. En el caso de que v termine en ρ^{-1} entonces vw no termina en w (ya que una parte de v se cancela con w) y tampoco termina en ρ^{-1} pues w no termina en ρ^{-1} , luego elegiríamos vw tal que $y=vw(x)$.

La unicidad de esta representación depende del hecho de que los únicos elementos que fijan x son las potencias de w . Para demostrar esto , fijémonos que si u fija x , entonces como se verifica la conmutatividad local u fija también $w(x)$ y por una importante propiedad de los grupos libres u y v tienen que ser potencias de alguna palabra $t \in F$, es decir $u = t^j, w = t^k$ con $j, k \in \mathbb{Z}$. Pero como w es minimal, se verifica $|j| > |k|$ y se puede escribir $j = lk + r, l, r \in \mathbb{Z}$. Se tiene entonces $x = u(x) = t^r(t^k)^l(x) = t^r(x)$ pues $t^k(x) = x$, pero utilizando de nuevo el argumento sobre la minimalidad de w r debe ser cero y por lo tanto k divide a j y u es potencia de w , como se decía anteriormente.

Ahora, si $y = v(x) = u(x)$ son dos representaciones de y en la forma deseada, entonces $u^{-1}v(x) = x$ y por lo tanto o $u^{-1}v$ o $v^{-1}u$ debe ser una potencia positiva de w . Supongamos que lo es $u^{-1}v$, entonces una posibilidad es que u^{-1} comience por ρ ya que esa condición es cumplida por toda palabra que es potencia de w pues el termino más a la izquierda de w es ρ . Pero esto contradice el hecho de que u no termine en ρ^{-1} . La otra posibilidad es que todos los literales de u^{-1} se cancelan con parte de v implicando que v termina en w , lo cual también es una contradicción.

Habiendo demostrado que esta representación es válida, comencemos la asignación de puntos en \mathcal{O} a los conjuntos $A_i^* = \{g(M) : g \in A_i, i = 1, \dots, 4\}$. Colocaremos y en A_i^* si A_i^* es el subconjunto de F que contiene a v , siendo $v(x)$ la representación única de y definida previamente. Veamos que esta asignación es la correcta. Consideremos la primera relación $(\sigma(A_2^*) = A_2^* \cup A_3^* \cup A_4^*)$ y supongamos que $y \in A_2^*$, es decir y tiene la representación $v(x)$, $v \in A_2^*$. Observemos $\sigma(y)$. Si $\sigma v(x)$ es la representación correcta de $\sigma(y)$ entonces como $y \in A_2^*$, entonces $\sigma(v) \in A_2 \cup A_3 \cup A_4$. Pero es posible que σv termine en w o ρ^{-1} (en el caso $\sigma v = w$). En el primer caso ya que v no finaliza en w , σv debe ser igual a w (o una potencia suya), luego $\sigma(y) = w(x) = x = e(x)$. Como σv está en $A_2 \cup A_3 \cup A_4$, w también estará en $A_2 \cup A_3 \cup A_4$ y también $\{e\}$. Ya que $e(x)$ es la única representación de x , esto implica que x , y por lo tanto $\sigma(y) = \sigma(x)$ está en $A_2^* \cup A_3^* \cup A_4^*$.

En el segundo caso (σv termina en ρ^{-1}), como v no lo hace, v debe ser $\{e\}$, luego $y = vx = x, \sigma v = \sigma$, y $\rho^{-1} = \sigma$ y por lo tanto w comienza por σ^{-1} . Luego como $w \in A_2^*$ entonces $v = e \in A_2^*$. Por lo tanto $\sigma(w)$ está en $A_2 \cup A_3 \cup A_4$.

Pero como $\sigma w(x)$ es la única representación de $\sigma(x)$ (ya que $w(x)$ deja fijo x), entonces $\sigma(y) = \sigma(x) \in A_2^* \cup A_3^* \cup A_4^*$. De la misma manera se realiza la asignación para las otras relaciones de contención señaladas en la demostración del Teorema 3.5.2.

Corolario 3.5.5. *El menor numero de piezas con la que \mathbb{S}^2 es SO_3 -paradójica es cuatro.*

Demostración

Ya que la acción de las rotaciones sobre la esfera es localmente conmutativa y como SO_3 tiene un subgrupo libre de rango 2, la conclusión de este corolario se sigue del teorema anterior(3.5.4)

Hay otros ejemplos de acciones localmente conmutativas a las que se puede aplicar el Teorema 3.5.4. Sea SL_2 , el grupo de transformaciones en \mathbb{R}^2 que conservan el área y la orientación, es localmente conmutativo en su acción sobre $\mathbb{R}^2 \setminus \{0\}$. Si σ_1 y σ_2 fijan p , entonces σ_i tiene una representación del tipo:

$$\begin{pmatrix} 1 & a_i \\ 0 & 1 \end{pmatrix}$$

donde $\det(\sigma_i) = 1$. Es inmediato ver que si $i \neq j$, $\sigma_i \sigma_j = \sigma_j \sigma_i$. Más adelante demostraremos que estas matrices representan transformaciones independientes en SL_2 , lo cual hace que $\mathbb{R}^2 \setminus \{0\}$ admita una descomposición paradójica en 4 piezas.

Análisis para la bola tridimensional B

Aunque el trabajo realizado en la sección anterior se podría aplicar a la bola B en G_3 , la situación es diferente por dos motivos:

1. G_3 no actúa sobre la bola unidad en \mathbb{R}^3 (pensemos en la traslación o la reflexión respecto a un eje que no pase por el centro de la misma).
2. La acción de G_3 no es localmente conmutativa. Dos rotaciones con diferentes ejes que pasan por el centro de la bola tienen éste como punto fijo y sin embargo no conmutan.

A pesar de esto, ya vimos en la Proposición 3.4.2 que B es SO_3 -paradójica. El asunto sobre cuál es el mínimo número de piezas para ello no está claro ya que depende de disponer de un conjunto de transformaciones sobre un conjunto en sí mismo. El siguiente teorema y su demostración resuelve la cuestión planteada por medio de un argumento geométrico.

Teorema 3.5.6. *Una bola sólida en \mathbb{R}^3 no admite una descomposición paradójica en menos de 5 piezas y esa descomposición existe para toda bola.*

Demostración

Haremos la demostración para la bola unidad de \mathbb{R}^3 centrada en $\mathbf{0}$. Denotemos por S la esfera unidad que es la superficie de B . Supongamos $B = B_1 \cup B_2 \cup B_3 \cup B_4$ siendo $\{B_i\}_{i=1}^{i=4}$ disjuntos y $\sigma_1 B_1 \cup \sigma_2 B_2 = B = \sigma_3 B_3 \cup \sigma_4 B_4$ con $\{\sigma_i\}_{i=1}^{i=4}$ isometrías. Partimos de este supuesto ya que sería el único caso en que se podría dar la descomposición con menos de 5 piezas. Es evidente que no todas las isometrías pueden fijar $\{0\}$. En caso contrario, sea por ejemplo el caso en que $\{0\}$ está en B_4 . Como los $\{B_i\}_{i=1}^{i=4}$ son disjuntos y $B_1 \cup B_2 \not\subseteq 0$ tenemos que $\sigma_1 B_1 \cup \sigma_2 B_2 \not\subseteq 0$ no cumpliéndose $\sigma_1 B_1 \cup \sigma_2 B_2 = B$. Supongamos ahora que σ_4 no fija 0, entonces $\sigma_4(B) \neq B$, luego existe una superficie hemisférica cerrada $H \in S$ que es disjunta con $\sigma_4(B)$. Sea r la recta que une $\sigma_4(0)$ y 0 y sea H la superficie hemisférica simétrica respecto al punto de intersección de S con la recta r . Ya que $\sigma_3(B_3)$ debe contener a H entonces $B_3 \supset \sigma_3^{-1}(H)$. Como $\{B_i\}_{i=1}^{i=4}$ son disjuntos, entonces $(\sigma_1 B_1 \cup \sigma_2 B_2) \cap S$ estarán contenidos en una superficie hemisférica abierta de S , concretamente el complemento de $\sigma_3^{-1}H$. Como ni B_1 ni B_2 contienen un hemisferio cerrado σ_1 y σ_2 deben fijar el origen y por lo tanto llevan S sobre S . Y $(\sigma_1 B_1 \cup \sigma_2 B_2) \cap S = (\sigma_1 B_1 \cap S) \cup (\sigma_2 B_2 \cap S)$ es un conjunto que está contenido en el conjunto S' unión de dos semiesferas abiertas con $S' \subset S$, lo cual está en contradicción con que $\sigma_1 B_1 \cup \sigma_2 B_2 = B$.

Para intentar construir una descomposición paradójica de B en 5 piezas, trabajaremos separadamente en cada S_r , $0 < r \leq 1$ (dejamos aparte en un principio S y $\{0\}$). El procedimiento es el habitual, encontraremos dos rotaciones independientes de B y a través del grupo libre que generan, teniendo en cuenta el Teorema 3.5.4 particionaremos cada S_r $0 < r < 1$ en A_i^r , $i = 1, \dots, 4$ verificando $\sigma(A_2^r) = A_2^r \cup A_3^r \cup A_4^r$ y $\tau(A_4^r) = A_1^r \cup A_2^r \cup A_4^r$. Por lo tanto tendríamos una descomposición paradójica en cuatro piezas para la bola abierta menos el origen. Para S necesitaríamos una partición en 5 piezas, A_i^1 , $i = 1, \dots, 4$ y un único punto p , de tal manera que $\sigma(A_2^1) = A_2^1 \cup A_3^1 \cup A_4^1 \cup \{p\}$ y $\tau(A_4^1) = A_1^1 \cup A_2^1 \cup A_4^1 \cup \{p\}$

Podemos construir esa partición seleccionando una órbita, \mathcal{O} , de puntos que no sean fijos. Debe haber una órbita así porque F tiene un conjunto numerable de conjuntos fijos. Por lo tanto seleccionamos $p \in \mathcal{O}$ y asignaremos los otros puntos q en \mathcal{O} a las particiones $A_i^1, i = 1, \dots, 4$ según w comience por $\sigma, \sigma^{-1}, \tau$ y τ^{-1} con $q = w(p)$ siendo w única (ver demostración del Teorema 3.5.4). Con esta elección B ahora puede ser particionada en $B_i, i = 1, \dots, 4$ y $\{p\}$ siendo $B_1 = \{0\} \cup \{A_1^r : 0 < r \leq 1\}$ y $B_i = \{A_i^r : 0 < r \leq 1\}, i = 2, 3, 4$. Considerando ρ la traslación que lleva p a $\{0\}$, obtenemos la descomposición paradójica en 5 piezas: $B_1 \cup \sigma(B_2) = \{0\} \cup (\bigcup_{i=1}^4 \{A_i^r : 0 < r \leq 1\}) = B$ y $B_3 \cup \tau(B_4) \cup \rho(\{p\}) = (\bigcup_{i=1}^4 \{A_i^r : 0 < r \leq 1\}) \cup \{0\} = B$. De la misma forma, (sustrayendo un punto a S para generar el origen de la bola) si consideramos una esfera de radio arbitrariamente grande se puede encontrar una descomposición paradójica de todo el espacio \mathbb{R}^3 . Es más, se puede encontrar una descomposición paradójica en menos piezas (cuatro) ya que existe un grupo libre de isometrías que actúa sobre todo \mathbb{R}^3 sin puntos fijos no triviales.

Tratamiento en otras dimensiones

Una vez expuestos los resultados de la paradoja en 3 dimensiones, cabe preguntarse si en dimensiones superiores o inferiores podemos encontrar una paradoja del mismo tipo. En este capítulo abordaremos este análisis. El problema es bastante diferente para dimensión superior a 3 que para dimensión inferior. De hecho para $n > 3$ el grupo de rotaciones $SO(3)$ está contenido en un grupo de rotación $SO(n)$ y según vimos éste último es paradójico pues contiene al primero, el cual es paradójico. Para dimensiones inferiores a 3 ($n = 1, 2$), sin embargo el hecho de que el grupo de isometrías de la recta o el plano no contenga subgrupos no abelianos (veamos que cualquier transformación isométrica conmuta entre sí en estas dimensiones) hace que no se pueda reproducir la paradoja para dimensión menor que 3. La resolubilidad de G_1 y G_2 , en particular hace posible que exista una medida finitamente aditiva e invariante por isometría, en \mathbb{R} y \mathbb{R}^2 .

4.1. La paradoja de Banach Tarski en dimensión mayor que 3

Teorema 4.1.1. *Supongamos $n > 3$. Se verifica lo siguiente:*

1. *Cualquier esfera en \mathbb{R}^n es paradójica respecto a su grupo de rotaciones.*
2. *Cualquier bola sólida en \mathbb{R}^n es paradójica y también \mathbb{R}^n mismo.*
3. *Cualesquiera dos conjuntos acotados de \mathbb{R}^n con interior no vacío son equidescomponibles.*

Demostración

Presentaremos la demostración para el apartado 1. El resto de apartados se demuestra de una manera similar a cómo se hizo en caso de $n = 3$ (Proposiciones 3.4.1 y 3.4.2). Como el resultado es cierto para $n = 3$, procederemos por inducción. Partimos de la esfera \mathbb{S}^n , la esfera unidad en \mathbb{R}^{n+1} . Supongamos que tenemos una partición $A_i, B_j \subseteq \mathbb{S}^{n-1}$ y una familia de transformaciones σ_i, τ_j con la que se verifica que \mathbb{S}^{n-1} es paradójico. Consideremos las particiones A_i^*, B_j^* de \mathbb{S}^n excluyendo los polos, $\mathcal{S} = \mathbb{S}^n \setminus (0, 0, \dots, 0, \pm 1)$:

1. $A_i^* = \{(x_1, x_2, \dots, x_n, z) \in \mathcal{S} \mid (x_1, x_2, \dots, x_n) / \mid (x_1, x_2, \dots, x_n) \mid \in A_i\}$
2. $B_i^* = \{(x_1, x_2, \dots, x_n, z) \in \mathcal{S} \mid (x_1, x_2, \dots, x_n) / \mid (x_1, x_2, \dots, x_n) \mid \in B_i\}$

Y consideremos las transformaciones extendidas σ_i^*, τ_j^* :

$$\sigma_i^* = \begin{pmatrix} & & & & 0 \\ & & & & \vdots \\ & & \sigma_i & & \vdots \\ & & & & \vdots \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}, \tau_i^* = \begin{pmatrix} & & & & 0 \\ & & & & \vdots \\ & & \tau_i & & \vdots \\ & & & & \vdots \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}$$

De tal manera que $A_i^*, B_i^*, \sigma_i^*, \tau_j^*$ generan una descomposición paradójica de \mathcal{S} . Pero podemos utilizar una rotación de orden infinito, como hicimos en la demostración del Teorema 3.4.1, para demostrar que $\mathcal{S} = \mathbb{S}^n \setminus (0, 0, \dots, 0, \pm 1) \sim 2\mathbb{S}^n$ y por la Proposición 3.4.2 q.e.d.

A continuación veremos que se puede obtener un sistema de generadores de un grupo no abeliano sin puntos fijos para \mathbb{S}^3 tomando como base las rotaciones $SO(4)$.

Teorema 4.1.2. *El grupo $SO(4)$ tiene un subgrupo libre de rango 2 cuya acción sobre la esfera \mathbb{S}^3 no tiene puntos fijos no triviales.*

Demostración

Sea θ un ángulo cuyo coseno es trascendente (por el teorema de Hermite-Lindemann sabemos que $\theta=1$ cumple los requisitos) y sean las rotaciones $\sigma, \tau \in SO(4)$:

$$\sigma = \begin{pmatrix} \cos \theta & -\operatorname{sen} \theta & 0 & 0 \\ \operatorname{sen} \theta & \cos \theta & 0 & 0 \\ 0 & 0 & \cos \theta & -\operatorname{sen} \theta \\ 0 & 0 & \operatorname{sen} \theta & \cos \theta \end{pmatrix}, \tau = \begin{pmatrix} \cos \theta & 0 & 0 & -\operatorname{sen} \theta \\ 0 & \cos \theta & -\operatorname{sen} \theta & 0 \\ 0 & \operatorname{sen} \theta & \cos \theta & 0 \\ \operatorname{sen} \theta & 0 & 0 & \cos \theta \end{pmatrix}$$

Si denotamos por $U(\sigma, \rho)$ aquellas palabras reducidas que comienzan por σ o σ^{-1} y terminan en ρ o ρ^{-1} , observemos que cualquier palabra no trivial w será de la forma $U(\sigma, \rho), U(\rho, \sigma), U(\sigma, \sigma)$ o $U(\rho, \rho)$. Veamos también que $U(\rho, \sigma)$ se reduce a $U(\sigma, \rho)$ si la sometemos a la operación de inversión, por ejemplo si tenemos $\rho\rho\rho\rho^{-1}\sigma$ que es de la forma $U(\rho, \sigma)$, entonces:

$$\rho\rho\rho\rho^{-1}\sigma \xrightarrow{()^{-1}} \sigma^{-1}\rho\rho^{-1}\rho^{-1}\rho^{-1}$$

siendo esta última palabra de la forma $U(\sigma, \rho)$.

Por otro lado si tenemos una palabra de los tipos $U(\sigma, \sigma)$ o $U(\rho, \rho)$, se pueden reducir a una de las otras formas aplicando conjugación repetida. Por ejemplo, sea $w = \rho\rho\rho\rho^{-1}\sigma\rho$ que es de la forma $U(\rho, \rho)$:

$$\rho\rho\sigma\rho^{-1}\sigma\rho \xrightarrow{c} \rho\rho\rho\sigma\rho^{-1}\sigma$$

Basándonos en que:

1. Las potencias de:

$$\sigma_{2 \times 2} = \begin{pmatrix} \cos \theta & -\operatorname{sen} \theta \\ \operatorname{sen} \theta & \cos \theta \end{pmatrix}$$

son de la forma

$$\begin{pmatrix} \cos m\theta & -\operatorname{sen} m\theta \\ \operatorname{sen} m\theta & \cos m\theta \end{pmatrix}$$

2. Al ser $(\cos \theta)$ trascendente, θ no es un múltiplo racional de π (sino sería solución de una ecuación del tipo $e^{2n\pi i} = z^n = 0$) y por lo tanto no tiene puntos fijos,

entonces podemos reducir el estudio de todas las palabras reducidas al estudio de una palabra de tipo $\sigma^{\pm 1} \dots \rho^{\pm 1}$. Veamos que cada matriz $\sigma^{\pm 1}, \rho^{\pm 1}$ tiene la forma:

$$\begin{pmatrix} P & -Q & -R & -S \\ Q & P & -S & -R \\ R & S & P & -Q \\ S & R & Q & P \end{pmatrix} \tag{4.1}$$

siendo P y R polinomios con coeficientes enteros en $\cos \theta$, y Q y S el producto de esos polinomios con $\operatorname{sen} \theta$. Por otro lado veamos que:

$$\sigma_{\tau}^{\pm 1} = \begin{pmatrix} \cos^2 \theta & -\operatorname{sen} \theta \cos \theta & \pm(1 - \cos^2 \theta) & \mp(\operatorname{sen} \theta \cos \theta) \\ \operatorname{sen} \theta \cos \theta & \cos^2 \theta & \mp(\operatorname{sen} \theta \cos \theta) & \mp(1 - \cos^2 \theta) \\ \mp(1 - \cos^2 \theta) & \pm(\operatorname{sen} \theta \cos \theta) & \cos^2 \theta & -\operatorname{sen} \theta \cos \theta \\ \pm(\operatorname{sen} \theta \cos \theta) & \pm(1 - \cos^2 \theta) & \operatorname{sen} \theta \cos \theta & \cos^2 \theta \end{pmatrix}$$

$$\sigma^{\pm 1} \tau = \begin{pmatrix} \cos^2 \theta & \operatorname{sen} \theta \cos \theta & \mp(1 - \cos^2 \theta) & \mp(\operatorname{sen} \theta \cos \theta) \\ -\operatorname{sen} \theta \cos \theta & \cos^2 \theta & \mp(\operatorname{sen} \theta \cos \theta) & \pm(1 - \cos^2 \theta) \\ \pm(1 - \cos^2 \theta) & \pm(\operatorname{sen} \theta \cos \theta) & \cos^2 \theta & \operatorname{sen} \theta \cos \theta \\ \pm(\operatorname{sen} \theta \cos \theta) & \mp(1 - \cos^2 \theta) & -\operatorname{sen} \theta \cos \theta & \cos^2 \theta \end{pmatrix}$$

de donde podemos deducir que toda palabra en $\sigma^{\pm 1} \dots \rho^{\pm 1}$ tiene la forma mencionada. Calculemos el polinomio característico de la matriz (4.1):

$$\lambda^4 - 4P\lambda^3 + (6P^2 + 2(Q^2 + R^2 + S^2))\lambda - 4P(P^2 + Q^2 + R^2 + S^2)\lambda + (P^2 + Q^2 + R^2 + S^2) = 0$$

Como (4.1) es ortogonal, el elemento a_{11} del producto de la matriz por su inversa es 1, esto es:

$$P^2 + Q^2 + R^2 + S^2 = 1$$

y la ecuación característica se convierte en:

$$\lambda^4 - 4P\lambda^3 + (4P^2 + 2)\lambda - 4P\lambda + 1 = 0$$

Si existiera algún punto fijo tendríamos $\lambda=1$ y por lo tanto la ecuación:

$$4P^2 - 8P + 4 = 0$$

tendría solución. Pero la anterior ecuación es una ecuación en $\cos \theta$ y ya que $\cos \theta$ era trascendente por hipótesis, sólo nos queda demostrar que P , que es un polinomio en $\cos \theta$ no es una constante, lo cual haremos a continuación.

Analicemos la forma de σ^m :

$$\sigma^m = \begin{pmatrix} \cos m\theta & \mp \text{sen } m\theta & 0 & 0 \\ \pm \text{sen } m\theta & \cos m\theta & 0 & 0 \\ 0 & 0 & \cos m\theta & \mp \text{sen } m\theta \\ 0 & 0 & \pm \text{sen } m\theta & \cos m\theta \end{pmatrix}$$

Si tenemos en cuenta las identidades:

$$\cos m\theta = 2^{m-1} \cos^m \theta + \Gamma(\cos \theta)$$

$$\text{sen } m\theta = \text{sen } \theta (2^{m-1} \cos^{m-1} \theta + \Delta(\cos \theta))$$

donde $\Gamma(\cos \theta)$ y $\Delta(\cos \theta)$ representan polinomios en $\cos \theta$ de orden menor que m . A partir de ahora trabajaremos con los términos de orden mayor. De esta manera se tiene:

$$\sigma^{\pm m} = 2^{m-1} \cos^{m-1} \theta \begin{pmatrix} \cos \theta & \mp \text{sen } \theta & 0 & 0 \\ \pm \text{sen } \theta & \cos \theta & 0 & 0 \\ 0 & 0 & \cos \theta & \mp \text{sen } \theta \\ 0 & 0 & \pm \text{sen } \theta & \cos \theta \end{pmatrix}$$

Hay otra representación similar para τ^k . Si multiplicamos las dos representaciones:

$$\sigma^{\epsilon m} \tau^{\delta k} = 2^{m+k-2} \cos^{m+k-1} \theta \begin{pmatrix} \cos \theta & -\epsilon \text{sen } \theta & -\epsilon \delta \cos \theta & -\delta \text{sen } \theta \\ \epsilon \text{sen } \theta & \cos \theta & -\delta \text{sen } \theta & \epsilon \delta \cos \theta \\ \epsilon \delta \cos \theta & \delta \text{sen } \theta & \cos \theta & -\epsilon \text{sen } \theta \\ \delta \text{sen } \theta & -\epsilon \delta \cos \theta & \epsilon \text{sen } \theta & \cos \theta \end{pmatrix}$$

siendo $\epsilon, \delta = \pm 1$.

Ahora demostraremos que si $w = \sigma^{\epsilon_n m_n} \tau^{\delta_n k_n} \dots \sigma^{\epsilon_1 m_1} \tau^{\delta_1 k_1}$ y $\Sigma = |k_1| + |m_1| + \dots + |k_n| + |m_n|$, entonces:

$$w = 2^{\Sigma-n-1} \cos^{\Sigma-1} \theta \begin{pmatrix} \xi \cos \theta & -\mu \sin \theta & -\zeta \cos \theta & -\nu \sin \theta \\ \mu \sin \theta & \xi \cos \theta & -\nu \sin \theta & \zeta \cos \theta \\ \zeta \cos \theta & \nu \sin \theta & \xi \cos \theta & -\mu \sin \theta \\ \nu \sin \theta & -\zeta \cos \theta & \mu \sin \theta & \xi \cos \theta \end{pmatrix} \quad (4.2)$$

donde $\xi, \mu, \nu, \zeta = \pm 1$. Demostremos esto último por inducción. Supongamos que w tiene la forma correcta y calculemos $(\sigma^{\epsilon m} \tau^{\delta k})w$. Al ser una matriz antisimétrica y debido a la relación entre los diversos elementos, sólo nos hará falta calcular los elementos $a_{11}, a_{12}, a_{13}, a_{14}$:

1. $a_{11} = \kappa(-\epsilon \mu_n - \delta \nu_n + (\xi_n + \epsilon \mu_n + \delta \nu_n - \epsilon \delta \zeta_n) \cos \theta)$
2. $a_{12} = \kappa(-\mu_n - \epsilon \xi_n - \epsilon \delta \nu_n + \delta \zeta_n) \sin \theta$
3. $a_{13} = \kappa(-\epsilon \nu_n - \delta \mu_n + (-\zeta_n + \epsilon \nu_n + \delta \mu_n - \epsilon \delta \xi_n) \cos \theta)$
4. $a_{14} = \kappa((- \nu_n - \epsilon \zeta_n - \delta \xi_n + \epsilon \delta \mu_n) \sin \theta)$

donde $\kappa = 2^{\Sigma^*-n-2} \cos^{\Sigma^*-1} \theta$ y $\Sigma^* = |k_1| + |m_1| + \dots + |k_n| + |m_n| + |k| + |m|$.

Teniendo en cuenta que $\xi, \mu, \nu, \zeta = \pm 1$, $a_{11} = \xi_{n+1}$, $a_{12} = \mu_{n+1}$, $a_{13} = \zeta_{n+1}$, $a_{14} = \nu_{n+1}$, y obviando un factor 2, obtenemos las siguientes relaciones

1. $\xi_{n+1} = \xi_n + \epsilon \mu_n + \delta \nu_n - \epsilon \delta \zeta_n$
2. $\mu_{n+1} = \mu_n + \epsilon \xi_n + \delta \zeta_n - \epsilon \delta \nu_n$
3. $\zeta_{n+1} = \zeta_n - \epsilon \nu_n + \delta \mu_n + \epsilon \delta \xi_n$
4. $\nu_{n+1} = \nu_n - \epsilon \zeta_n + \delta \xi_n + \epsilon \delta \mu_n$

Por otro lado, de la condición de ortogonalidad se puede demostrar que para $\sigma^{\epsilon m} \tau^{\delta k}$ se cumple $\mu \nu = \xi \zeta$ y esto permanece cierto para $\sigma^{\epsilon m} \tau^{\delta k}$. Aplicando esta propiedad:

$$\xi_{n+1} = \xi_n + \epsilon \mu_n + \delta \nu_n - \frac{\epsilon \delta \mu_n \nu_n}{\xi_n} = \xi_n + \epsilon \mu_n + \delta \nu_n - \xi_n (\epsilon \mu_n) (\delta \nu_n) \quad (4.3)$$

Se puede ver que para cualquier combinación de los valores $\{\epsilon, \delta, \mu_n, \nu_n, \xi_n\} \in \{-1, 1\}$ la ecuación (4.3) toma el valor 2. Lo mismo ocurre para μ_{n+1}, ζ_{n+1} y ν_{n+1} . Sacando factor común 2 se demuestra finalmente (4.2).

Por lo que finalmente queda demostrado que el elemento $a_{11}(= P)$ de una palabra genérica en $\sigma^{\pm 1}, \rho^{\pm 1}$ es un polinomio no constante en $\cos \theta$ y la ecuación característica de la representación de esa palabra no tendrá solución por ser $\cos \theta$ trascendente. Esto nos lleva a la conclusión final que w no tiene autovalores 1 y por lo tanto no hay puntos fijos no triviales.

Corolario 4.1.3. *Si $n \geq 3$ y $n \neq 5$, entonces $SO(n)$ tiene un subgrupo libre de rango 2 que es localmente conmutativo en su acción sobre \mathbb{S}^{n-1} . Si n es un múltiplo de 4 entonces $SO(n)$ tiene un subgrupo libre de rango 2 sin ningún punto fijo no trivial sobre \mathbb{S}^{n-1} .*

Demostración

Veamos que cualquier $n \geq 3$, excepto $n = 5$ puede ser escrito como $3k + 4l$ con k, l enteros positivos. Sean σ_3, τ_3 y σ_4, τ_4 los pares de generadores independientes para $SO(3)$ y $SO(4)$ respectivamente. Podemos definir $\sigma, \tau \in SO(n)$ usando los anteriores generadores en subespacios de dimensiones 3 y 4 respectivamente de \mathbb{R}^n , conforme a la descomposición de n en $3k + 4l$. Por lo tanto la matriz de σ será diagonal por bloques con k bloques de dimensión 3×3 $(\sigma_3s)^1$, y l bloques de dimensión 4×4 de n en $3k + 4l$. Por lo tanto la matriz de σ será diagonal por bloques con k bloques de dimensión 3×3 (σ_3s) y l bloques de dimensión de dimensión 3×3 (σ_3s) , y l bloques de dimensión 4×4 (σ_4s) . Para $k = 0$ la inexistencia de puntos fijos no triviales asegura el mismo hecho para rotaciones representadas por la matriz compuesta únicamente por bloques de σ_4 en espacios \mathbb{R}^n con n múltiplo de 4. Supongamos sin embargo que $k > 0$, y sean u, v dos palabras que tienen un punto fijo común. Como comparten un punto fijo en \mathbb{S}^{n-1} , deben compartir un punto fijo no trivial en uno de esos k subespacios de dimensión 3. Pero esto implicaría que u y v conmutan en la parte de los generadores σ_3, τ_3 . La independencia de σ_3 y τ_3 implicaría que u y v conmutarían consideradas como palabras formadas por $\sigma^{\pm 1}, \tau^{\pm 1}$.

Veamos que los grupos de transformaciones $SO(4n + 2)$ y $SO(5)$ han quedado fuera del análisis anterior. Se puede demostrar que $SO(m)$, con m impar tiene un autovalor $+1$ y por lo tanto un punto fijo en \mathbb{S}^{m-1} . Por otro lado para el caso $SO(4n + 2)$, también es demostrable la existencia de grupos libres de rango 2 sin puntos fijos no triviales[3].

4.2. La paradoja de Banach-Tarski en dimensión menor que 3

La inexistencia de grupos libres no abelianos construidos con isometrías en la recta y en el plano es la principal causa de que no se verifique la paradoja de Banach-Tarski en dimensiones 1 y 2. Como apuntamos al principio de este capítulo, el hecho de que los grupos de isometrías en el espacio euclídeo de 1 y 2 dimensiones, G_1 y G_2 , sean resolubles tiene como consecuencia la existencia de medidas totales ² finitamente aditivas invariantes segun isometrías. Sin embargo, existe la opción de manejar un grupo más amplio de transformaciones, *el grupo de las transformaciones afines*, A_n . Éste consta de todas las transformaciones de la forma $\sigma = \tau l$, siendo τ una traslación y l una transformación lineal. Se puede

¹Plural de σ_3

²Medidas definidas sobre todos los posibles conjuntos del espacio considerado, en este caso $\mathcal{P}(\mathbb{R}^n)$, $n = 1, 2$

definir el subgrupo especial SA_n como el grupo de las transformaciones afines de determinante $+1$, esto es, aquellas transformaciones afines que conservan la orientación y el área. Es intuitivo ver que para dimensión 1 el grupo G_1 coincide con A_1 no siendo así obviamente para el resto de dimensiones. A continuación se muestra un ejemplo de subgrupo libre no abeliano en SA_2 .

Proposición 4.2.1. *Las transformaciones pertenecientes a $SL_2(\mathbb{Z}) (\subset SA_2(\mathbb{Z}))$ ³:*

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

son independientes.

Demostración

Es fácil comprobar que se cumplen las siguientes relaciones:

$$A^n = \begin{pmatrix} 1 & 2n \\ 0 & 1 \end{pmatrix}, B^n = \begin{pmatrix} 1 & 0 \\ 2n & 1 \end{pmatrix}$$

$$A^n B^{n'} = \begin{pmatrix} 4nn' + 1 & 2n \\ 2n' & 1 \end{pmatrix}, B^n A^{n'} = \begin{pmatrix} 1 & 2n' \\ 2n & 4nn' + 1 \end{pmatrix}$$

donde $n, n' \in \mathbb{Z} \setminus \{0\}$

Teniendo en cuenta lo anterior, es trivial comprobar que cualquier palabra reducida en $A^{\pm 1}, B^{\pm 1}$ no podrá ser igual a la identidad y por lo tanto el sistema $\{A, B\}$ es independiente.

Como la acción de $SL_2(\mathbb{Z})$ es localmente conmutativa sobre $\mathbb{R}^2 \setminus \{0\}$ entonces $\mathbb{R}^2 \setminus \{0\}$ es $SL_2(\mathbb{Z})$ -paradójico. Además como $\mathbb{R}^2 \setminus \{0\} \sim_{\tau} \mathbb{R}^2$ siendo $\tau \in SA_2(\mathbb{Z})$ la traslación $\tau : (x, y) \rightarrow (x + 1, y)$ concluimos que \mathbb{R}^2 es SA_2 -paradójico. Una cuestión diferente es si se permite una descomposición paradójica de \mathbb{R}^2 en 4 piezas. De hecho $SA_2(\mathbb{Z})$ no tiene un subgrupo libre de rango 2 cuya acción sobre \mathbb{R}^2 sea localmente conmutativa por lo que la respuesta a esa cuestión sería negativa.

Descomposición paradójica en \mathbb{R}^2

Nuestro siguiente objetivo es encontrar una descomposición paradójica de un conjunto medible de área finita positiva. Sea \mathcal{J} el cuadrado unidad semiabierto $[0, 1) \times [0, 1)$. Existe una dificultad inicial y es que el grupo $SL_2(\mathbb{Z})$ no actúa en el cuadrado. Por ejemplo, si:

³ $SL_2(\mathbb{Z})$ es el grupo especial de transformaciones lineales definidas por $Ax, x \in \mathbb{R}^2$ siendo A las matrices cuadradas de dimensión 2×2 con determinante $+1$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \in \mathcal{J} \implies A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ \frac{1}{2} \end{pmatrix} \notin \mathcal{J}$$

La solución será utilizar una versión en el plano de aritmética módulo 1. A estos efectos definiremos la siguiente relación de equivalencia en \mathbb{R}^2 : $P \approx Q$ si y sólo si $Q = P + (m, n)$ con $m, n \in \mathbb{Z}$.

Sea $SA_2(\mathbb{Z})$ el grupo de las transformaciones afines de la forma τl siendo τ cualquier traslación y $l \in SL_2(\mathbb{Z})$ y sea $P \in \mathbb{R}^2$ y $\hat{P} \in \mathcal{J}$ el único punto en \mathcal{J} tal que $\hat{P} \approx P$. Para $\sigma \in SA_2(\mathbb{Z})$, consideremos la función $\hat{\sigma} : \mathcal{J} \rightarrow \mathcal{J}$ definida por $\hat{\sigma}(P) = \widehat{\sigma(P)}$.

Proposición 4.2.2. *La aplicación $\sigma \rightarrow \hat{\sigma}$ es un homomorfismo del grupo de biyecciones afines a trozos y que conservan el área, de \mathcal{J} en \mathcal{J} . Cuando el dominio de esta aplicación se restringe a $SL_2(\mathbb{Z})$, entonces es un isomorfismo sobre su imagen.*

Demostración

Mostraremos primero que se verifica lo siguiente: si $\sigma \in SL_2(\mathbb{Z})$ y $P \approx Q$ entonces $\sigma(P) \approx \sigma(Q)$. Teniendo en cuenta que para $\sigma \in SA_2(\mathbb{Z})$, trivialmente se cumple: $\sigma(\mathbb{Z}^2) \subseteq \mathbb{Z}^2$, entonces si $P = Q + R$, $R \in \mathbb{Z} \times \mathbb{Z}$, esto implica $\sigma(P) = \sigma(Q) + \sigma(R)$ con $\sigma(R) \in \mathbb{Z} \times \mathbb{Z}$. Por lo tanto $\sigma(P) \approx \sigma(Q)$. De igual manera como $\sigma^{-1} \in SL_2(\mathbb{Z})$ y $\sigma^{-1}(\mathbb{Z}^2) \subseteq \mathbb{Z}^2$ entonces si $P \approx Q$ se deduce $\sigma^{-1}(P) \approx \sigma^{-1}(Q)$.

Comprobaremos que $\hat{\sigma}$ es biyectiva sobre \mathcal{J} . Veamos primero que es inyectiva. Supongamos que no lo es. Existen entonces $P, P' \in \mathcal{J}$, $P \neq P'$ y $\hat{\sigma}(P) = \hat{\sigma}(P')$. Como $P, P' \in \mathcal{J}$, $\hat{\sigma}(P) = \hat{\sigma}(P') \implies \sigma(P) = \sigma(P')$. Como σ es lineal se tiene $\sigma(P) - \sigma(P') = 0$ lo cual implica $P = P'$ y esto es una contradicción. Por lo tanto $\hat{\sigma}$ es inyectiva. Por otro lado para ver la suprayectividad, sea $P' \in \mathcal{J}$. Entonces $P' = \widehat{P'} = \widehat{\sigma\sigma^{-1}(P')} = \widehat{\sigma(\widehat{\sigma^{-1}(P')})}$, con $\widehat{\sigma^{-1}(P')} \in \mathcal{J}$.

Es intuitivo comprobar además que la aplicación $\hat{\sigma}$ conserva el área y es afín a trozos pues σ tiene estas propiedades. De hecho, $\hat{\sigma}$ se puede descomponer en muchas funciones $\tau_i\sigma$ siendo τ_i una traslación por un punto en $\mathbb{Z} \times \mathbb{Z}$.

Ahora, probaremos que $\sigma \rightarrow \hat{\sigma}$ es un homomorfismo. Si $P \in \mathcal{J}$ se puede ver que la propiedad $(\widehat{\quad})$ es conservada por la composición:

$$\widehat{\sigma_1\widehat{\sigma_2(P)}} = \widehat{\sigma_1\sigma_2(P)} = \widehat{\sigma_1(\widehat{\sigma_2(P)})} = \widehat{\sigma_1\sigma_2(P)} = \widehat{\sigma_1\sigma_2(P)}$$

Y por último comprobaremos que $l \rightarrow \hat{l}$ es un isomorfismo, pero esto se ve de manera sencilla. Para ello veremos que la única transformación de $SL_2(\mathbb{Z})$ que actúa como la identidad en \mathcal{J} es la identidad en $SL_2(\mathbb{Z})$. Sea $l \in SL_2(\mathbb{Z})$ tal que $\hat{l} = \mathcal{I}_{\mathcal{J}}$ (identidad en \mathcal{J}). Como hemos dicho antes, \hat{l} se puede escribir como τl sobre algún polígono y por tanto $\mathcal{I}_{\mathcal{J}} = \hat{l} = \tau l$. Como τl es afín, esto implica que $\tau l = \mathcal{I}$, luego necesariamente $l = \mathcal{I}$.

La consecuencia de la anterior proposición es que si G' es un subgrupo libre de $SL_2(\mathbb{Z})$ entonces $\widehat{G'} = \{\widehat{g'} : g' \in G'\}$ es un subgrupo libre con el mismo rango actuando sobre \mathcal{J} . El procedimiento para llegar a la paradoja será básicamente el mismo que el seguido en el capítulo anterior para el caso de \mathbb{S}^2 . Se encontrará primero una descomposición paradójica de todo el plano menos los puntos fijos (entre ellos

el origen) y luego eliminaremos el conjunto de puntos fijos, encontrando un conjunto que los contenga y sea equidescomponible con su diferencia respecto al mismo conjunto de puntos fijos (ver demostración del Teorema 3.4.1).

La paradoja de Von Neumann en el plano

Manejando los conceptos ya expuestos en la sección inmediatamente anterior, enunciaremos a continuación la existencia de una descomposición paradójica para el cuadrado unidad \mathcal{J} .

Teorema 4.2.3. *Sea G el subgrupo de $SA_2(\mathbb{Z})$ generado por σ_1, σ_2 , transformaciones independientes de $SL_2(\mathbb{Z})$ y \mathcal{T} el grupo de todas las traslaciones, entonces \mathcal{J} es G -paradójico. Además dos subconjuntos acotados cualesquiera de \mathbb{R}^2 con interior no vacío son G -equidescomponibles.*

Demostración

Sea G el grupo generado por σ_1, σ_2 y \mathcal{T} el grupo de todas las traslaciones, y sean las imágenes de \widehat{G} y \widehat{T} por la reducción módulo \mathbb{Z}^2 . Si D es el conjunto de puntos en \mathcal{J} fijados por algún elemento no trivial de \widehat{G} , entonces \widehat{G} actúa sin puntos fijos no triviales en $\mathcal{J} \setminus D$. Ya que \widehat{G} es generado libremente por σ_1, σ_2 entonces $\mathcal{J} \setminus D$ es G -paradójico. Finalmente sólo quedará demostrar que $\mathcal{J} \sim_{\mathcal{T}} \mathcal{J} \setminus D$.

Ya que todo punto en D es un punto fijo de un conjunto afín y por lo tanto es un subespacio afín, y como G es numerable, D estará formado por un conjunto numerable de puntos D_0 y un conjunto numerable de segmentos D_1 . Empleando la misma técnica que la utilizada en la demostración del Teorema 3.4.1 podemos demostrar que para cualquier conjunto numerable C de \mathcal{J} , $\mathcal{J} \sim_{\mathcal{T}} \mathcal{J} \setminus C$ usando dos piezas. Basta con encontrar una traslación τ que verifique $C \cap \widehat{\tau}^n(C) = \emptyset$ con $n \geq 1$.

Veremos ahora cómo construir esa traslación. Como D_0 ya es numerable nos concentraremos en D_1 . En primer lugar enumeremos los segmentos en D_1 como S_0, S_1, \dots . Sean n, i y j enteros tales que $0 \leq i, j$, $i \neq j$. Si S_i y S_j no son paralelos, no hay problema porque intersectan a lo sumo en un punto y por lo tanto $\widehat{\tau}^n(S_j)$ también tendrá sólo un punto de intersección con S_i . Si S_i y S_j son paralelos, entonces $\widehat{\tau}^n(S_j)$ podría coincidir con S_i y por lo tanto su intersección sería una recta la cual tiene un conjunto no numerable de puntos. Esto lo podemos solventar si elegimos $p \in S_j$ y excluimos todas las traslaciones τ tales que para algún $(m_1, m_2) \in \mathbb{Z} \times \mathbb{Z}$, $\tau^n(p)$ está en la doble línea que contiene a $p' + S_i$ con $p' \in \mathbb{Z} \times \mathbb{Z} \subset \mathbb{R}^2$. De esta manera, es inmediato ver que τ pertenece al conjunto de traslaciones a considerar si y sólo si $S_i \cap \widehat{\tau}^n(S_j) = \emptyset$. Como partimos inicialmente de un conjunto numerable de segmentos habrá un conjunto no numerable de traslaciones posibles. Observemos que en cada línea de D_1 que cumple el requisito de que no intersecta a otra de D_1 , sólo descartamos un conjunto numerable de traslaciones y debe haber, al menos, una traslación τ con las propiedades que deseamos. Por lo tanto $D_1 \cap \widehat{\tau}^n(D_1) = \emptyset$ es numerable ya que si no lo fuera entonces algún segmento en D_1 se solaparía con alguno en $\widehat{\tau}^n(D_1)$, lo cual no puede ocurrir ya que hemos excluido todas aquellas traslaciones que llevan un segmento de D_1 a otro de D_1 .

Sea por tanto $C = \bigcup \{D \cap \widehat{\tau}^n(D) : n \in \mathbb{Z} \setminus \{0\}\}$. C es un subconjunto numerable de D y los conjuntos $\widehat{\tau}^n(D \setminus C)$, $n \geq 0$ son disjuntos por pares y disjuntos de C , luego si A es el complemento de $B = \widehat{\tau}^n(D) \setminus C$ en $\mathcal{J} \setminus C$:

$$\begin{aligned} \mathcal{J} \setminus C &= \left(\bigcup_{i=0}^{\infty} \hat{\tau}^n(D \setminus C) \right) \cup A = \hat{\tau} \left(\left(\bigcup_{i=1}^{\infty} \hat{\tau}^n(D \setminus C) \right) \cup A \right) \implies \mathcal{J} \setminus C \sim \left(\bigcup_{i=1}^{\infty} \hat{\tau}^n(D \setminus C) \right) \cup A \\ &= (\mathcal{J} \setminus C) \setminus (D \setminus C) = \mathcal{J} \setminus D \end{aligned}$$

De una manera similar a cómo se hizo en la demostración de la Proposición 3.4.2, como $\mathcal{J} \setminus D$ es paradójico y $\mathcal{J} \setminus C \sim \mathcal{J} \setminus D$, entonces $\mathcal{J} \sim \mathcal{J} \setminus D$, y como $\mathcal{J} \setminus D$ es G -paradójico entonces por la Proposición 3.3.5 \mathcal{J} es G -paradójico.

Aunque la demostración ha sido realizada para el cuadrado unidad \mathcal{J} se puede extrapolar para cuadrados de tamaño arbitrario utilizando transformaciones s de la forma $s(x, y) = (\kappa x \ \kappa y)$, $\kappa > 0$, y tales que $sG^{-1}s^{-1} \in G$. Para demostrar la equidescomponibilidad de dos conjuntos acotados cualesquiera con interior no vacío se utiliza el teorema de Banach-Schröder-Bernstein como se aplicó en la demostración del Teorema 3.4.3 utilizando cuadrados de cualquier tamaño en lugar de bolas.

Descomposición paradójica en \mathbb{R}

El intentar encontrar descomposiciones paradójicas en la recta real presenta un problema que no aparecía cuando se intentaba lo mismo en el plano. Las transformaciones afines en la recta se reducen a aplicaciones del tipo $x + a$ con $a \in \mathbb{R}$ las cuales son isometrías conmutativas en este espacio, por lo que no son útiles para nuestro objetivo. Una solución alternativa es considerar el grupo $G(\lambda)$ de todas las biyecciones $f: \mathbb{R} \rightarrow \mathbb{R}$ siendo f y f^{-1} medibles Lebesgue y que conservan la medida de Lebesgue λ , esto es: $\lambda(f(A)) = \lambda(A) = \lambda^{-1}(f(A))$ siendo A medible Lebesgue. Este grupo contiene al grupo de las isometrías y es capaz de proporcionar descomposiciones paradójicas en la recta.

Teorema 4.2.4. (AE)

Cualquier intervalo de la recta real es $G(\lambda)$ -paradójico. Dos subconjuntos acotados de \mathbb{R} cualesquiera con interior no vacío son $G(\lambda)$ -equidescomponibles.

Demostración

Un importante resultado de teoría de la medida dice que existe una biyección $f: [0, 1) \rightarrow \mathbb{S}^2$ donde f y f^{-1} convierten conjuntos medibles en conjuntos medibles y donde también f y f^{-1} conservan la medida, considerando ésta la medida Lebesgue normalizada a la superficie de la esfera ($\lambda(\mathbb{S}^2) = 1$). Apoyándonos en esta idea, es demostrable que para una transformación $\sigma \in SO_3$, $f^{-1}\sigma f$ es una biyección que conserva la medida de $[0, 1)$ en sí mismo, por lo cual si \mathbb{S}^2 es SO_3 -paradójica con las transformaciones σ_i y los subconjuntos A_i , entonces $[0, 1)$ sera $G(\lambda)$ -paradójico con piezas $f^{-1}(A_i)$ y transformaciones $f^{-1}\sigma f$.

La paradoja de Von Neumann en la recta

En 1929, el matemático John Von Neumann encontró una clase diferente de paradoja en la recta real utilizando transformaciones lineales fraccionarias. Concretamente el grupo a considerar es L , grupo de las transformaciones sobre $\mathbb{R} \cup \{\infty\}$ de la forma $x \rightarrow \frac{ax+b}{cx+d}$, $ad - bc = 1$. Es fácil observar que $\sigma \in L$ representa

una acción de grupo sobre $\mathbb{R} \cup \{\infty\}$ y es una función creciente en cualquier intervalo que no contenga a $x = -\frac{d}{c}$. Además como el semiplano superior es isomorfo a PSL_2 ⁴, luego no habrá problema de ahora en adelante, si para cada $\sigma \in L$ establecemos una correspondencia con los coeficientes de la matriz

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2$$

Teorema 4.2.5. *Sean los números algebraicamente independientes $a_k, b_k, c_k, d_k, k = 1, \dots, n$ sobre los números racionales y sean las transformaciones lineales fraccionarias de la forma $\sigma_k : x \rightarrow \frac{a_k x + b_k}{c_k x + d_k}, a_k d_k - b_k c_k = 1, a_k, b_k, c_k, d_k \in \mathbb{R}$, entonces σ_k generan un grupo libre.*

Demostración

Debemos probar que si $m_1, \dots, m_r \in \mathbb{Z} \setminus \{0\}$ y $k_i \neq k_{i-1}, i = 2, \dots, r$, entonces $\alpha = \sigma_{k_1}^{m_1} \sigma_{k_2}^{m_2} \dots \sigma_{k_n}^{m_n}$ no es la identidad. Es sencillo comprobar que $\alpha(x) = \frac{Ax+B}{Cx+D}$, siendo A, B, C y D polinomios en a_k, b_k, c_k, d_k con coeficientes enteros y $AD - BC \neq 0$. Si α fuera la identidad entonces se tendría que cumplir $B = C = 0$ y $A = D$. Luego entonces como a_k, b_k, c_k, d_k son algebraicamente independientes, entonces estas ecuaciones son identidades. Esto implica si n es el numero de índices diferentes entre k_1, \dots, k_r entonces para transformaciones lineales fraccionarias arbitrarias $\beta_1, \dots, \beta_n, \beta_{k_1}^{m_1} \beta_{k_2}^{m_2} \dots \beta_{k_n}^{m_n}$ es la identidad. Pero se puede comprobar que si tenemos las transformaciones lineales fraccionarias $\sigma_A = x + 2, \sigma_B = \frac{1}{2x+1}$ asociadas a las matrices A, B de la Proposición 4.2.1 al ser éstas últimas independientes, como vimos, $\sigma_A, \sigma_B \in L$ también lo son y por lo tanto no existen m_1, \dots, m_n tal que la transformación lineal fraccionaria asociada $\beta_{k_1}^{m_1} \beta_{k_2}^{m_2} \dots \beta_{k_n}^{m_n}$ con $\beta_{k_i} = \sigma_A$ o $\sigma_B, i = 1, \dots, n$ actúe sobre x como la identidad.

Definición 4.2.6. Una transformación $\sigma \in L$ es una ϵ -contracción (o Lipschitz) con respecto a un intervalo \mathcal{I} de \mathbb{R} si $|\sigma(x) - \sigma(y)| \leq \epsilon |x - y| \forall x, y \in \mathcal{I}$.

Definición 4.2.7. Una transformación $\sigma \in L$ es una contracción a trozos si existe una partición finita A_1, \dots, A_n tal que la restricción $f|_{A_i}$ es una contracción $\forall i = 1, \dots, n$.

Lema 4.2.8. *Si α, β son dos cualesquiera números algebraicos independientes, tal que $0 < \alpha < \beta < 1$ entonces $\hat{\sigma}$ y $\hat{\tau}$ son biyecciones independientes de $[0, \gamma)$ donde:*

$$\gamma = \frac{\beta(2-\beta)}{(1+\beta)(2-\beta)}, \hat{\sigma}(x) = \frac{(1+\beta)x}{\frac{1}{1+\beta} + (2-\beta)x}$$

$$\hat{\tau}(x) = \begin{cases} x + \alpha & \text{si } x + \alpha < \gamma \\ x + \alpha - \gamma & \text{en el caso contrario} \end{cases}$$

⁴grupo especial lineal proyectivo representado por las matrices cuadradas de dimensión 2×2 con determinante distinto de cero

Demostración

En primer lugar veamos que las matrices asociadas a las transformaciones lineales fraccionarias definidas ($\hat{\sigma}$ y $\hat{\tau}$) son:

$$M_{\hat{\sigma}} = \begin{pmatrix} 1 + \beta & 0 \\ 2 - \beta & \frac{1}{1 + \beta} \end{pmatrix}, M_{\hat{\tau}}^+ = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}, M_{\hat{\tau}}^- = \begin{pmatrix} 1 & \alpha - \gamma \\ 0 & 1 \end{pmatrix} \in SL_2$$

Es trivial ver que $\hat{\tau}(x)$ es la traslación módulo γ de la traslación $\tau(x) = x + \alpha$. Por otro lado $\hat{\sigma}$ tiene como dominio $[0, \gamma)$ donde 0 y γ son puntos fijos. Sea una palabra w en $\hat{\sigma}^{\pm 1}, \hat{\tau}^{\pm 1}$ que es igual a la identidad en $[0, \gamma)$. Como w es la identidad y γ es una función racional de β habrá cuatro funciones racionales resultantes de la multiplicación de potencias de $M_{\hat{\sigma}}, M_{\hat{\tau}}^- y M_{\hat{\tau}}^+$ en w , $R_{ij}(\alpha, \beta, \gamma)$, $i, j = 1, 2$ que deben verificar $R_{ij} = \delta_{ij}$. Además siendo α y β algebraicamente independientes esas funciones racionales deben ser iguales a 0 ó 1, y por lo tanto w será la identidad independientemente de qué valores tengan α y β . Eligiendo $\alpha = 2$ y $\beta = 0$ resulta $\gamma = 0$ y entonces tendríamos w sería una palabra en:

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

pero esto es una contradicción ya que estas dos matrices son independientes (ver Proposición 4.2.1)

Teorema 4.2.9. (Paradoja de Von Neumann para la recta)(AE) *El intervalo unidad es paradójico usando contracciones, de hecho para cualquier ϵ existe una descomposición paradójica que emplea ϵ -contracciones con respecto a $[0, 1]$ para transformar los conjuntos que particionan el intervalo. Además para cualesquiera conjuntos A, B acotados cuyo interior es no vacío, si $\{A_i\}_1^n$ es una partición de A entonces existe una familia de ϵ -contracciones σ_i con respecto a un intervalo que contiene a A , de tal manera que $\sigma_i\{A_i\}$ es una partición de B .*

Demostración

Definiremos inicialmente una ϵ -contracción a trozos de $[0, 1]$ a un conjunto que contenga a $[0, 2]$. Elegiremos una transformación que contraiga $[0, 1]$ a $[0, \gamma]$ e iremos aplicando las transformaciones del Lema 4.2.8 para ir cubriendo $[0, 2]$. En primer lugar sean $\alpha, \beta, \gamma, \hat{\sigma}$ y $\hat{\tau}$ conforme al Lema 4.2.8. Como el conjunto de puntos fijos es numerable (sólo dos puntos) se puede aplicar la técnica aplicada en el Teorema 3.4.1, es decir encontrar una transformación τ_1 tal que $[0, \gamma]$ sea equidescomponible con $[0, \gamma) \setminus D$. Aplicando la Proposición 3.1.4 y la Proposición 3.3.5 podremos afirmar que $[0, \gamma]$ es paradójico con las transformaciones consideradas. Si hacemos la primera derivada de la contracción $\hat{\sigma}$:

$$\frac{d\sigma}{dx} = \frac{1}{[(2-\beta)x + \frac{1}{1+\beta}]^2}$$

que toma su valor máximo en $x = 0$ siendo $(1 + \beta)^2$. Aplicando el teorema del valor medio:

$$|\sigma(x) - \sigma(y)| \leq (1 + \beta)^2 |x - y|$$

Entonces el procedimiento sería primero encoger $[0, 1]$ a $[0, \gamma]$ y luego aplicar la descomposición paradójica de $[0, \gamma]$ el suficiente número de veces m como para cubrir $[0, 2]$. Como en cada iteración se multiplica por dos el número de copias existente se duplica el intervalo luego se requerirán en total $m = -\log_2(\gamma)$. De esta forma, la transformación lineal final por trozos de $[0, 1]$ a un conjunto que contenga a $[0, 2]$, obtenida mediante el procedimiento anterior encogerá las distancias por un factor de $\gamma(1 + \beta)^{2m}$. Estudiemos el límite de la anterior expresión conforme $\beta \rightarrow 0$ Veamos que:

$$\gamma(1 + \beta)^{2m} = \frac{\gamma}{\beta}(\beta)(1 + \beta)^{-2\log_2(\frac{\gamma}{\beta}\beta)}$$

Como conforme $\beta \rightarrow 0$ $\frac{\gamma}{\beta} \rightarrow 1$, entonces:

$$\gamma(1 + \beta)^{2m} = \frac{\gamma}{\beta}(\beta)(1 + \beta)^{-2\log_2(\frac{\gamma}{\beta}\beta)} \xrightarrow{\frac{\gamma}{\beta} \rightarrow 1} (\beta)(1 + \beta)^{-2\log_2(\beta)} = \frac{\beta}{(1 + \beta)^{2\log_2(\beta)}} \xrightarrow{\beta \rightarrow 0} 0$$

Luego podemos elegir β tan pequeño como para que $\gamma(1 + \beta)^{2m} \leq \epsilon$ y por lo tanto la transformación de $[0, 1]$ a $[0, 2]$ sea una ϵ -contracción como se precisa. Es sorprendente que si $\epsilon \leq 1$, entonces como $\epsilon^n < \epsilon$ y por lo tanto la transformación final contrae más que la original. Para demostrar la segunda parte del teorema, sean A y B tal que A está contenido en un intervalo I y como tiene interior no vacío, contiene un intervalo $[a, b]$. Sea $\epsilon > 0$. Por el teorema de Banach-Schröder-Bernstein basta con encontrar una ϵ -contracción de A a un subconjunto C de B y otra ϵ -contracción de A a un superconjunto D de B ($D \supset B$). La primera parte se puede acometer de una manera muy sencilla usando $\alpha x + n$ con α suficientemente pequeño como para que el intervalo I se transforme en C . Y para lo segundo podemos utilizar el procedimiento iterativo descrito anteriormente considerando $\alpha < \min\{1, (b - a)\epsilon\}$ de tal manera que consigamos generar una α -contracción de $[0, 1]$ a $D \supset B$. Por supuesto, en primer lugar se debe aplicar una $\frac{1}{b-a}$ -contracción para transformar $[a, b]$ en $[0, 1]$ y de este modo obtenemos una ϵ -contracción a trozos de $[a, b]$ a D conteniendo a B , como queríamos.

La paradoja de Von Neumann para \mathbb{R} , por lo tanto, nos hace comprobar la inexistencia de una medida finitamente aditiva definida en $\mathcal{P}(\mathbb{R})$ y normalizada al intervalo unidad que tenga la propiedad de encoger bajo contracciones.

Medida y descomposiciones paradójicas

En los capítulos anteriores, el método para encontrar la paradoja de Banach Tarski se basaba en la idea principal de encontrar una descomposición paradójica para el grupo de transformaciones primero y luego aplicando la inexistencia de puntos fijos no triviales, trasladar esa descomposición al conjunto sobre el que actuaba. Como vimos, la posibilidad de encontrar descomposiciones paradójicas para ciertos conjuntos está estrechamente ligada a la posibilidad de existencia de una medida finitamente aditiva en esos conjuntos. En este aspecto se profundizará en la primera parte de este capítulo. Por otro lado, fue Von Neumann quien se dió cuenta de que si se pudiera encontrar una medida finitamente aditiva invariante por la izquierda definida en todos los subconjuntos del grupo de transformaciones, G , ese hecho podría ser utilizado para generar una medida finitamente aditiva G -invariante definida en todos los subconjuntos de X , introduciendo la noción de grupo amenable. La segunda sección de este capítulo está dedicada al estudio particular de este tipo de grupos.

5.1. Teorema de Tarski

En esta sección hablaremos sobre la compatibilidad de la existencia de medidas finitamente aditivas con la existencia de conjuntos que se pueden descomponer de manera paradójica. Se finalizará la sección con el enunciado del teorema de Tarski que directamente expresa que ambos conceptos no son compatibles. Nos conviene introducir previamente varios conceptos como tipo de equidescomponibilidad que nos ayudarán en el tratamiento posterior.

Definición 5.1.1. Supongamos que el grupo G actúa sobre X . Definimos una acción extendida de la siguiente manera: Sea $X^* = X \times \mathbb{N}$ y sea $G^* = \{(g, \pi), g \in G \text{ siendo } \pi \text{ una permutación de } \mathbb{N}\}$ y sea el grupo G^* actuando sobre X^* del siguiente modo: $(g, \pi)(x, n) = (g(x), \pi(n))$. Si $A \subseteq X^*$, entonces aquellos elementos n cuyo A tiene al menos un elemento con segunda coordenada n son denominados niveles de A .

Vemos ahora que la G -equidescomponibilidad y la G^* -equidescomponibilidad son conceptos bastante

relacionados, por ejemplo, si $E_1, E_2 \subseteq X$, entonces $E_1 \sim_G E_2 \iff E_1 \times \{m\} \sim_{G^*} E_2 \times \{n\}$ para todo $m, n \in \mathbb{N}$. Intuitivamente, la información añadida que tenemos con esta nueva relación de equivalencia es el tipo de equidescomponibilidad de un conjunto. En este sentido si E es G -paradójico, la siguiente relación $E \times \{0\} \sim_{G^*} E \times \{0, 1\}$ es similar a decir $E = 2E$.

Definición 5.1.2. Sean G, X, G^*, X^* , entonces un subconjunto A de X^* se dirá acotado si tiene un número finito de niveles. La clase de equivalencia con respecto a G^* -equidescomponibilidad de un conjunto acotado $A \subseteq X^*$ se representará como $[A]$ y es denominada tipo de A . La colección de tipos de conjuntos acotados se representará como \mathcal{S} .

A estos efectos y con el objetivo de demostrar el teorema de Tarski trabajaremos en \mathcal{S} , que es el semigrupo de acciones de G sobre X . Como cualquier medida finitamente aditiva en $\mathcal{P}(X)$ es invariante bajo G -equidescomponibilidad, tal medida, μ , inducirá una medida ν de \mathcal{S} en $[0, \infty]$ tal que si $A \in \mathcal{S}$, $\nu([A]) = \sum(A_n)$ con $A = \bigcup A_n \times \{n\}$.

De esta manera Tarski utiliza una medida definida sobre un semigrupo para llegar al teorema que expondremos al final.

Proposición 5.1.3. Sea (\mathcal{F}, x) un semigrupo conmutativo dotado de elemento neutro, $(\mathcal{F}, 0)$. \mathcal{F} admite una ordenación definida por $\alpha \prec \beta$ con $\alpha, \beta \in \mathcal{F}$ si $\alpha + \delta = \beta$ para algún $\delta \in \mathcal{F}$. Además hay una multiplicación $\cdot : \mathcal{F} \rightarrow \mathcal{F}$ definida como $n \cdot \alpha$ ($n\alpha$) donde $n \in \mathbb{N}$.

Teniendo en cuenta que existe una ordenación para el semigrupo y una multiplicación, se pueden aplicar conceptos como los que aplicamos para conjuntos como los números naturales o enteros. Por ejemplo diremos que un semigrupo \mathcal{F} es acotado si fijado $\epsilon \in \mathcal{F}$ para algún $n \in \mathbb{N}$ se cumple $\alpha \prec n\epsilon$.

Teorema 5.1.4. (AE) Sea $(\mathcal{F}, +, 0)$ un semigrupo conmutativo y un elemento determinado del mismo $\epsilon \in \mathcal{F}$, Entonces las siguientes afirmaciones son equivalentes:

1. Para todo $n \in \mathbb{N}$, $(n+1)\epsilon \not\prec \epsilon$.
2. Existe una medida $\mu : \mathcal{F} \rightarrow [0, \infty)$ tal que se verifica $\mu(\epsilon) = 1$ y $\mu(\alpha + \beta) = \mu(\alpha) + \mu(\beta)$ para todo $\alpha, \beta \in \mathcal{F}$.

Demostración

(2) \implies (1): De la propiedad aditiva de la medida definida es sencillo deducir que $\mu(\alpha) \leq \mu(\beta)$ si $\alpha \prec \beta$ y que $\mu(n\epsilon) = n\mu(\epsilon)$, luego claramente se verifica $(n+1)\epsilon \not\prec \epsilon$

(1) \implies (2) no es tan evidente y precisaremos de una técnica que es bastante utilizada cuando se intenta construir medidas. Se basa en la compacidad del espacio producto $[0, \infty)^{\mathcal{F}}$ la cual se deduce trivialmente del conocido teorema de Tychonoff que afirma que el producto de cualquier familia de espacios compactos es compacto. Merece la pena mencionar que la demostración de este teorema también hace uso del Axioma de Elección.

Partiremos de la asunción de que todos los elementos de \mathcal{F} están acotados para ϵ , y extenderemos la medida definida para ellos a los elementos no acotados asignándoles medida ∞ .

Proposición 5.1.5. Si \mathcal{F}_0 es un subconjunto finito de \mathcal{F} el cual contiene a ϵ , entonces existe una función $\mu : \mathcal{F}_0 \rightarrow [0, \infty)$ que verifica:

1. $\mu(\epsilon) = 1$.
2. Para todo $\{\phi_i\}, \{\theta_j\} \in \mathcal{F}_0$, si $\phi_1 + \dots + \phi_m \leq \theta_1 + \dots + \theta_n$, entonces $\sum \mu(\phi_i) \leq \sum \mu(\theta_j)$

Demostración

La demostración se realizará por inducción sobre el tamaño de \mathcal{F}_0 . Si $|\mathcal{F}_0| = 1$ entonces claramente $\mathcal{F}_0 = \{\epsilon\}$ y $\mu(\epsilon) = 1$, en cuyo caso el apartado 2 de la proposición se reduce a demostrar que si $m\epsilon \leq n\epsilon$ entonces $m \leq n$. Pero esto se deduce del Teorema 5.1.4 1) porque si tuviéramos $m\epsilon \leq n\epsilon$ y $m \geq n + 1$ debería cumplirse $n\epsilon \geq m\epsilon \geq (n + 1)\epsilon$, lo cual es una contradicción. Supongamos que $|\mathcal{F}_0| > 1$ con $\alpha \in \mathcal{F}_0 \setminus \{\epsilon\}$. Utilizaremos la hipótesis de inducción para encontrar una función ν que sobre \mathcal{F}_0 satisfaga lo enunciado en la proposición. Como partimos de la hipótesis de acotación, todos los elementos deben estar acotados por algún $n\epsilon$, luego ν debe tomar solo finitos valores. Definamos μ dejando que ν coincida con μ en $\mathcal{F}_0 \setminus \alpha$ y para α asignamos $\mu(\alpha) = \inf(\frac{\sum \nu(\beta_k) - \sum \nu(\gamma_l)}{r})$ donde \inf es el ínfimo sobre el conjunto de todos los enteros positivos r y $\beta_1, \dots, \beta_l, \gamma_1, \dots, \gamma_p$ que satisfacen

$$\gamma_1 + \dots + \gamma_p + r\alpha \leq \beta_1 + \dots + \beta_l \tag{5.1}$$

. Como α está acotado, $\mu(\alpha)$ es el mayor límite inferior de un conjunto no vacío. Si $\alpha \leq n\epsilon$ entonces $\mu(\alpha) \leq n$. Se puede detectar la similitud con el clásico método de construcción de la medida exterior, μ debe ser lo más grande posible siempre dentro de los límites que establece la proposición para poder ser cumplida.

Una consecuencia de la propiedad mencionada en el apartado 2 de la proposición que estamos intentando demostrar es que $\mu(\alpha) \geq 0$. Para esto veamos el ejemplo $\epsilon \leq \epsilon + \alpha$, luego $1 \leq 1 + \mu(\alpha)$. Nos queda por demostrar que μ tal y como la hemos definido verifica el apartado 2) de la Proposición 5.1.5. Supongamos $\phi_1 + \dots + \phi_m + s\alpha \leq \theta_1 + \dots + \theta_n + t\alpha$ donde como antes $\phi_i, \theta_j \in \mathcal{F}_0 \setminus \alpha; s, t \in \mathbb{N}$. Si $s = t = 0$ entonces la desigualdad se verifica ya que ν cumple lo propuesto en $\mathcal{F}_0 \setminus \{\alpha\}$. Sea ahora el caso $s = 0$ y $t > 0$, luego debemos demostrar $\sum \nu(\phi_i) \leq t\mu(\alpha) + \sum \nu(\theta_j)$ o lo que es lo mismo si despejamos $\mu(\alpha)$, $\mu(\alpha) \geq w = \frac{\sum \nu(\phi_i) - \sum \nu(\theta_j)}{t}$. Bastará con demostrar que $\frac{\sum \nu(\beta_k) - \sum \nu(\gamma_l)}{r} \geq w$. Si multiplicamos $\phi_1 + \dots + \phi_m \leq \theta_1 + \dots + \theta_n + t\alpha$ por r y sumando a ambos lados $t\gamma_1 + \dots + t\gamma_p$ resulta:

$$r\phi_1 + \dots + r\phi_m + t\gamma_1 + \dots + t\gamma_p \leq r\theta_1 + \dots + r\theta_n + t\alpha + t\gamma_1 + \dots + t\gamma_p$$

Si sustituimos (5.1) entonces:

$$r\phi_1 + \dots + r\phi_m + t\gamma_1 + \dots + t\gamma_p \leq r\theta_1 + \dots + r\theta_n + t\beta_1 + \dots + t\beta_l$$

Luego la hipótesis de inducción sobre ν proporciona:

$$r \sum \nu(\phi_i) + t \sum \nu(\gamma_l) \leq r \sum \nu(\theta_j) + t \sum \nu(\beta_k)$$

lo cual implica $\frac{\sum \nu(\beta_k) - \sum \nu(\gamma_l)}{r} \geq w$ como queríamos. Nos queda el caso en que

$$\phi_1 + \cdots + \phi_m + s\alpha \leq \theta_1 + \cdots + \theta_n + t\alpha \quad (5.2)$$

donde $s > 0$. En este caso nos es suficiente con demostrar que:

$$\sum \nu(\phi_i) + s\mu(\alpha) \leq x_1 + \dots + x_t + \sum \nu(\theta_j)$$

siendo x_1, \dots, x_t cualquiera de los números cuyo mayor límite inferior define $\mu(\alpha)$. Si elegimos el más pequeño de x_1, \dots, x_t podemos asumir que son todos el mismo. Es decir supongamos que:

$$\gamma_1 + \cdots + t\gamma_p + r\alpha \leq \beta_1 + \cdots + \beta_l$$

con $\gamma_l, \beta_k \in \mathcal{F}_0 \setminus \alpha$ y sea $x = \frac{\sum \nu(\beta_k) - \sum \nu(\gamma_l)}{r}$. Tenemos que demostrar: $s\mu(\alpha) + \sum \nu(\phi_i) \leq tx + \sum \nu(\theta_j)$. Multiplicando (5.2) por r y sumando a ambos lados $t\gamma_1 + \cdots + t\gamma_p$ resulta:

$$r\phi_1 + \cdots + r\phi_m + t\gamma_1 + \cdots + t\gamma_p \leq r\theta_1 + \cdots + r\theta_n + rt\alpha + t\gamma_1 + \cdots + t\gamma_p$$

Y utilizando (5.1):

$$r\phi_1 + \cdots + r\phi_m + t\gamma_1 + \cdots + t\gamma_p + rs\alpha \leq r\theta_1 + \cdots + r\theta_n + t\beta_1 + \cdots + t\beta_l$$

siendo ésta última desigualdad similar a (5.1), que es la utilizada para definir $\mu(\alpha)$ y por lo tanto $s\mu(\alpha) + \sum \nu(\phi_i)$ está acotado por :

$$\sum \nu(\phi_i) + s\frac{1}{rs}(r \sum \nu(\theta_j) + t \sum \nu(\beta_k) - r \sum \nu(\phi_i) - t \sum \nu(\gamma_l))$$

lo que es igual que $tx + \sum \nu(\theta_j)$, como queríamos.

Después de haber demostrado 5.1.5 veremos cómo podemos utilizar la idea de compacidad para producir la medida deseada en \mathcal{F} . Para cualquier \mathcal{F}_0 conforme a la Proposición 5.1.5 definamos $\mathcal{M}(\mathcal{F}_0)$ como el conjunto $f \in [0, \infty]^{\mathcal{F}}$ que satisfacen $f(\epsilon) = 1$ y $f(\alpha + \beta) = f(\alpha) + f(\beta)$ con $\alpha, \beta \in \mathcal{F}_0$. La propiedad de aditividad es una consecuencia directa de 5.1.5 2), luego se deduce que cada $\mathcal{M}(\mathcal{F}_0)$ es no vacío. Por otra parte para ver la compacidad del espacio producto $[0, \infty]^{\mathcal{F}}$ razonemos de la siguiente manera: Si una colección de subconjuntos cerrados de $[0, \infty]^{\mathcal{F}}$ tiene la propiedad de intersección finita (cualquier intersección de un número finito de conjuntos es no vacía), entonces la intersección de todos los conjuntos en la colección es no vacía. Veamos también que $\mathcal{M}(\mathcal{F}_0)$ es cerrado. Esto es así porque $f \in \mathcal{M}(\mathcal{F}_0)$ depende únicamente de un número finito de coordenadas en $f \upharpoonright_{\mathcal{F}_0}$ y si f no cumple (1) o (2), entonces $f(\epsilon) \neq 1$ o $f(\alpha + \beta) \neq f(\alpha) + f(\beta)$. En cualquiera de los dos casos habría un subconjunto abierto de $[0, \infty]^{\mathcal{F}}$ que contiene a $f \upharpoonright_{\mathcal{F}_0}$ en el cual la condición no es satisfecha. Luego $\mathcal{M}(\mathcal{F}_0)$ es cerrado. Veamos también que $\mathcal{M}(\mathcal{F}_1) \cap \cdots \cap \mathcal{M}(\mathcal{F}_n) \supseteq \mathcal{M}(\mathcal{F}_1 \cup \cdots \cup \mathcal{F}_n)$ donde cada \mathcal{F}_i es un subconjunto finito de $\mathcal{F} \supset \{\epsilon\}$. Como $\bigcup \mathcal{F}_i$ es finito, por la Proposición 5.1.5 entonces $\mathcal{M}(\bigcup \mathcal{F}_i) \neq \emptyset$. Luego $\{\mathcal{M}(\mathcal{F}_0) : \mathcal{F}_0\}$ es un subconjunto finito de $\mathcal{F} \supset \{\epsilon\}$ y que tiene la propiedad de intersección finita. Por compacidad debe existir $\mu \in \mathcal{M}(\mathcal{F}_0)$ siendo μ una medida con las propiedades deseadas, esto es: Como $\mu \in \mathcal{M}(\{\epsilon\})$ $\mu(\epsilon) = 1$ y la propiedad aditiva se deduce del hecho que $\mu \in \mathcal{M}(\{\epsilon, \alpha, \beta, \alpha + \beta\})$.

Teorema 5.1.6. (Teorema de Tarski)(AE) Sea G actuando sobre X y $E \subseteq X$. Entonces existe una medida finitamente aditiva G -invariante $\mu : \mathcal{P}(X) \rightarrow [0, \infty)$ tal que $\mu(\epsilon) = 1 \iff E$ no es G -paradójico.

Demostración

Partamos de la hipótesis de que G no es paradójico y sea \mathcal{S} el semigrupo de acciones de G sobre X . Por ser G no paradójico, ningún elemento de G , por ejemplo $\epsilon = E \times \{0\}$ es equidescomponible con varias copias de sí mismo, es decir $\epsilon \not\sim E \times X$, donde X esta formado por al menos dos elementos en \mathbb{N} . En particular $n\epsilon \not\sim 2n\epsilon, \forall n \in \mathbb{N}$ y para $n = 1$ $\epsilon \not\sim 2\epsilon \implies \epsilon \prec 2\epsilon$. Por otro lado si se verificara $(n+1)\epsilon \prec \epsilon$ entonces tendríamos:

$$(n+1)\epsilon \prec n\epsilon \implies n\epsilon + \epsilon \prec n\epsilon \implies (n+1)\epsilon + \epsilon \prec n\epsilon \implies \dots \implies n\epsilon + n\epsilon \prec n\epsilon \implies 2n\epsilon \prec n\epsilon \implies 2\epsilon \prec \epsilon$$

Luego entonces $\epsilon \sim 2\epsilon$ lo cual es una contradicción. Por lo tanto $(n+1)\epsilon \not\prec \epsilon \forall n \in \mathbb{N}$ verificándose el Teorema 5.1.4 1. Por el mismo teorema, entonces, existe una medida ν en \mathcal{S} normalizando ϵ y tal que $\mu(A) = \nu([A \times \{0\}])$ es la deseada medida G -invariante en $\mathcal{P}(X)$.

5.2. Grupos amenables

En esta sección se introduce la definición de grupo amenable, que es aquél en el que se puede definir una medida con las características ya estudiadas y por lo tanto es un concepto en cierta manera opuesto al del grupo paradójico. Veremos también ciertas propiedades que los hacen especialmente interesantes.

Definición 5.2.1. Amenabilidad. Si, para un grupo G , μ es una medida finitamente aditiva en $\mathcal{P}(G)$ siendo $\mu(G)=1$ e invariante por la izquierda ($\mu(g(A)) = \mu(A)$ para $g \in G$ y $A \subseteq G$), entonces μ es una medida en G . Un grupo amenable (también denominado promediable) es un grupo que está dotado de tal medida. Se designa como AG a la clase de todos los grupos amenables.

Proposición 5.2.2. *Si G es paradójico con respecto a la traslación por la izquierda, entonces G no es amenable.*

Demostración

Sea G un grupo amenable, por lo que existe una medida finitamente aditiva sobre G con las características expresadas en la definición. Si G es G -paradójico entonces existen $A_1, \dots, A_m, B_1, \dots, B_n \in G$, y $g_1, \dots, g_m, h_1, \dots, h_n \in G$ tal que:

$$G = \bigcup_{i=1}^m g_i(A_i), G = \bigcup_{j=1}^n h_j(B_j)$$

Teniendo en cuenta lo anterior:

$$\mu(G) \geq \mu(A_1 \cup A_2 \cup \dots \cup A_m \cup B_1 \cup B_2 \cup \dots \cup B_n)$$

Por ser μ finitamente aditiva:

$$\mu(A_1 \cup A_2 \cup \dots \cup A_m \cup B_1 \cup \dots \cup B_n) = \mu(A_1) + \mu(A_2) + \dots + \mu(A_m) + \mu(B_1) + \mu(B_2) + \dots + \mu(B_n)$$

y por ser G amenable:

$$\begin{aligned} & \mu(A_1) + \mu(A_2) + \dots + \mu(A_m) + \mu(B_1) + \mu(B_2) + \dots + \mu(B_n) = \\ & \mu(g_1(A_1)) + \mu(g_2(A_2)) + \dots + \mu(g_m(A_m)) + \mu(h_1(B_1)) + \mu(h_2(B_2)) + \dots + \mu(h_n(B_n)) \end{aligned}$$

Aplicando otra vez la finita aditividad de μ , llegamos a la conclusión:

$$\mu(G) \geq \mu\left(\bigcup_{i=1}^m g_i(A_i)\right) \cup \left(\bigcup_{j=1}^n h_j(B_j)\right) = \mu(2G) = 2\mu(G)$$

lo cual está en contradicción con la definición de la medida que estamos utilizando.

Por lo tanto todo grupo que tenga un subgrupo libre de rango 2 no es amenable.

A continuación veremos que si consideramos un grupo amenable G y definimos $B(G)$ como el conjunto de todas las funciones acotadas reales actuando sobre G , como existe una medida bien definida sobre G podremos aplicar la construcción de la integral en base a una medida sobre el espacio de medida $(G, \mathcal{P}(G), \mu)$. El hecho de que todos los conjuntos de G sean medibles implica que todas las funciones son medibles y por lo tanto $\int f d\mu$ define un funcional lineal en $B(G)$, con las propiedades ya conocidas:

1. $\int af + bg d\mu = a \int f d\mu + \int g d\mu \quad \forall a, b \in \mathbb{R}$.
2. $\int f d\mu \geq 0$ si para todo $g \in G$, $f(g) \geq 0$.
3. $\int \chi_G d\mu = 1$ donde χ_G es la función característica de G , es decir:

$$\chi_G(g) = \begin{cases} 0 & \text{si } g \text{ no pertenece a } G \\ 1 & \text{si } g \text{ pertenece a } G \end{cases}$$

4. Invariancia por la izquierda: sea $g \in G, f \in B(G)$, entonces $\int_g f d\mu = \int f d\mu$ donde $(_g f)(h) = f(g^{-1}h)$.

A todo funcional lineal en $B(G)$ que satisface las anteriores propiedades se le denomina media invariante por la izquierda, por lo que en un grupo amenable siempre se puede definir un funcional de este tipo.

Podemos definir la propiedad de ser invariante por la derecha de un modo similar. Si μ es invariante por la izquierda podemos definir una medida μ_0 invariante por la derecha de la siguiente manera: $\mu_0(A) = \mu(A^{-1})$ y por lo tanto la integral asociada es invariante por acción derecha de G en $B(G)$ dada por $f_g(h) = f(hg^{-1})$.

Proposición 5.2.3. *Si G es amenable entonces hay una medida ν en \mathcal{P} que es invariante por la derecha y por la izquierda.*

Demostración

Sea μ una medida definida sobre G y μ_0 su correspondiente medida invariante por la derecha asociada. Si $A \subseteq G$, definamos f_A en $B(G)$ por $f_A = \mu_0(Ag^{-1})$, tal que f_A está acotada por 1. Definamos el funcional ν en $\mathcal{P}(G)$ como $\nu(A) = \int f d\mu_0$. Es trivial ver que $\nu(G) = 1$ pues $f_G(g)=1$ y que es finitamente aditiva pues μ_0 lo es. Además puesto que el grupo es invariante por conjugación y μ_0 es invariante por la derecha, $f_{gA} = f_A$ y $f_{Ag} = (f_A)_g$ luego ν es invariante por la derecha y por la izquierda.

Proposición 5.2.4. *Todo grupo finito es amenable.*

Demostración

Sea G grupo finito de orden n y definimos la medida μ sobre subconjuntos A de G de la siguiente manera: $\mu(A) = \frac{|A|}{n}$. Es trivial ver que $\mu(G) = 1$. Por otro lado también es finitamente aditiva. Para $\{A_i\}_1^{n'}$ $\subset G, n' < n$ disjuntos:

$$\mu\left(\bigcup_{i=1}^{n'} A_i\right) = \frac{\left|\bigcup_{i=1}^{n'} A_i\right|}{n} = \frac{|A_1|+|A_2|+\dots+|A_{n'}|}{n} = \mu(A_1) + \mu(A_2) + \dots + \mu(A_{n'}) = \sum_{i=1}^{n'} \mu(A_i)$$

Y también es invariante por la izquierda: $\mu(g(A)) = \frac{gA}{n} = \frac{A}{n} = \mu(A)$ con $g \in G$ y $A \subset G$

Por tanto cumple con todas las condiciones para ser un grupo amenable.

Proposición 5.2.5. *Todo subgrupo de un grupo amenable es amenable.*

Demostración

Sea μ una medida definida sobre G , grupo amenable, con las características enunciadas en la Definición 5.2.1 y sea H un subgrupo de G . Sea M un conjunto de representantes de las clases de equivalencia de las clases laterales derechas de H en G . Definamos $\nu : \mathcal{P}(H) \rightarrow [0, 1]$ de la siguiente manera:

$$\nu(A) = \mu(\bigcup\{Ag : g \in M\}) \text{ para } A \subset H$$

Trivialmente $\nu(H) = \nu(G) = 1$ y ν es finito-aditiva pues μ lo es y la disyunción de subconjuntos $A_i \in H$ es trasladada a $A_i g$, ya que el subgrupo actúa libremente sobre el grupo en el que está contenido. Para comprobar que es invariante por la izquierda veamos que si $g, g' \in G$:

$$\nu(g'A) = \mu(\bigcup\{(g'A)g : g \in M\}) = \mu(\bigcup\{g'(Ag) : g \in M\}) = \nu(A) \text{ para } A \subset H$$

donde hemos aplicado la propiedad asociativa por ser G un grupo y el hecho de que μ es invariante por la izquierda.

Proposición 5.2.6. (AE)

Si N es un subgrupo normal de un grupo amenable G , entonces G/N es amenable.

Demostración

Sea μ una medida con las características enunciadas en la Definición 5.2.1 . Como en las demostraciones anteriores construiremos una medida $\nu : \mathcal{P}(G/N) \rightarrow [0, 1]$ a partir de μ de la siguiente manera:

$$\nu(A) = \mu(\bigcup A) \text{ para } A \subset G/N$$

Como N es normal se tiene que si $n \in N$ y $g \in G$ entonces $ng = gn \in G/N$, por lo cual:

$$\nu(G/N) = \mu(\bigcup g/N) = \mu(\bigcup \{ng : g \in G\}) = \mu(\bigcup \{gn : g \in G, n \in N\}) = \mu(N) = 1$$

Además ν es finito-aditiva puesto que μ lo es. Y para comprobar la invariancia por la izquierda veamos que si $A \subset G/N$ y G' es un subconjunto de G tal que $G'N = A$, entonces:

$$\nu(gA) = \mu(\bigcup \{gNg' : g' \in G'\}) = \mu(\bigcup \{Ng' : g' \in G'\}) = \nu(A)$$

Proposición 5.2.7. (AE)

Si N es un subgrupo normal de un grupo G y $N \subset G$ y G/N son amenables, entonces G es amenable.

Demostración

Partimos de N amenable, subgrupo normal de G con grupo cociente G/N también amenable. Por lo tanto existen dos medidas ν_1, ν_2 definidas en N y G/N respectivamente con las propiedades enunciadas en 5.2.1. Para cualquier $A \in G$ definamos $f_A : G \rightarrow \mathbb{R}$ como:

$$f_A(g) = \nu_1(N \cap g^{-1}A)$$

Entonces si g_1 y g_2 definen la misma clase lateral izquierda de N en G , se tiene $g_1^{-1}g_2 = n \in N$, luego

$$f_A(g_1) = \nu_1(N \cap g_1^{-1}A) = \nu_1(N \cap ng_2^{-1}A) = \nu_1(n(N \cap g_2^{-1}A)) = \nu_1(N \cap g_2^{-1}A) = f_A(g_2)$$

,esto es, la restricción de f_A a G/N , es una función real acotada con dominio G/N . Como partimos de la suposición de que G/N es amenable, entonces se puede definir un funcional lineal que es media invariante por la izquierda, como vimos al principio de la sección, definiendo $\mu : \mathcal{P}(G/N) \rightarrow \mathbb{R}$ como $\mu(A) = \int f_A d\nu_2$. Veamos que la medida así definida cumple con las propiedades pertinentes. En primer lugar, veamos que $f_g(G) = 1$ siendo $g \in G$ y por ser $G/N \subseteq G$, $f_g(gN) = 1$ siendo $gN \in G/N$. Con el objetivo de demostrar la finito-aditividad, es sencillo ver que si $A, B \in G$, y $A \cap B = \emptyset$, entonces $g^{-1}A \cap g^{-1}B = \emptyset$ y por lo tanto como ν_1 también es finitamente aditiva, $f_{A \cup B}(g) = f_A(g) \cup f_B(g)$, siendo así μ finitamente aditiva. Finalmente, para verificar que μ es invariante por la izquierda, basta tener en cuenta que $f_{gA}(g_0) = \nu_1(N \cap (g_0)^{-1}gA) = \nu_1(N \cap (g_0^{-1}g)A) = f_A(g_0^{-1}g) = (f_A)(g_0)$ y de aquí se deduce la invariancia por la izquierda de la integral definida por $\int f_A d\nu_2$, luego μ también es invariante por la izquierda, y por lo tanto G es amenable.

Proposición 5.2.8. (AE)

Todo grupo abeliano es amenable.

Hay varias demostraciones de la Proposición 5.2.8. La más conocida hace uso del Teorema de compacidad de Tychonoff el cual utiliza el Axioma de Elección. No la expondremos aquí pudiendo consultarse por ejemplo en [9].

Grupos amenables en \mathbb{R} y \mathbb{R}^2

Una orientación diferente a la hora de estudiar la existencia de descomposiciones paradójicas en \mathbb{R} y \mathbb{R}^2 se basa en la utilización de los grupos amenables definidos anteriormente.

Sabemos que en \mathbb{R} el grupo de isometrías de \mathbb{R} , G_1 , coincide con el grupo de traslaciones en la recta real el cual es un grupo abeliano y por lo tanto, por la Proposición 5.2.8, amenable. Luego no existen descomposiciones paradójicas en \mathbb{R} considerando este tipo de transformaciones, pero sí existe una medida finitamente aditiva G_1 -invariante por la izquierda definida sobre $\mathcal{P}(\mathbb{R})$.

En \mathbb{R}^2 el grupo de isometrías, G_2 , no es abeliano, sin embargo el subgrupo $SO(2) \in G_2$ es subgrupo normal y sí lo es y también G_2/S_2 . Luego en virtud de la proposición 5.2.7 G_2 es amenable y en él es posible la definición de una medida con las características enunciadas en la Definición 5.2.1.

Comentarios finales

En este trabajo se ha pretendido presentar y explicar con cierta profundidad uno de los resultados más sorprendentes producidos en la historia de las matemáticas: la paradoja de Banach-Tarski. La versión popular e informal de la misma se le conoce como paradoja del guisante y el Sol, y dice que podemos trocear un guisante en un número finito de trozos, transformarlos sólo mediante traslaciones y rotaciones y unir los trozos transformados para formar una bola del tamaño del Sol. Esta versión recoge la idea esencial detrás de la paradoja, sin embargo como hemos visto a lo largo del trabajo, no sería posible llevarlo a la práctica puesto que tendríamos que partir la bola inicial (guisante) en trozos tan extraordinariamente complejos que físicamente sería imposible. A lo largo de todo el trabajo se ha hecho hincapié en la idea de que la paradoja es posible por ciertas características intrínsecas del grupo de transformaciones en relación al conjunto sobre el que actúa. Como vimos al principio del capítulo 3, dos características especiales de este grupo: el poder encontrar transformaciones que no conmutan y el que cada transformación tenga un conjunto limitado de puntos fijos posibilitan que la esfera tridimensional pueda ser descompuesta paradójicamente. En el capítulo 4 al estudiar la paradoja en otras dimensiones, vimos que como estas dos características se mantienen para las isometrías en dimensión mayor que 3, la paradoja también es posible para esas dimensiones, sin embargo para dimensiones 1 y 2 no es posible ya que las transformaciones consideradas, en esos espacios, no conmutan. Introduciendo un grupo más amplio (menos restrictivo) de transformaciones finalmente conseguimos encontrar descomposiciones paradójicas tanto en la recta como en el plano. Finalmente en el capítulo 5 estudiamos la relación entre el concepto de medida y el de descomposición paradójica. Se expone el Teorema de Tarski en el cual se plasma la imposibilidad de definir una medida finitamente aditiva en un conjunto paradójico. Se acaba el capítulo introduciendo la noción de grupos amenables que son aquéllos en los que sí se puede definir una medida finitamente aditiva y por lo tanto son grupos que no se pueden descomponer paradójicamente. Por otro lado, aparte de todo lo anterior se ha visto que el Axioma de Elección interviene en la demostración de los teoremas principales relacionados con la paradoja, y se ha mencionado que su utilización todavía a día de hoy no deja de ser polémica, pues suele conducir a resultados en contra de la intuición, y que por lo tanto pudieran ser discutibles. En el apéndice A se hablará con más detalle de este Axioma y sus implicaciones. Con el

objetivo de que la extensión del trabajo fuera la adecuada y que sólo sirviera como introductorio, fuera del tratamiento que se ha realizado, quedan muchos desarrollos que se han hecho y se continúan haciendo en la actualidad en relación a esta paradoja. Basta citar el estudio de descomposiciones paradójicas en conjuntos con ciertas propiedades, verbigracia, conjuntos que tienen la propiedad de Baire o con otro tipo de geometrías como la hiperbólica. Por último, merece la pena mencionar que la motivación de estudiar en profundidad la paradoja y sus implicaciones ha propiciado indirectamente la consecución de resultados en diversas áreas de la matemática que a priori no estaban relacionadas de un modo claro con ella.

El papel del Axioma de Elección

Como mencionamos en la introducción de este trabajo, desde su descubrimiento, la Paradoja de Banach-Tarski ha reforzado la visión crítica que muchos matemáticos tienen del Axioma de Elección. De hecho tan pronto como la Paradoja de Hausdorff fue propuesta, Emile Borel fue uno de los matemáticos que se posicionó claramente en contra de ésta por el uso del Axioma de Elección, sin embargo previamente el propio Borel inconscientemente lo había utilizado en la demostración del resultado obtenido por Cantor de que todo conjunto infinito contiene un conjunto numerable.

Los mismos Banach y Tarski ya anticiparon cuando la expusieron, que en su demostración habían utilizado este axioma, y que este aspecto debía ser tenido en cierta consideración. Para poder entender lo que vamos a explicar a continuación, introduciremos cierta notación de teoría de conjuntos. Si T es una colección de sentencias en el lenguaje de teoría de conjuntos, por ejemplo $T=ZF$ (Zermelo-Fraenkel) o $T=AE$, entonces $Con(T)$ es la aserción de que T es consistente. Por ejemplo, Gödel demostró en 1938 que $Con(ZF)$ implica $Con(ZF+AE)$, es decir, el Axioma de Elección no contradice lo propuesto en el conjunto axiomático de Zermelo-Fraenkel.

Desde que se propuso la existencia de conjuntos no medibles, muchos matemáticos se preguntaron si había alguna manera de construir un conjunto de este tipo sin utilizar el Axioma de Elección. En 1964 Robert M. Solovay, usando el método de forzado propuesto por Paul Cohen un año antes, respondió de manera negativa a la pregunta anterior obteniendo $Con(ZF+LM)$ ¹ es consistente. De aquí se deduce que la no veracidad de la Paradoja de Banach-Tarski es consistente con ZF y por lo tanto la Paradoja de Banach-Tarski no se puede deducir de ZF . Por otro lado la no utilización del Axioma de Elección también puede llevar a resultados sorprendentes cuando se considera la medida de Lebesgue junto con ZF . De hecho la afirmación de que \mathbb{R} es unión numerable de conjuntos numerables es consistente con $ZF+LM$. Sin embargo, si esto fuera cierto, la medida Lebesgue no podría ser numerablemente aditiva. Con lo cual parece que debemos permitir cierta parte de Axioma de Elección para que la medida se comporte de la manera adecuada pero no tanto como para que los números reales puedan ser bien ordenados o que

¹LM es la sentencia que dice que todos los subconjuntos de \mathbb{R}^n son medibles Lebesgue

un conjunto no medible Lebesgue pueda ser construido. En este sentido surge otro Axioma: El Axioma de Elección Dependiente (DE), el cual garantiza la existencia de un conjunto de elección numerable. Es decir restringe el Axioma de Elección a colecciones numerables de conjuntos. La introducción de este Axioma representa una solución intermedia entre las opciones extremas de utilizar AE o no. Sin embargo se puede demostrar también que la paradoja de Banach no se puede deducir de $ZF+DE$. Otro punto particular interesante es que la Paradoja de Banach no es completamente equivalente a AE ya que esta última requiere que todo conjunto pueda ser bien ordenado mientras que la primera sólo requiere que lo sea \mathbb{R} . Hay muchos más desarrollos al respecto, pero no establecen de forma clara y definitiva si es correcto utilizar el Axioma de Elección o no. A día de hoy son mayoritarios los matemáticos que aceptan el uso del Axioma de Elección. Es generalmente asumido que conjuntos no medibles conducen a situaciones curiosas que contradicen la realidad del mundo físico, pero el tipo de matemáticas que se suele emplear para describir problemas físicos suelen desarrollarse sobre conjuntos que son medibles. Además como punto a favor de AE, desde el punto de vista de teoría de conjuntos, se puede demostrar que $ZF+AE$ establece unos principios fundacionales más coherentes que ZF sólo.

Bibliografía

- [1] A. Beshenov. *Curso de Álgebra*, El Salvador. 2018.
- [2] L. Cohn. *Measure Theory*. Birkhäuser, Boston. 1980.
- [3] P. Deligne, D. Sullivan *Division Algebras and The Hausdorff-Banach-Tarski Paradox*. End. Math 29. 1983.
- [4] F. Hausdorff *Grundzuge der Mengenlehre*. Chelsea Publishing Company reprint(1949). 1914.
- [5] K. Juschenko. *Amenability of discrete groups by examples*. 2015.
- [6] Hans Kurzweil, Bernd Stellmacher. *Theory of finite groups*. Springer. 2004.
- [7] M. Laczkovich. *Von Neumann's paradox with translations*. Fundamenta Mathematicae, vol 131, 1. 1988.
- [8] Eli Maor. *The infinity and beyond*. Birkhäuser, Boston. 1987.
- [9] M. Vicaria. *Definable groups in Models of Presburger Arithmetic and $G00$* . Msc thesis. 2016.
- [10] John Von Neumann. *Zur Allgemein theorie des masses*. Fundamenta Mathematicae 13.1, 1929.
- [11] Stan Wagon. *The Banach-Tarski paradox*. Cambridge University Press. 2016.