00324

# UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

2

FACULTAD DE CIENCIAS

## AXIOMA DE ELECCION Y TEORIA DE LA MEDIDA.

# T E S I S

QUE PARA OBTENER EL TITULO DE:

**M A T E M A T I C A**

P R E S E N T A :

**A N A  A L V A R E Z  V E L A S C O**

DIRECTOR DE TESIS: DRA. ANA MEDA GUARDIOLA
DR. JOSE ALFREDO AMOR MONTAÑO



FACULTAD DE CIENCIAS
UNAM



2003



1

PAGINACION DISCONTINUA

VNIVERSDAD NACIONAL
AVPNºMA DE
MEXICO

**DRA. MARÍA DE LOURDES ESTEVA PERALTA**
**Jefa de la División de Estudios Profesionales de la**
**Facultad de Ciencias**
**Presente**

Comunicamos a usted que hemos revisado el trabajo escrito:

Axioma de Elección y Teoría de la Medida

realizado por Ana Álvarez Velasco

con número de cuenta 09320386-3 , quién cubrió los créditos de la carrera de Matemáticas

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis
Propietario Dra. Ana Meda Guardiola

Director de Tesis
Propietario Dr. José Alfredo Amor Montaño

Propietario Dr. Guillermo Grabinsky Steider

Suplente Dra. Magali Folch Gabayet

Suplente M. en C. Alejandro Bravo Mojica

**Consejo Departamental de** Matemát

M. en C. JOSE ANTONIO GOMEZ ORTEGA

FACULTAD DE CIENCIAS

MATEMATICAS

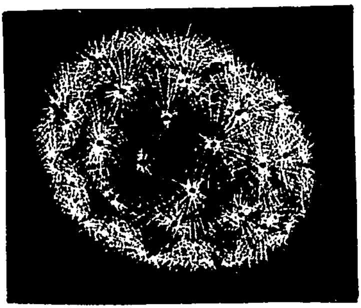
TESIS CON
FALLA DE ORIGEN

1-a

# AXIOMA DE ELECCIÓN
# Y
# TEORIA DE LA MEDIDA

Ana Álvarez Velasco

Septiembre, 2003





1. b

# Puntos de Fuga

Il y avait en effet environ cinq mille hommes.
Il a dit à ses disciples: « Faites-les s'installer
par groupes d'une cinquantaine». Ils firent
ainsi et les installèrent tous. Jésus prit les
cinq pains et les deux poissons et levant son
regard vers le ciel, il prononça sur eux la
bénédiction, les rompit, et il les donnait aux
disciples pour les offrir à la foule.
Ils mangèrent et furent tous rassasiés, et ce
qu'ils avaient eu de reste fut emporté:
douze couffins de morceaux.

Saint Luc 9:14-17.



*A quien me demostró que el vacío existe.*

# Confesiones Capitales

La realización de esta tesis ha sido un ejercicio interdisciplinario dentro de las matemáticas que ha involucrado el trabajo de muchas personas. No sólo de quienes introdujeron por primera vez los resultados que aquí se presentan (Solovay, Banach y Ulam, entre otros) o de quienes los sintetizaron con especial maestría (Oxtoby, Kunen y Wagon), sino también de quienes tuvieron la disposición y el entusiasmo para tender puentes entre el área de su especialidad y otras ramas de las matemáticas. Me refiero, en particular, a la Dra. Ana Meda Guardiola (Análisis Matemático y Probabilidad) y al Dr. José Alfredo Amor Montaño (Teoría de Conjuntos y Lógica Matemática), a quienes se debe la concreción de este trabajo. Además ha sido un privilegio tener como interlocutores al Dr. Guillermo Grabinsky, a la Dra. Magali Folch, al M.en. C. Alejando Bravo y a los miembros del seminario de Forcing a cargo del Dr. José Alfredo Amor (Dr. Alejandro Odgers y Mat. Miguel Angel Mota). El conjunto de cualidades de todos ellos no es medible y, como los de Vitali, hace que el mundo de las matemáticas sea fascinante.

# Índice general

# Introducción

En 1970 Robert Solovay publicó un artículo bajo el título "Un modelo de la teoría de conjuntos en el que todos los conjuntos de reales son Lebesgue medibles": contraviniendo, al menos en apariencia, uno de los principales resultados de teoría de la medida. Recordemos que en 1905 Giuseppe Vitali había demostrado la existencia de conjuntos de reales que no son Lebesgue medibles mediante el axioma de elección. Ya desde entonces el propio Lebesgue, que creía haber encontrado una forma de medir el tamaño geométrico de todos los conjuntos de reales, veía con recelo el uso del axioma de elección para la construcción de conjuntos. Sin embargo, aunque el ejemplo de Vitali hacía un claro uso de dicho axioma, no había la certeza de que fuera condición necesaria para la existencia de conjuntos no Lebesgue medibles, ya que las mismas propiedades de la medida podían ser un factor determinante.

El interés por identificar qué era exactamente lo que deshacía la ilusión de Lebesgue derivó en la formalización y depuración de los vínculos entre la teoría de la medida y la teoría de conjuntos, que incluyen puntos de encuentro con la teoría de grupos y la topología. El presente trabajo quisiera dar fe de esa compleja relación pasional que tan hermosos resultados, como la paradoja de Banach-Tarski, ha dado.

La tesis se ha dividido en dos partes. La primera mucho más cercana a la teoría de la medida y el análisis matemático y la segunda a la teoría de conjuntos y la lógica matemática. Sin embargo, puesto que uno de los objetivos de la tesis es justamente mostrar que esos límites son difusos, sería más adecuado decir que los resultados de la primera parte demuestran la relevancia, para la teoría de la medida, del teorema de Solovay, cuya demostración se basa en una compleja red de resultados de la lógica que se presenta en la segunda parte.

La explicación sobre cómo es posible tener un modelo de la teoría de conjuntos en el que todos los conjuntos de reales son Lebesgue medibles se deja para el último capítulo de la tesis (Capítulo 6). En él se expone el método de Forcing que introdujo Cohen en 1963 para demostrar, entre otras cosas, que la negación del axioma de elección es consistente con el resto de los axiomas de la teoría de conjuntos.

Por otro lado, a pesar de su título, el artículo de Solovay no sólo se refiere a teoría de la medida, sino que incluye resultados interesantes respecto a la topología y cardinalidad de los conjuntos de reales. En el Capítulo 5 se analiza el significado y relevancia de éstos. En particular, se discute la propiedad de Baire y la propiedad del conjunto perfecto. Asimismo se ve que en el modelo de Solovay son válidos algunos resultados fundamentales del análisis matemático como el teorema de Heine-Borel y algunas hipótesis como la del continuo, que se suelen aceptar en la práctica matemática cotidiana. Sin embargo, para entender de manera más profunda el valor del trabajo de Solovay para el análisis matemático es necesario presentar los problemas que acompañaron el nacimiento de la teoría de la medida, lo que establece el puente con la primera parte de la tesis.

El problema de cómo generalizar la noción de longitud es el punto de partida de esta tesis (Capítulo 1) y aunque este trabajo está muy lejos de ser un recuento histórico de dicho problema, sí podemos decir que uno de sus objetivos es mostrar que el lazo que ata los resultados que presentamos es el intento por resolver el problema de la medida. La solución sugerida por Lebesgue es un eslabón central en esta cadena, no sólo porque a ella se refiere el artículo de Solovay, sino porque es la única forma de medir adecuadamente los conjuntos borelianos. Es por ello que en el Capítulo 3 se demuestra la unicidad de la medida de Lebesgue y se incluye una construcción que permite entender por qué dicha medida extiende la noción de longitud.

Por otro lado, en los Capítulos 1, 2 y 3 se presentan tres conjuntos no medibles (Vitali, Banach-Tarski y Bernstein) construidos sobre la base de un buen orden para los reales. En el Capítulo 2 se construye la medida de Banach, que junto los conjuntos anteriores permite ver que ninguna de las propiedades que se le piden a una medida en el contexto euclideano son condiciones necesarias para tener conjuntos no medibles.

Los teoremas de Ulam, que se presentan en el Capítulo 4, establecen un vínculo entre la teoría de conjuntos y la teoría de la medida que, a diferencia de los resultados de los capítulos anteriores, va más allá del axioma de elección y pone en la mesa de discusión la relación entre cardinales inaccesibles y conjuntos no medibles. Lo interesante es que dicha discusión es punto de partida del trabajo de Solovay y, por lo tanto, un nuevo enlace entre la segunda y la primera parte de esta tesis.

Finalmente, cabe señalar que aunque los resultados de la primera parte de la tesis se incluyeron por su vínculo con la existencia de conjuntos no medibles, tienen valor teórico propio. Algunos porque nos enfrentan a paradojas que requieren ser aclaradas, como lo es el caso de la descomposición rígida de una esfera en dos esferas del mismo tamaño, pero otros simplemente porque nos muestran la aplicación de algunos resultados importantes del análisis matemático, como el uso del teorema de Hahn-Banach en la construcción de una medida finito-aditiva definida sobre toda la recta real y conocida como la medida de Banach.

La tesis incluye cuatro apéndices. En el primero se demuestra que aceptar el axioma de elección significa aceptar que para todo conjunto existe un buen orden, en particular que existe uno para el conjunto de reales. En el segundo se demuestra que un conjunto "totalmente perforado" no necesariamente tiene medida cero; es decir, que hay conjuntos densos en ninguna parte con medida positiva. El tercero contiene detalles técnicos relacionados con el teorema de Banach-Tarski. En el último apéndice se incluyen algunas reflexiones sobre las preguntas que quedan abiertas respecto al teorema de Hahn-Banach y el artículo de Solovay, pues con esta tesis se confirma que se cumple el análogo del teorema de incompletud de Gödel para la teoría de titulación: "No todas las preguntas sobre un tema se pueden responder en una tesis".

# Notación y convenciones

**ZERMELO FRAENKEL (ZF)**

**Axioma de Existencia:** El conjunto vacío existe.

**Axioma de Extensionalidad:** Sean $A, B$ dos conjuntos. Si $\forall x(x \in A \longleftrightarrow x \in B)$ entonces $A = B$.

**Esquema de Axiomas de Separación:** Sea $P(x)$ un propiedad de $x$, entonces para todo conjunto $A$ existe un conjunto $B$ tal que $B = \{x \in A : P(x)\}$.

**Axioma del Par:** Sean $A, B$ dos conjuntos, entonces existe un conjunto $C$ tal que $C = \{A, B\}$.

**Axioma de la Unión:** Para todo conjunto $X$ existe un conjunto $\cup X$ tal que $y \in \cup X$ si y sólo si $y \in A$ para alguna $A \in X$.

**Axioma de la Potencia:** Para todo conjunto $X$ existe un conjunto $\wp(X)$ tal que $y \in \wp(X)$ si y sólo si $y \subseteq X$.

**Axioma del Infinito:** Existe un conjunto $A$ tal que $\emptyset \in A$ y $\forall x(x \in A \Longrightarrow x \cup \{x\} \in A)$

**Esquemas de Axiomas de Reemplazo:** Sea $P(x, y)$ una propiedad tal que para todo $x$ hay un único $y$ tal que $P(x, y)$ se cumple, entonces para todo conjunto $A$, existe un conjunto $B$ tal que $B = \{y : \exists x \in A$ tal que $P(x, y)\}$

**Axioma de Regularidad o Buena Fundación:** Para todo conjunto no vacío A existe $y \in A$ tal que $\forall x(x \in y \rightarrow x \notin A)$.

**Observación 1** *ZF tiene una cantidad infinita de axiomas, ya que en los casos de separación y reemplazo por cada propiedad, respecto a los conjuntos, que se pueda expresar mediante una fórmula del lenguaje de ZF hay un axioma.*

**VERSIONES EQUIVALENTES DEL AXIOMA DE ELECCIÓN**

**Axioma de Elección (AE):** Para todo conjunto de conjuntos no vacíos existe una función de elección.

$\forall A(\forall x \in A(x \neq \emptyset) \Longrightarrow \exists f$ tal que $f$ es función, $dom(f) = A$ y $\forall x \in A$, $f(x) \in x)$.

**Existencia de un Conjunto de Representantes (ECR):** Para toda partición existe un conjunto que contiene uno y sólo un elemento de cada pedazo.

**Principio del Buen Orden (PBO):** Para todo conjunto existe un buen orden.

**Lema de Zorn:** Sea A un conjunto no vacío parcialmente ordenado mediante la relación $<_A$. Si toda cadena $C \subseteq A$ está acotada en $A$ entonces hay un elemento maximal $m \in A$ con relación al orden $<_A$.

**VERSIONES DÉBILES DEL AXIOMA DE ELECCIÓN**

$BO_{\mathbb{R}}$ = Existe un buen orden para los reales.

$ED$ =Elección dependiente.

$EN$ =Elección numerable.

## TEORÍA DE CONJUNTOS

$OR$ = clase de los ordinales.
$CAR$ = clase de los cardinales.
$\omega = \aleph_0 = |\mathbb{N}|$.
$\wp(X)$ = conjunto potencia de $X$.
$A^c$ = complemento de A (que por lo general será con respecto a $\mathbb{R}$).
$\langle A, < \rangle$ denotará el conjunto A ordenado por la relación $<$.
Un conjunto A es **Numerable** si $\exists\ f : A \rightarrow \mathbb{N}$ tal que $f$ es biyectiva.
Un conjunton A es **Contable** si es finito o numerable.
Un **unitario** es un conjunto con un sólo elemento.

## ANÁLISIS MATEMÁTICO

$c = |\mathbb{R}|$.
$I_{\mathbb{R}} = \{(a, b) : a, b \in \mathbb{R}\}$.
Si $A \subseteq \mathbb{R}^1$ y $q \in \mathbb{R}$ entonces $\mathbf{A}_{trasq} = \{x + q : x \in A\}$.

Si $f$ es una función con dominio $X$ y $M \subseteq X$, entonces detonamos con $\mathbf{f}\ |_M$ a la restricción de $f$ a M. Es decir, $f\ |_M$es una función con domino $M$ tal que $\forall x \in M\ f\ |_M\ (x) = f(x)$.

## LÓGICA MATEMÁTICA

Si $T$ es una teoría y $\alpha$ un enunciado:
$T \vdash \alpha$ si y sólo si "$\alpha$ es teorema de $T$".
Sea $L$ un lenguaje.
Si $M$ es una interpretación de $L$ y $\varphi$ es una fórmula de $L$:
$M \models \varphi$ si y sólo si "$\varphi$ es verdadera en $M$".
$ssi$ = si y sólo si.

# Parte I

# El problema de la medida

# Capítulo 1

# El tamaño de un conjunto

La noción de tamaño de un subconjunto de la recta real ha sido uno de los conceptos más conflictivos dentro de la historia de las matemáticas y, en particular, dentro del análisis matemático, pues de ella ha dependido la clasificación de las funciones integrables[1]. En términos generales, podemos decir que existen tres formas de considerar el tamaño de un conjunto de la recta real.

La primera, que en realidad es común a todos los conjuntos, sin importar si sus elementos son números o calcetines, tiene que ver con la cantidad de elementos que un conjunto tiene. Sin embargo, esta forma teórico conjuntista de referirse al tamaño de un conjunto, que se conoce como la cardinalidad, no permite una buena clasificación para la teoría de la integración de los subconjuntos de la recta real. Esto es porque, por un lado, hay conjuntos que pueden tener la misma cardinalidad pero ser muy distintos en cuanto a tamaño geométrico, como por ejemplo los intervalos $(0,1)$ y $(0,2)$; pero, por el otro, y quizás esto es lo más importante, hay funciones que a pesar de tener la máxima cantidad posible de discontinuidades (es decir, tantas como los reales) son Riemann integrables[2].

En un cierto sentido es afortunado que las condiciones de integrabilidad no dependan de argumentos de cardinalidad, pues de otra manera estaríamos confrontados con el problema de determinar cuáles son las posibles cardinalidades de un conjunto cualquiera de reales. Lo que no es de ninguna manera trivial, pues se trata ni más ni menos que del problema del continuo, que no es posible resolver desde ZFE[3].

---

[1] Establecer cómo podía ser el conjunto de discontinuidades de una función de manera que ésta fuera integrable se convirtió en un desafío importante, sobre todo cuando la definición de función como conjunto de pares ordenados desplazó la idea de que una función debía tener una expresión analítica. De ahí partió no sólo el trabajo de Lebesgue, sino también el de Cantor. Para más detalles y referencias históricas, véase Hawkings [12].

[2] Por ejemplo, la función característica del conjunto de Cantor, que tiene $\mathfrak{c}$ discontinuidades, pero es Riemann integrable.

[3] Aunque para los subconjuntos de reales se conocen tres cardinalidades: finito, numerable y $\mathfrak{c}$, si la hipótesis del continuo (HC) se rechaza, entonces quién sabe cuántos casos más habría que considerar. El trabajo de Paul Cohen dejó abierta esta última posibilidad, al demostrar que la negación de HC es consistente con ZFE (ver Capítulo 7 de Kunen [18] ).

Otra forma de pensar el tamaño de un conjunto, pero que tome en cuenta características más específicas de la recta real, es mediante conceptos topológicos. La idea es clasificar los conjuntos de acuerdo a su densidad. En este caso se tienen básicamente tres tamaños: el de los conjuntos densos en ninguna parte, el de las uniones numerables de densos en ninguna parte (conocidos como conjuntos de primera categoría) y el resto de los conjuntos (conocidos como conjuntos de segunda categoría). Esta forma de considerar el tamaño permite diferenciar conjuntos que tienen la misma cardinalidad pero no la misma densidad, como es el caso del intervalo (0,1) y el conjunto de Cantor[4].

Cabe señalar que los conjuntos densos en ninguna parte, como el de Cantor, son por definición conjuntos tales que el interior de su cerradura es vacío. Esto, en términos más intuitivos, quiere decir que están llenos de hoyos. Parecería natural esperar que el tamaño geométrico de los conjuntos densos en ninguna parte es absolutamente despreciable y, por lo tanto, que la separación entre conjuntos de primera categoría y conjuntos de segunda categoría permite distinguir a los conjuntos que, por decirlo de algún modo, son geométricamente nulos de los que no lo son. Sin embargo, como veremos, esta aproximación topológica no permite ni siquiera esa sencilla clasificación geométrica.

El hecho de que la recta real no sólo sea un espacio topológico, sino también un espacio métrico, sugiere la tercera forma de pensar en el tamaño de un conjunto. Hasta ahora nos hemos referido al tamaño geométrico de una forma bastante intuitiva. Sin embargo, nótese que cuando hemos dicho que los intervalos (0,1) y (0,2) no son iguales en cuanto a tamaño geométrico, a pesar de serlo en cardinalidad y densidad, es porque se tenía en mente que sus longitudes son distintas. Esto señala el camino para la formalización de la noción de tamaño geométrico, pues lo que se quisiera es que la definición no sólo coincida con la longitud cuando se trate de intervalos, sino que preserve algunas de las propiedades geométricas que posee la longitud. En particular, que sea invariante bajo isometrías[5].

Por otro lado, para la teoría de la integración la pregunta sobre el tamaño geométrico de un conjunto es especialmente relevante cuando se trata de los puntos de discontinuidad de una función. En ese sentido, si pensamos en la integral de Riemann, es importante que la generalización de la noción de longitud garantice que si el tamaño de un conjunto es geométricamente despreciable, entonces se puede cubrir con una cantidad finita de intervalos tal que la suma de sus longitudes sea tan pequeña como se quiera[6]. Esto explica por qué la noción topológica del tamaño de un conjunto no permite distinguir los conjuntos geométricamente nulos de los que no lo son; ya que hay conjuntos, como algunas generalizaciones del conjunto de Cantor, que son densos en ninguna parte pero que no cumplen con la condición que se le exigiría a un conjunto de tamaño

[4] Recuérdese que el conjunto de Cantor se obtiene partiendo un intervalo en tres, removiendo el pedazo de en medio y repitiendo recursivamente esta operación sobre los intervalos que quedan. En la Sección 3.3 se presenta la construcción con más detalle.

[5] También conocidas como el grupo de transformaciones rígidas. En $\mathbb{R}^1$ es el grupo generado por las traslaciones y las reflexiones.

[6] En cuyo caso se dice que el contenido exterior del conjunto es cero.

geométrico nulo desde la perspectiva de la teoría de la integración planteada por Riemann[7].

La pregunta que nos queda por responder es: ¿Se puede realmente generalizar la noción de longitud de manera que se tenga una definición formal del tamaño geométrico de cualquier conjunto de la recta real y que además satisfaga las necesidades de la teoría de la integración?

## 1.1. Longitud y tamaño geométrico

Para responder nuestra pregunta sería conveniente comenzar con formalizar la idea de longitud de un intervalo, que no es sino una función $lh : I_{\mathbb{R}} \longrightarrow [0, \infty]$, con $I_{\mathbb{R}}$ el conjunto de los intervalos abiertos, tal que $lh((a, b)) = |a - b|$. Lo ideal sería encontrar una función $\mu : \wp(\mathbb{R}^1) \longrightarrow [0, \infty]$ que:

(1) **A los intervalos les asigne su longitud.** Si $I = (a, b) \subseteq \mathbb{R}^1$, entonces $\mu(I) = |a - b|$.

(2) **Sea invariante bajo traslaciones.** Sea $A \subseteq \mathbb{R}^1$ y $x \in \mathbb{R}^1$. Si $A_{trasx} = \{a + x \quad | \; a \in A\}$, entonces $\mu(A_{trasx}) = \mu(A)$.

(3) **Sea finito-aditiva.** Si $A$, $B \subseteq \mathbb{R}^1$ y $A \cap B = \varnothing$, entonces $\mu(A \cup B) = \mu(A) + \mu(B)$.

(4) **Sea $\sigma$-aditiva.** Si $B = \bigcup_{i \in \mathbb{N}} A_i$, con $A_i \subseteq \mathbb{R}^1$ y $A_i \cap A_j = \varnothing$ si $i \neq j$, entonces $\mu(B) = \sum_{i \in \mathbb{N}} \mu(A_i)$[8].

**Definición 1** *Diremos que una función $\mu$ es extensión de $lh$ si $I_{\mathbb{R}} \subsetneq dom(\mu)$ y $\mu$ cumple con las condiciones (1) a (4).*

Nótese que para que una función $\mu$ sea extensión de $lh$ no se pide que $dom(\mu) = \wp(\mathbb{R}^1)$, pues cualquier función que extienda la noción de tamaño geométrico a conjuntos más complejos que la unión de intervalos abiertos es un avance. Desde luego que sería deseable tener una extensión de $lh$ cuyo domino contenga a todos los subconjuntos de la recta real, pero no podemos asegurar *a priori* que sea posible tenerla. Nos referiremos a esta posibilidad como la hipótesis de la medida (HM).

En principio, parecería natural creer que podemos referir el tamaño geométrico de todo conjunto del espacio euclideano, pues justamente ocupa un "espacio". Sin embargo, la intuición no siempre tiene su contraparte formal.

[7] De hecho, ni siquiera por la planteada por Lebesgue, que sí permite una cantidad numerable de intervalos.

[8] Esta propiedad puede parecer artificial, pero es fundamental para extender la noción de integral. Por ejemplo, la función característica de los racionales ($X_{\mathbb{Q}}$) no es Riemann integrable, pues el ínfimo de las sumas superiores de Riemann es 1 y el supremo de las sumas inferiores de Riemann es 0. Sin embargo, sí es Lebesgue integrable, pues es igual a la función constante cero excepto en $\mathbb{Q}$, que tiene medida cero gracias a la $\sigma-$aditividad de la medida de Lebesgue. Además la $\sigma-$aditividad de la medida está presente en teoremas de convergencia fundamentales para al teoría de la integración (ver Secciones 2.2 y 2.3 de Folland [9]).

**Definición de medida**

**Definición 2** *Sea $X$ un conjunto no vacío. Una $\sigma-$álgebra sobre $X$ es un conjunto $\Im \subseteq \wp(X)$ tal que:*

1) $X \in \Im$
2) Si $A \in \Im$ entonces $A^c \in \Im$.
3) Si $\forall i \in \mathbb{N}$, $A_i \in \Im$ entonces $B = \bigcup_{i \in \mathbb{N}} A_i \in \Im$.

Estas propiedades indican que $\Im$ es cerrado bajo uniones finitas y numerables, así como bajo la operación complemento[9]. En lo sucesivo denotaremos con $\Im$ a las $\sigma$-álgebras.

**Definición 3** *Sea $\Im \subseteq \wp(X)$. Una medida no trivial sobre $\Im$ es una función $\mu : \Im \to [0, \infty]$:*

(1) Finito-aditiva. Si $A, B \in \Im$ y $A \cap B = \varnothing$, entonces $\mu(A \cup B) = \mu(A) + \mu(B)$.

(2) $\sigma -$ *aditiva.* Si $B = \bigcup_{i \in \mathbb{N}} A_i \in \Im$, con $A_i \in \Im$ y $A_i \cap A_j = \varnothing$ si $i \neq j$, entonces $\mu(B) = \sum_{i \in \mathbb{N}} \mu(A_i)$.

(3) No trivial. Existe $A \in \Im$ tal que $\mu(A) > 0$.

Con esta definición podemos plantear el problema de la medida de una forma más general. Es decir, quisiéramos saber si para todo conjunto $X$ es posible tener una medida definida sobre toda la potencia de $X$ o, lo que es lo mismo, si existe una función $\mu : \wp(X) \longrightarrow [0, \infty]$ que cumpla con las condiciones (1) a (3). Podríamos referirnos a esta pregunta como el problema generalizado de la medida y a la respuesta afirmativa como la hipótesis generalizada de la medida (HGM)[10]. Desde luego que el caso que más nos interesa aquí es cuando $X = \mathbb{R}^1$, pero más adelante se podrá constatar que esta aproximación más general al problema de la medida arroja luz sobre la causa de la existencia de subconjuntos de la recta real no medibles[11]. En particular, se podrá ver que la existencia de conjuntos no medibles no depende de las condiciones (3) y (4) que pedimos para que una medida $\mu$ sea extensión de $lh$.

Por último, antes de abordar con más detenimiento el problema de la medida, sería conveniente hacer algunas observaciones respecto a la forma de considerar el tamaño de un conjunto desde una perspectiva teórico conjuntista. En términos generales, podría decirse que la cardinalidad de un conjunto es la cantidad de elementos que tiene. Sin embargo, qué es lo que exactamente se entiende por eso es algo que depende de si se acepta o no el axioma de elección.

---

[9] Nótese que $\wp(X)$ es una $\sigma-$álgebra.

[10] Estas denominaciones se inspiran en la hipótesis del continuo (HC) y la hipótesis generalizada del continuo (HGC).

[11] Pues aunque no se tienen resultados respecto a HGM, sí se tienen resultados respecto a versiones que (suponiendo la hipótesis del continuo) están entre HM y HGM (ver Sección 4.1).

La forma más débil de pensar la cardinalidad es en términos puramente compa- rativos, en la que sólo se considera si un conjunto A tiene la misma cantidad de elementos que otro conjunto B (en cuyo caso decimos que son equipotentes), pero no pensamos en la cantidad de elementos de A como un "número"[12], pues para ello es necesario suponer el axioma de elección. En ese sentido, si denotamos con $c$ la cardinalidad de los reales, cuando se dice que un conjunto $A$ tiene cardinalidad $c$ muchas veces nos referimos a que existe una biyección entre el conjunto de los reales y el conjunto $A$. De modo que, para ser más precisos, en lugar de decir que $A$ tiene cardinalidad $c$ deberíamos decir que $A$ es equipotente con $\mathbb{R}$ y así reservar la notación de $c$ para el caso en que, suponiendo el axioma de elección, se considere el cardinal de los reales como el mínimo ordinal con el cual es biyectable[13]. Sin embargo, muchas pruebas, sobre todo de análisis matemático, utilizan o demuestran simplemente que un conjunto es equipotente con los reales, al margen de si existe un buen orden para ellos o no (es decir, al margen del axioma de elección) aunque suelan manejar la versión más fuerte de cardinalidad. Un ejemplo de ello es el problema del continuo y la hipótesis del continuo que se pueden plantear sin suponer el axioma de elección[14].

## 1.2. Conjuntos no medibles: Vitali

Regresando a nuestra discusión sobre si es posible tener o no una extensión de $lh$ cuyo dominio sea la potencia de la recta real, a continuación veremos que si se aceptan todos los axiomas de ZFE es imposible tener dicha función.

**Notación 1** $\langle \mathbb{R}, <_{\mathbb{R}} \rangle$ *denotará un buen orden sobre los reales.*

Definimos sobre el intervalo $[0,1)$ la siguiente relación de equivalencia $\sim_v$: $\forall x, y \in [0,1)[x \sim_v y \Leftrightarrow \exists r \in \mathbb{Q}(x - y = r)]$.

Esta relación, como toda relación de equivalencia, induce una partición del intervalo $[0,1)$. Es decir, tenemos una familia de conjuntos $\mathcal{F}_{\sim_v} = \{C_i\}_{i \in I}$ tal que:

1) $[0,1) = \bigcup\limits_{i \in I} C_i$.
2) $C_i \cap C_j = \varnothing$ si $i \neq j$.
3) $C_i \neq \varnothing$ para toda $i \in I$.

**Lema 1** *La relación $\sim_v$ induce $c$ clases de equivalencia sobre el intervalo* $[0,1)$.

---

[12] En el caso de los conjuntos infinitos, los números con los que contamos son los ordinales, que son la generalización de los naturales.

[13] Cabe señalar que hay otra forma, además de la equipotencia, de pensar el cardinal de un conjunto sin usar axioma de elección, aunque en ese caso se tiene que suponer el axioma de regularidad (ver Amor [1] p. 58).

[14] Estas dos versiones se introducen en la Sección 6.2.

Para demostrar este lema necesitamos primero demostrar que:

**Proposición 2** $\forall C_i \in \mathcal{F}$, $|C_i| = \aleph_0$ *(Cada una de las clases de equivalencia $C_i$ es numerable).*

**Demostración.** Sea $\langle \mathbb{R}, <_{\mathbb{R}} \rangle$ un buen orden. Sean $C_i \in \mathcal{F}$ y $(q_n)_{n=0}^{\infty}$ una numeración de los elementos de $\mathbb{Q} \cap [0,1)$. Sea $x_0 = \min_{<_{\mathbb{R}}} C_i$[15] y $a_n = x_o + q_n$. Entonces $\{a_n\}_{n=0}^{\infty}$ es una numeración de los elementos de $C_i$, ya que $a_n - q_n = x_o$. De modo que $a_n \sim_v x_0$ y $a_n \in C_i$. Por otro lado, si $x \in C_i$ entonces $x = x_0 + p$, para alguna $p \in \mathbb{Q} \cap [0,1)$. Así que para toda $x \in C_i$, existe $m \in \mathbb{N}$ tal que $x = x_0 + q_m$. Por lo tanto, $|C_i| = \aleph_0$. ■

Ahora sí podemos dar la demostración del Lema 1:
**Demostración.** Como $\mathcal{F} = \{C_i\}_{i \in I}$ es partición de [0,1), entonces $c = 2^{\aleph_0} = |[0,1)| = \left| \bigcup_{i \in I} C_i \right| = \max\{|I|, \aleph_0\}$. Por lo tanto, $c = |I|$ (es decir, hay un cantidad $c$ de clases de equivalencia)[16]. ■

**Definición 4** *Si nos apegamos a la notación anterior, decimos que un conjunto $V$ es de Vitali si y sólo si $\forall C_i \in \mathcal{F}\ \exists x(C_i \cap V = \{x\} \wedge \forall y(y \in C_i \cap V \rightarrow y = x))$. Es decir, un conjunto de Vitali es un conjunto que contiene exactamente un representante de cada clase de equivalencia $C_i$.*

**Observación 2** $\forall x, y \in V(x \neq y \implies x \nsim_v y)$, *pues* $x, y \in V(x \sim_v y) \implies \exists C_i \in \mathcal{F}(x, y \in C_i \cap V) \implies x = y$.

Es común que dada la definición de un conjunto se asuma que éste existe, sin verificar qué axiomas de la teoría de conjuntos nos garantizan su existencia. Así que aunque hemos definido al conjunto de Vitali, la pregunta por su existencia es pertinente.

**Lema 3** *Supongamos que existe un buen orden para los reales, entonces existe un conjunto de Vitali.*

**Demostración.** Tómese $V = \{x_i \mid C_i \in \mathcal{F} \wedge x_i = \min_{<_{\mathbb{R}}} C_i\}$. ■

**Observación 3** *El conjunto de Vitali se suele dar apelando al axioma de elección en su versión: "Toda partición tiene un conjunto de representantes". Sin embargo, acabamos de ver que se puede dar mediante una versión más débil de elección, a saber: "Existe un buen orden para los reales"*[17].

---

[15] Nótese que $x_0$ siempre existe puesto que $C_i \neq \emptyset \subset \mathbb{R}$ y $<_{\mathbb{R}}$ es un buen orden.

[16] Demostración alternativa. Supongamos que hay $\alpha < c$ clases de equivalencia. Entonces $c = |[0,1)| = \left| \bigcup_{\beta \in \alpha} C_\beta \right|$. Por otro lado, como $C_\zeta \cap C_\xi = \varnothing$ para $\zeta \neq \xi$, $\left| \bigcup_{\beta \in \alpha} C_\beta \right| = \sum_{\beta \in \alpha} |C_\beta| = \max\{\alpha, \aleph_0\} < c$. Esto querría decir que $c < c$! Por lo tanto, hay una cantidad $c$ de clases de equivalencia.

[17] Para tener esta versión basta con que exista una función de elección sobre la potencia de los reales, pues en general se tiene que *Un conjunto $A$ es bien ordenable ssi $\wp(A) - \{\emptyset\}$ tiene una función de elección* (ver Apéndice A).

**Teorema 4 (Vitali)** *Sea $V$ un conjunto de Vitali, entonces no existe $\mu$ una extensión de $lh$ tal que $dom(\mu) = \wp(\mathbb{R}^1)$.*

**Demostración.** Para demostrar este teorema induciremos otra partición sobre el intervalo $[0, 1)$ usando como base el conjunto de Vitali[18]. Cada clase de equivalencia será resultado de una partición en dos del conjunto de Vitali, seguida de una traslación de cada uno de los dos pedazos. Definimos, dada $q \in \mathbb{Q} \cap [0, 1)$, el siguiente conjunto:

$V_q = \{s + q \mid s \in V \wedge s + q < 1\} \cup \{(t + q) - 1 \mid t \in V \wedge t + q > 1\}$.

Sea $(p_n)_{n\in\mathbb{N}}$ una numeración de los elementos de $\mathbb{Q} \cap [0, 1)$. Afirmamos que $F' = \{V_{p_n}\}_{n\in\mathbb{N}}$ es una partición de $[0,1)$.

Como $\forall q \in \mathbb{Q} \cap [0, 1)[V_q \neq \emptyset \wedge V_q \subseteq [0, 1)]$. Sólo hay que verificar que:
a) $\forall q, p \in \mathbb{Q} \cap [0, 1)(p \neq q \Longrightarrow V_p \cap V_q = \varnothing)$.
b) $\forall x \in [0, 1)\exists r \in \mathbb{Q} \cap [0, 1)(x \in V_r)$[19].

(a) Sea $y \in V_p \cap V_q$, con $p \neq q$. Tenemos tres casos:
Caso 1. $\exists s \in V(y = s + p) \wedge \exists r \in V(y = r + q) \wedge p \neq q \Longrightarrow$
$\exists s, r \in V(s + p = r + q) \wedge p \neq q \Longrightarrow$
$\exists s, r \in V(s - r = q - p \in \mathbb{Q}) \wedge q - p \neq 0 \Longrightarrow$
$\exists s, r \in V(s \neq r \wedge s \sim_v r)$! (Observación 2).
Caso 2. $\exists s \in V(y = s + p) \wedge \exists r \in V(y = (r + q) - 1) \wedge p \neq q \Longrightarrow$
$\exists s, r \in V(s + p = (r + q) - 1) \wedge p \neq q \Longrightarrow$
$\exists s, r \in V(s - r = (q - p) - 1) \in \mathbb{Q} \wedge (q - p) - 1 \neq 0$ (pues $q, p \in \mathbb{Q} \cap [0, 1) \Longrightarrow$
$\exists s, r \in V(s \neq r \wedge s \sim_v r)$! (Observación 2).
Caso 3. $\exists s \in V(y = (s + p) - 1) \wedge \exists r \in V(y = (r + q) - 1) \wedge p \neq q \Longrightarrow$
$\exists s, r \in V(s - r = q - p \in \mathbb{Q}) \wedge q - p \neq 0 \Longrightarrow$
$\exists s, r \in V(s \neq r \wedge s \sim_v r)$! (Observación 2).
Como en los tres casos se llega a una contradicción, tenemos que $V_p \cap V_q = \varnothing$.

(b) Sea $x \in [0, 1)$. Si $x \in V$ entonces $x \in V_0 = V$. Si $x \notin V$ entonces existe $C_i \in \mathcal{F}$ tal que $x \in C_i$ (pues $[0, 1) = \bigcup_{i\in I} C_i$). De modo que existen $q \in \mathbb{Q}$ y $x_i \in V$ tales que $x - x_i = q$ (i.e. $x \sim_v x_i$). Por otro lado, como $x, x_i \in [0, 1)$, $q \in \mathbb{Q} \cap [0, 1)$ y $x = x_i + q < 1$. Así que $x \in V_q$.
Por lo tanto, $[0, 1) \subset \bigcup_{q\in\mathbb{Q}\cap[0,1)} V_q$.

Ahora completaremos la demostración del Teorema 4 :
Sea $\mu$ una extensión de $lh$. Supongamos que $dom(\mu) = \wp(\mathbb{R}^1)$.
Por un lado, como estamos suponiendo que $\mu$ es una extensión de $lh$, entonces $\mu([0, 1)) = 1$[20].

---

[18] Es interesante notar que con la partición inducida por $\sim_v$ se tenían $2^{\aleph_0}$ clases de equivalencia, cada una de cardinal $\aleph_0$. Mientras que ahora daremos una partición en $\aleph_0$ clases de equivalencia, cada una de cardinal $2^{\aleph_0}$. Lo que nos permitirá usar la $\sigma-$aditividad de $\mu$.

[19] Nótese que la unicidad de $r$ se desprende del inciso anterior, así que no hace falta demostrarla.

[20] Por hipótesis tenemos que $\mu((0, 1)) = 1$. Además, como se verá cuando se trabaje la unicidad de la medida de Lebesgue (Sección 29), dada $a \in \mathbb{R}$, $\mu(\{a\}) = 0$. Así que usando la finito aditividad tendríamos que $\mu([0, 1)) = 1$.

Por otro lado, dada $p \in \mathbb{Q} \cap [0,1)$, podemos partir el conjunto de Vitali de la siguiente manera:

$V = V^p_{<1} \cup V^p_{>1}$, con $V^p_{<1} = \{s \in V \mid p+s<1\}$ y $V^p_{>1} = \{t \in V \mid t+p>1\}$.

Está claro que $V^p_{<1} \cap V^p_{>1} = \varnothing$. Así que por las propiedades de $\mu$ tenemos que $\mu(V) = \mu(V^p_{<1}) + \mu(V^p_{>1})$. Además $V_p = (V^p_{<1})_{trasp} \cup (V^p_{>1})_{tras(p-1)}$. Es decir, $V_p$ es la traslación de los dos pedazos en los que se puede describir al conjunto de Vitali, aunque cada uno trasladado por un racional distinto. Como $\mu$ es invariante bajo traslaciones, tenemos que $\mu(V^p_{<1}) = \mu((V^p_{<1})_{trasp})$ y $\mu(V^p_{>1}) = \mu((V^p_{>1})_{tras(p-1)})$. Por lo tanto, $\mu(V) = \mu(V^p_{<1}) + \mu(V^p_{>1}) = \mu((V^p_{<1})_{trasp}) + \mu((V^p_{>1})_{tras(p-1)}) = \mu(V_p)$. De modo que dada $p \in \mathbb{Q} \cap [0,1)$, $\mu(V) = \mu(V_p)$.

Además como $\mu$ es $\sigma-$aditiva y $V_n \cap V_m = \varnothing$ si $m \neq n$, entonces $\mu([0,1)) = \mu(\bigcup_{n\in\mathbb{N}} V_n) = \sum_{n\in\mathbb{N}} \mu(V_n) = \sum_{n\in\mathbb{N}} \mu(V)$. Así que:

$\mu(V) = 0 \Longrightarrow \mu([0,1)) = 0!$ y

$\mu(V) > 0 \Longrightarrow \mu([0,1)) = \infty!$

Por lo tanto, no existe $\mu$ una extensión de $lh$ tal que $dom(\mu) = \wp(\mathbb{R}^1)$[21]. ■

Tenemos entonces que el conjunto de Vitali impide tener una medida que extienda de manera adecuada a $lh$[22]. ¿Qué clase de conjunto extraño es Vitali? La construcción que hemos dado no parece encerrar patología alguna, pues simplemente hemos tomado el mínimo de cada una de las clases de equivalencia. Pero, ¿realmente sabemos cómo es Vitali? La repuesta es no, pues aunque nuestra construcción parece ser bastante explícita, se basa en la existencia de un buen orden para los reales del que sólo sabemos eso: que existe, pero no cómo es. De modo que también del conjunto de Vitali sólo sabemos que existe, pero no cómo es. En resumen lo que tenemos es que:

$$\begin{array}{ll} ZF + BO_{\mathbb{R}} & \vdash \exists V(\forall C_i \in \mathcal{F}\exists! x_i(C_i \cap V = \{x_i\})). \\ ZF & \vdash \exists V(\forall C_i \in \mathcal{F}\exists! x_i(C_i \cap V = \{x_i\})) \longrightarrow \neg HM. \\ ZF + BO_{\mathbb{R}} & \vdash \neg HM. \end{array}$$

El resultado más famoso es $ZFE \vdash \neg HM$, que claramente se desprende del anterior, pues el axioma de elección implica que existe un buen orden para los reales (ver Apéndice A). Así que cualquiera que acepte el axioma de elección se verá obligado a aceptar que $HM$ es falsa.

Aquí se ha querido presentar la demostración de Vitali bajo la hipótesis más débil de que existe un buen orden para los reales por dos razones: una porque con esto se ve que para rechazar la hipótesis de la medida no hace falta ser un fervoroso creyente del axioma de elección, sino que basta suponer una versión más débil de elección. La segunda razón, y quizás la más importante, tiene que ver con la amplia gama de versiones de elección que se tienen[23] y los alcances y deficiencias de cada una de ellas.

[21] Nótese que de la contradicción no se concluye necesariamente que Vitali no sea $\mu-$ medible, pues la contradicción se evitaría simplemente si para alguna $p \in [0,1) \cap \mathbb{Q}$ sucede que $V^p_{<1}$ ó $V^p_{>1}$ no es $\mu-$ medible.

[22] Es decir, que esté definida para todos los subconjuntos de reales.

[23] Rubin presenta más de cien versiones en su libro (ver Rubin [20]).

Como la hipótesis de que exista un buen orden para los reales es central en la demostración que hemos dado, parecería natural preguntarse qué sucedería con el problema de la medida si desechamos esa hipótesis. Es decir, ¿será que $ZF \nvdash \neg HM$ o dicho de otro modo, que $ZF + HM$ es consistente? Desde un punto de vista lógico-conjuntista esta pregunta puede ser muy interesante, sin embargo no lo es para teoría de la medida, pues no podemos olvidar que el problema de la medida tiene como horizonte la teoría de la integración, que sería inconcebible sin alguna forma de elección infinita[24]; así que no podemos darnos el lujo de desechar toda forma de elección infinita. Sin embargo, podríamos pensar la misma pregunta pero con versiones de elección más débiles que la existencia de un buen orden para los reales. Es decir, quisiéramos saber si existe una versión débil de axioma de elección que sea compatible con la hipótesis de la medida y que nos garantice la mayor cantidad posible de resultados del análisis matemático.

En la demostración que hemos dado de la negación de la hipótesis de la medida no sólo se utilizó la existencia de un buen orden para los reales, sino también dos de las propiedades que se le piden a la función $\mu$ (invarianza bajo traslaciones y $\sigma - aditivad$). De modo que, antes de considerar la posibilidad de desechar el axioma de elección o debilitarlo, valdría la pena preguntarse qué pasa si debilitamos las condiciones que se le piden a $\mu$.

[24] La equivalencia entre las distintas definiciones de convergencia de una sucesión, de continuidad de una función, entre otros resultados de análisis, dependen de elección numerable (ver Hrbacek y Jech [14] p.145).

# Capítulo 2

# Aditividad

En este capítulo veremos qué sucede si en lugar de pedir que $\mu$ sea $\sigma-aditiva$ pedimos tan sólo que sea finito-aditiva. Es decir, si $HM_{-4}$ denota la hipótesis de la medida sin la condición de que sea $\sigma-aditiva$, entonces nos preguntamos lo siguiente: ¿Será que ZFE $\vdash \neg HM_{-4}$?

## 2.1. Medida de Banach

La respuesta a esta pregunta es especialmente interesante, pues no sólo se tiene que ZFE $\nvdash \neg HM_{-4}$, sino un resultado aún más fuerte y ciertamente bastante sorprendente: ZFE $\vdash HM_{-4}$.

En la demostración de este resultado se hace uso indirecto del axioma de elección, mediante una hermosa aplicación del teorema de Hahn-Banach. Veremos que es posible dar, utilizando dicho teorema, una medida finito-aditiva $\mu : \wp(\mathbb{R}^1) \longrightarrow [0, \infty]$ que a los intervalos les asigne su longitud y que sea invariante bajo traslaciones. Así que tenemos que ZF + Hahn-Banach $\vdash HM_{-4}$ y, como el teorema de Hahn-Banach se demuestra utilizando el lema de Zorn (equivalente al axioma de elección), también tenemos que ZFE $\vdash HM_{-4}$.

Antes de dar la demostración de este resultado conocido como la medida de Banach[1], daremos algunas definiciones y enunciaremos, aunque no demostraremos, el teorema de Hahn-Banach[2].

**Definición 5** *Sea $X$ un espacio vectorial sobre $\mathbb{R}$. Una función $p$: $X \longrightarrow \mathbb{R}$ es un funcional sublineal si y sólo si es*

a) Positivo homogénea. $\forall \alpha \geq 0$ y $\forall x \in X$ $(p(\alpha x) = \alpha p(x))$.
b) Subaditiva. $\forall x, y \in X$, $p(x+y) \leq p(x) + p(y)$.

---

[1] Banach dió con este resultado al trabajar sobre el problema de la medida (ver Bachman y Narici [4]).
[2] La demostración puede consultarse en Folland [9] (pp.149-150).

**Teorema 5 (Hahn-Banach)** *Sea $\mathcal{M}$ un subespacio de un espacio vectorial $X$ sobre $\mathbb{R}$, sea $p$ un funcional sublineal y sea $f$ un funcional lineal sobre $\mathcal{M}$ tal que $\forall x \in \mathcal{M}$ :*

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

*Entonces existe un funcional lineal $F$ definido en todo $X$ tal que $F$ extiende a $f$. Es decir, $\forall x \in \mathcal{M}$, $F(x) = f(x)$ y además $\forall x \in X$ :*

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

.

**Teorema 6 (Medida de Banach)** *Existe $\mu : \wp([0,1)) \to [0,\infty]$ tal que:*

(1) A los intervalos les asigna su longitud. $\forall A = (a,b) \subset [0,1)$, $\mu(A) = |a-b|$.

(2) Es invariante bajo traslaciones. $\forall A \subset [0,1), \forall x \in \mathbb{R}$ $(A_{trasx} \subset [0,1) \Longrightarrow \mu(A_{trasx}) = \mu(A))$.

(3) Es finito-aditiva. $\forall A = \{A_i\}_{i=1}^{n}[A_i \subset [0,1) \wedge (i \neq j \Rightarrow A_i \cap A_j = \emptyset) \Longrightarrow \mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i)]$.

**Demostración.** Como ya hemos señalado antes, para la demostración de este teorema se utilizará el teorema de Hahn-Banach. Así que lo primero es dar el espacio vectorial sobre el que se va a trabajar. Sea $X$ la clase de las funciones de $\mathbb{R}^1$ en $\mathbb{R}^1$ acotadas y periódicas, con periodo 1. Es decir,

$$X = \{f \mid f : \mathbb{R}^1 \to \mathbb{R}^1 \wedge \exists C \in \mathbb{R}\ (\forall x \in \mathbb{R}(|f(x)| < C \wedge f(x) = f(x+1))\}.$$

Claramente $X$ es un espacio vectorial[3]. Lo siguiente será dar el funcional sublineal $p$, un subespacio $\mathcal{M}$ y un funcional lineal $f$ sobre $\mathcal{M}$, para poder aplicar el teorema de Hahn-Banach y obtener un funcional lineal $F$ sobre todo $X$, a partir del cual se construirá lo que se conoce como la medida de Banach y que cumple $HM_{-4}$. Sea $S = \{\{a_i\}_{i=1}^{n} : n \in \mathbb{N} \wedge a_i \in \mathbb{R}\}$, el conjunto de sucesiones finitas de reales. Definimos $p(g)$, para $g \in X$, de la siguiente manera:

$$p(g) = \inf_{s \in S} \left\{ \sup_{0 \leq x < 1} \{\frac{1}{n} \sum_{i=1}^{n} g(x + a_i) : a_i \in s = \{a_i\}_{i=1}^{n}\} \right\}.$$

Obsérvese que, aunque el supremo se toma haciendo variar a $x$ sobre el intervalo $[0,1)$(es decir, tomando una sucesión fija), el ínfimo corre sobre todas las posibles sucesiones finitas de reales.

[3] Las funciones de $\mathbb{R}^1$ en $\mathbb{R}^1$ que son acotadas forman un espacio vectorial. Así que sólo habría que probar que $X$ es un subespacio:

(1) La función constante 0 tiene periodo 1.

(2) Si $f, g \in X$ y $\alpha, \beta \in \mathbb{R}$ entonces $(\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x) = \alpha f(x+1) + \beta g(x+1) = (\alpha f + \beta g)(x+1)$.

(3) Evidentemente $(f \in X \Longrightarrow -f \in X)$.

En otras palabras, dada una sucesión $s = \{a_i\}_{i=1}^n$, cada $x \in [0,1)$ induce una partición $p_x = \{x + a_o, x + a_1, ..., x + a_n\}$ del intervalo [0,1) a partir de la cual se toma el promedio de los valores asígnados por $g$ a cada uno de los puntos de la partición. Para cada sucesión se toma el supremo de los promedios cuando se deja correr libremente a $x$ sobre el intervalo [0,1). Y por último se toma el ínfimo del conjunto de los supremos generados por todas las posibles sucesiones finitas de reales.

Hay que verificar que $p$ es un funcional sublineal.

(a) $p$ es positivo homegénea. En general, se tiene que $\forall A \subset \mathbb{R}$ y $a \geq 0$, $\inf aA = a \inf A$ y $\sup aA = a \sup A$. Usando esto es muy fácil demostrar que para $a \geq 0$ y $g \in X$, $p(ag) = ap(g)$.

Si para $s = \{a_i\}_{i=1}^n \subset \mathbb{R}$ tomamos $A_s = \left\{\frac{1}{n}\sum_{i=1}^{n} g(x + a_i) : 0 \leq x < 1\right\}$ entonces $\sup_{0\leq x<1} \frac{1}{n}\sum_{i=1}^{n} ag(x + a_i) = \sup_{0\leq x<1} a(\frac{1}{n}\sum_{i=1}^{n} g(x + a_i)) = \sup aA_s$ y como $A_s \subset \mathbb{R}$, tenemos que $\sup_{0\leq x<1} \frac{1}{n}\sum_{i=1}^{n} ag(x + a_i) = a \sup A_s$. Pero además, si tomamos $A' =\{\sup_{0\leq x<1} A_s \mid s \in S\}$ entonces

$$\inf_{s\in S}\left\{a \sup_{0\leq x<1}\{\frac{1}{n}\sum_{i=1}^{n} g(x + a_i)\right\} = \inf aA'.$$

Como $A' \subset \mathbb{R}$ y $a \geq 0$, tenemos que $\inf aA' = a \inf A'$. Así que $p(ag) = \inf aA' = a \inf A' = a \inf\{\sup_{0\leq x<1} A_s \mid s \in S\} = ap(g)$

Por lo tanto, $p$ es positivo homogénea.

(c) $p$ es sublineal. Sean $h, g \in X$. Hay que demostrar que $p(h + g) \leq p(h) + p(g)$. Como $p$ involucra en última instancia un ínfimo, entonces dada $\epsilon > 0$ existen $s_n = \{a_i\}_{i=1}^n$ y $s_m = \{b_j\}_{j=1}^m$ sucesiones de reales tales que:

$\sup_{0\leq x<1}(\frac{1}{n}\sum_{i=1}^{n} g(x + a_i)) \leq p(g) + \epsilon$ y $\sup_{0\leq x<1}(\frac{1}{m}\sum_{i=1}^{m} h(x + b_i)) \leq p(h) + \epsilon$.

Sea $s_{mn} = \{a_i + b_j\}_{i=1,j=1}^{n,m} = \{c_l\}_{l=1}^{mn}$.

Entonces por definición de $p$ :

$p(g + h) \leq \sup_{0\leq x<1}(\frac{1}{mn}\sum_{l=1}^{mn}(g + h)(x + c_l))$.

Donde $\frac{1}{mn}\sum_{l=1}^{mn}(g + h)(x + c_l) = \frac{1}{mn}\sum_{l=1}^{mn} g(x + c_l) + \frac{1}{mn}\sum_{l=1}^{mn} h(x + c_l) =$

$\frac{1}{m}\sum_{j=1}^{m}\frac{1}{n}\sum_{i=1}^{n} g(x + a_i + b_j) + \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m}\sum_{j=1}^{m} h(x + b_j + a_i)$.

Como $g$ tiene periodo 1, si tomamos a $j$ fija entonces:

$\sup_{0\leq x<1} \frac{1}{n}\sum_{i=1}^{n} g(x + a_i + b_j) = \sup_{b_j\leq x<1+b_j} \frac{1}{n}\sum_{i=1}^{n} g(x + a_i) =$

$\sup_{0\leq x<1} \frac{1}{n}\sum_{i=1}^{n} g(x + a_i)$.

Análogamente para $h$ con $i$ fija. De modo que :

$\frac{1}{m}\sum\limits_{j=1}^{m}\left(\frac{1}{n}\sum\limits_{i=1}^{n} g(x+a_i+b_j)\right)+\frac{1}{n}\sum\limits_{i=1}^{n}\left(\frac{1}{m}\sum\limits_{j=1}^{m} h(x+b_j+a_i)\right)\leq$

$\frac{1}{m}\sum\limits_{j=1}^{m}\left(\sup_{0\leq x<1}\frac{1}{n}\sum\limits_{i=1}^{n} g(x+a_i)\right)+\frac{1}{n}\sum\limits_{i=1}^{n}\left(\sup_{0\leq x<1}\frac{1}{m}\sum\limits_{j=1}^{m} h(x+b_j)\right)=$

$\sup_{0\leq x<1}\frac{1}{n}\sum\limits_{i=1}^{n} g(x+a_i)+\sup_{0\leq x<1}\frac{1}{m}\sum\limits_{j=1}^{m} h(x+b_j)$

$\leq p(g)+p(h)+2\epsilon.$

Así que dada $\epsilon>0$, existe una sucesión $s=\{c_l\}_{l=1}^{mn}$ tal que $\sup_{0\leq x<1}(\frac{1}{mn}\sum\limits_{l=1}^{mn}(g+h)(x+c_l))\leq p(g)+p(h)+2\epsilon$. De donde se puede concluir que:

$\inf_{s\in S}\left\{\sup_{0\leq x<1}\frac{1}{mn}\sum\limits_{l=1}^{mn}(g+h)(x+a_l)\right\}\leq p(g)+p(h).$

Por lo tanto, $p(g+h)\leq p(g)+p(h)$.

Ahora daremos el subespacio $\mathcal{M}$ y el funcional lineal $f$ acotado por el funcional sublineal $p$.

Sea $\mathcal{M}=\{g\mid\forall x\in\mathbb{R},\ g(x)=0\}$. Nótese que $\mathcal{M}$ es el counjunto que tiene como único elemento a la función constante cero, que denotaremos con $\hat{O}$, así que trivialmente se tiene que $\mathcal{M}$ es un subespacio de X. Sea $f=p\mid_{\mathcal{M}}$, entonces $f$ es un funcional lineal sobre $\mathcal{M}$. Es decir, $\forall g,h\in\mathcal{M}$ y $\alpha,\beta\in\mathbb{R}$, $f(\alpha g+\beta h)=\alpha f(g)+\beta f(h)$. Pero como $\mathcal{M}$ tiene a $g=\hat{O}$ como único elemento, en realidad sólo hay que demostrar que $\forall\alpha\in\mathbb{R}$, $p(\alpha\hat{O})=\alpha p(\hat{O})$. Esto se sigue inmediatamente del hecho de que $\alpha\hat{O}=\hat{O}$ y de que $p(\hat{O})=0$. Entonces tenemos $\mathcal{M}$ un subespacio de $X$, $p$ un funcional sublineal definido en todo $X$ y $f$ un funcional lineal sobre $\mathcal{M}$ tal que $\forall g\in\mathcal{M}$:.

$f(g)\leq p(g)$[1].

Entonces, por el teorema de Hahn-Banach, existe un funcional lineal $F$ definido sobre todo $X$ que extiende a $f$ y tal que $\forall g\in X$ :

$$F(g)\leq p(g).$$

Para dar la medida de Banach necesitamos la siguiente definición:

Sea $\{z_i\}_{i=1}^{\infty}$ una numeración de $\mathbb{Z}$. Entonces dada $E\subset[0,1)$ sea $\Gamma_E=\sum\limits_{i=1}^{\infty}\chi_{E_{tras\,z_i}}$. Nótese que $\Gamma_E$ es la versión periódica (con periodo 1) de la función característica de un conjunto contenido en [0,1), ya que $E_{tras(z_i)}=\{x+z_i:x\in E\}$.

Definimos la medida de Banach $\mu_B$ de la siguiente manera:

$$\forall E\subset[0,1),\mu_B(E)=F(\Gamma_E).$$

[1] Esto se cumple porque $f$ es la restricción de $p$ a $\mathcal{M}$.

Nótese que $\mu_B$ está bien definida, pues dada $E \subset [0,1)$, $\Gamma_E \in X$[5]. Lo siguiente será demostrar que esta medida es finito-aditiva, invariante bajo traslaciones y que asigna la longitud a los intervalos.

**(1)** $\mu_B$ **es finito-aditiva.** Sea $E = \bigcup_{j=1}^{n} E_j \subset [0,1)$, con $E_i \cap E_j = \emptyset$ si $i \neq j$.

Entonces $\chi_E = \sum_{j=1}^{n} \chi_{E_j}$ y, por lo tanto, $\Gamma_E = \sum_{j=1}^{n} \Gamma_{E}$. Como $F$ es un funcional lineal, entonces $\mu_B(E) = F(\Gamma_E) = \sum_{j=1}^{n} F(\Gamma_{E_j}) = \sum_{j=1}^{n} \mu_B(\Gamma_{E_j})$.

**(2)** $\mu_B$ **es invariante bajo traslaciones.** Puesto que dadas $E, J \subset [0,1)$, si existe $a \in \mathbb{R}$ tal que $E_{tras(a)} = J$ entonces $\Gamma_E(x) = \Gamma_J(x+a)$, basta demostrar que, dada $a \in \mathbb{R}$, si $x' = x + a$ entonces $\forall g \in X$ , $F(g(x)) = F(g(x'))$[6]. Ya que en ese caso tendríamos que $\mu_B(E) = F(\Gamma_E) = F(\Gamma_J(x')) = F(\Gamma_J(x)) = \mu_B(J)$.

Sean $a \in \mathbb{R}$, $g \in X$ y $h(x) = g(x) - g(x+p)$. Hay que demostrar que $F(h) = 0$.

Sea $\{a_i\}_{i=0}^{n}$ la sucesión $a_0 = 0, a_1 = a, a_2 = 2a, \ldots, a_n = na$, entonces tenemos la siguiente serie telescópica:

$\sum_{i=0}^{n} h(x+a_i) = \sum_{i=0}^{n} g(x+a_i) - g(x+a+a_i) = g(x) - g(x+a) + g(x+a) - g(x+2a) + \ldots - g(x+na) + g(x+na) - g(x+(n+1)a) = g(x) - g(x+(n+1)a)$.

De modo que $\frac{1}{n+1}\sum_{i=0}^{n} h(x+a_i) = \frac{1}{n+1}(g(x) - g(x+(n+1)a))$. Por otro lado, como $g \in \mathcal{M}$, $\exists C \in \mathbb{R}(\forall x \in \mathbb{R}, |g(x)| < C)$. Así que $\forall x \in \mathbb{R}$, $(g(x) - g(x+(n+1)a) < 2C$. Entonces $\forall x \in \mathbb{R}$, $\frac{1}{n+1}\sum_{i=0}^{n} h(x+a_i) < \frac{2C}{n+1}$. De modo que $\sup_{0 \le x < 1} \frac{1}{n+1}\sum_{i=0}^{n} h(x+a_i) < \frac{2C}{n+1}$.

Como la sucesión que se dió en realidad está definida para cualquier $n$ y $\lim_{n\to\infty} \frac{2C}{n+1} = 0$ entonces:

$$p(h) = \inf_{s \in S}\left\{ \sup_{0 \le x < 1} \frac{1}{n+1}\sum_{i=0}^{n} h(x+a_i)\right\} = 0.$$

Por lo tanto $F(h) \le 0$, pues $F(h) \le p(h)$.

Falta ver que $0 \le F(h)$. Si tomamos la misma sucesión que antes, pero para $-h = g(x+a) - g(x)$, tendríamos que $\frac{1}{n+1}\sum_{i=0}^{n} -h(x+a_i) = \frac{1}{n+1}(-g(x) + g(x+(n+1)a))$. Y como también se tiene que $\forall x \in \mathbb{R}$, $(-g(x) + g(x+(n+1)a) <$

---

[5] Justamente esta es la razón por la que en lugar de tomar la función característica, que no es periódica, definimos a $\Gamma_E$.

[6] Desde luego que esto es un abuso de notación, pues en realidad tendríamos que definir $h_g = g \circ x'$ (se puede ver facilmente que $h_g \in X$) . Así que habría que demostrar que $\forall g \in X(F(g) = F(h_g))$. Sin embargo, la otra forma de expresarlo permite entender mejor qué es lo que se está haciendo.

$2C$, siguiendo exactamente el mismo razonamiento que antes llegamos a que $p(-h) = 0$. Por otro lado, si consideramos que:

**Observación 4** *$\forall g \in X$. $-p(-g) \leq F(g)$. Como $F$ es lineal, $F(-g) \leq p(-g) \Longrightarrow -F(g) \leq p(-g) \Rightarrow -p(-g) \leq F(g)$ y, por lo tanto, $\forall g \in X$ :*

$$-p(-g) \leq F(g) \leq p(g).$$

Así que $0 = -p(-h) \leq F(h)$. Y como ya habíamos visto que $F(h) \leq 0$, entonces $F(h) = 0$.

Finalmente, como $F$ es lineal, $F(h) = F(g(x)) - F(g(x+a)) = 0$, de modo que $F(g(x+a)) = F(g(x))$.

**(3) $\mu_B$ asigna a los intervalos su longitud.**

Lo primero que demostraremos es que $\mu_B([0,1)) = 1$. Como $\forall x(\Gamma_{[0,1)}(x) = 1)$, lo que hay que demostrar es que $F$ de la función constante $\hat{1}$ es 1. Para lo cual basta notar que $p(\hat{1}) = 1$ y $p(-\hat{1}) = -1$, pues por la Observación 4 tenemos que $-p(-\hat{1}) \leq F(\hat{1}) \leq p(\hat{1})$.

Para demostrar que $\mu_B$ asigna a cualquier intervalo $(a,b) \subset [0,1)$ su longitud, es suficiente demostrar que lo hace para los intervalos de la forma $(0,b)$, pues ya se demostró que $\mu_B$ es invariante bajo traslaciones. Sea $(0,b) \subset [0,1)$. Lo que haremos es demostrar que la medida de Banach de este conjunto coincide con la integral de Riemann de su función característica.

Sea $S_n(g\mid_{[0,1]}) = \sum_{i=1}^{n} \frac{1}{n} M_i$, con $M_i = \text{máx}\{g(x) : x \in [\frac{i-1}{n}, \frac{i}{n}]\}$ y $s_n(g\mid_{[0,1]})$-$= \sum_{i=1}^{n} \frac{1}{n} m_i$, con $m_i = \text{mín}\{g(x) : x \in [\frac{i-1}{n}, \frac{i}{n}]\}$ . Es decir, estamos tomando las sumas superior e inferior de Riemann de $g$ restringida al intervalo [0,1], con las particiones del intervalo [0,1] en una cantidad $n$ de pedazos del mismo tamaño. Las denotamos como $S_n(g)$ y $s_n(g)$, para indicar su dependencia de la función $g$, pero a partir de este momento escribiremos simplemente $S_n$ y $s_n$ para facilitar la notación. Demostraremos que $\forall g \in X$ tal que $g$ es no negativa y $\forall n \in \mathbb{N}$, $p(g) \leq S_n$ y $s_n \leq -p(-g)$. Dada $n \in \mathbb{N}$, tomamos la sucesión de reales $t_n = \{a_j : a_j = \frac{j-1}{n}\}_{j=1}^{n}$ (Nótese que $t_n = \{0, \frac{1}{n}, \ldots, \frac{n-1}{n}\}$). Afirmamos que $\forall x \in [0,1)$ . $\sum_{j=1}^{n} \frac{1}{n} g(x+a_j) \leq \sum_{i=1}^{n} \frac{1}{n} M_i$. Si $x = a_k = \frac{k-1}{n}$ entonces $\{g(x+a_j)\}_{j=1}^{n} = \{g(a_i)\}_{i=1}^{n}$, pues aunque $x + a_j > 1$, $g$ es periódica con periodo 1 y $x + a_j \equiv a_i \pmod 1$ para alguna $i \leq n$. Por otro lado, por definición de $M_i$, tenemos que $g(a_i) \leq M_i$, así que $\sum_{j=1}^{n} \frac{1}{n} g(x+a_j) \leq \sum_{i=1}^{n} \frac{1}{n} g(a_i) \leq \sum_{i=1}^{n} \frac{1}{n} M_i$. Si, en cambio, $x \in [0,1) - \{a_j\}_{j=1}^{n}$ entonces basta demostrar que existe una biyección $h$: $\{x+a_j\}_{j=1}^{n} \to \{y_i\}_{i=1}^{n}$, donde $y_i \in P_i = [\frac{i-1}{n}, \frac{i}{n}]$ y tal que $g(x+a_j) = g(h(x+a_j))$. En otras palabras, basta demostrar que en cada intervalo $P_i$ hay un sólo elemento de la sucesión $\{x+a_j\}_{j=1}^{n}$ módulo 1. Esto nos permitirá usar el mismo argumento que en el caso anterior.

Definimos la función $h$ de la siguiente manera:

$$h(x+a_j) = x + a_j - [x+a_j]$$

donde $[x+a_j] = \max\{m \in \mathbb{N} : m \leq x + a_j\}$; es decir, nos quedamos sólo con la expansión decimal de $x + a_j$. Así que claramente $h(x+a_j) \equiv x + a_i (mod 1)$ y como $g$ es periódica con periodo 1, tenemos que $g(x+a_j) = g(h(x+a_j))$. Por otro lado, $h(x+a_j) \in [0,1) \Longrightarrow \exists i \leq n(h(x+a_j) \in P_i = [\frac{i-1}{n}, \frac{i}{n}])$. Sea $i_j = \min\{i \leq n : h(x+a_j) \in P_i\}$. Falta ver que si $j \neq k$, $i_j \neq i_k$, para garantizar que en cada intervalo está la imagen de sólo un elemento de $\{x + a_j\}_{j=1}^n$. Para ello basta ver que $|h(x+a_k) - h(x+a_j)| > \frac{1}{n}$, ya que $x, y \in P_{i_j} \Longrightarrow |x - y| \leq \frac{1}{n}$. Nótese que $|h(x+a_k) - h(x+a_j)| = |a_k - a_j - ([x+a_k] + [x+a_j])|$ y que $[x+a_k]+[x+a_j]$ sólo puede ser 0, 1 ó 2. Como $a_k - a_j = \frac{k-j}{n}$, con $k - j \neq 0$, en los tres casos tenemos que $|a_k - a_j - ([x+a_k] + [x+a_j])| \geq \frac{1}{n}$. Si existe $m \leq n$ tal que $h(x+a_k)$ y $h(x+a_j)$ pertenecen a $P_m$ entonces $|h(x+a_k) - h(x+a_j)| = \frac{1}{n}$. Supongamos, sin pérdida de generalidad, que $h(x+a_k) \leq h(x+a_j)$. Puesto que la distancia entre $h(x+a_k)$ y $h(x+a_j)$ es en este caso $\frac{1}{n}$, tenemos que $h(x+a_k) \in P_{m-1}$ y $h(x+a_j) \notin P_{m-1}$, así que $i_j \neq i_k$. Eso quiere decir que $\{h(x+a_j)\}$ tiene un sólo elemento en cada intervalo $[\frac{i-1}{n}, \frac{i}{n}]$. Usando otra vez la definición de $M_i$ y el hecho de que $g(x+a_j) = g(h(x+a_j))$ tenemos que:

$$\sum_{j=1}^{n} \frac{1}{n} g(x+a_j) \leq \sum_{j=1}^{n} \frac{1}{n} g(h(x+a_j)) \leq \sum_{i=1}^{n} \frac{1}{n} M_i.$$

Una vez verificados ambos casos podemos concluir que:

$$\forall x \in [0,1), \sum_{j=1}^{n} \frac{1}{n} g(x+a_j) \leq \sum_{i=1}^{n} \frac{1}{n} M_i.$$

Por lo tanto, $\sup_{0 \leq x < 1}\{\frac{1}{n}\sum_{i=1}^{n} g(x+a_j) \leq S_n$. Como dada $n \in \mathbb{N}$ se tiene la sucesión de reales $t_n$ tal que lo anterior pasa, entonces:

$$p(g) = \inf_{s \in S} \left\{ \sup_{0 \leq x < 1} \{\frac{1}{n} \sum_{i=1}^{n} g(x+a_i) \mid a_i \in s = \{a_i\}_{i=1}^n\} \right\} \leq S_n.$$

El otro caso es totalmente análogo si $g$ es no negativa, lo que nos es de gran utilidad pues las funciones $\Gamma_E$ son no negativas. Se puede demostrar, usando las mismas sucesiones que en el caso anterior, que $\forall x \in [0,1)$ $\sum\limits_{i=1}^{n} \frac{1}{n} m_i \leq \sum_{j=1}^{n} \frac{1}{n} g(x+a_j)$. De modo que $s_n \leq \inf_{0 \leq x < 1}\{\frac{1}{n}\sum_{i=1}^{n} g(x+a_j)\}$. Como también en este caso dada $n \in \mathbb{N}$ se tiene la sucesión de reales $t_n$ tal que lo anterior pasa, entonces:

$$s_n \leq \sup_{s \in S} \left\{ \inf_{0 \leq x < 1} \{\frac{1}{n} \sum_{i=1}^{n} g(x+a_i) \mid a_i \in s = \{a_i\}_{i=1}^n\} \right\} = -p(-g).$$

Finalmente, por lo dicho hasta ahora y por la Observación 4 para cualquier función $g$ que pertenezca a X y sea positiva, dada $n \in \mathbb{N}$ :

$$s_n \leq -p(-g) \leq F(g) \leq p(g) \leq S_n.$$

En particular, como $\Gamma_{(0,b)}$ es positiva y pertenece a X, tenemos que dada $n \in \mathbb{N}$ :

$$s_n(\chi_{(0,b)}) \leq -p(-\Gamma_{(0,b)}) \leq F(\Gamma_{(0,b)}) \leq p(\Gamma_{(0,b)}) \leq S_n(\chi_{(0,b)}).$$

Afirmamos que $\lim_{n\to\infty} s_n = \int \chi_{(0,b)} = b = \lim_{n\to\infty} S_n$, de donde se tendría inmediatamente, por la desigualdad anterior y la definición de $\mu_B$, que $\mu_B((0,b)) = F(\Gamma_{(0,b)}) = |0-b| = b$. Sea $\epsilon > 0$ entonces existe $n_\epsilon \in \mathbb{N}$ tal que $\frac{1}{n_\epsilon} < \epsilon$. Nótese además que existe $m \leq n_\epsilon$ tal que $b \in [\frac{m-1}{n_\epsilon}, \frac{m}{n_\epsilon}]$, de modo que $S_{n_\epsilon} = \frac{m}{n_\epsilon}$ y $\forall n \in \mathbb{N}(n_\epsilon \leq n \rightarrow \frac{m-1}{n_\epsilon} \leq S_n \leq \frac{m}{n_\epsilon})$, pues de otro modo $b \notin [\frac{m-1}{n_\epsilon}, \frac{m}{n_\epsilon}]$. Por lo tanto, $\forall n \in \mathbb{N}(n_\epsilon \leq n \rightarrow |S_n - b| \leq \frac{1}{n_\epsilon} < \epsilon)$. Esto quiere decir que $\lim_{n\to\infty} S_n = b$. Para el otro caso se tiene que $s_{n_\epsilon} = \frac{m-1}{n_\epsilon}$ y también que $\forall n \in \mathbb{N}(n_\epsilon \leq n \rightarrow \frac{m-1}{n_\epsilon} \leq s_n \leq \frac{m}{n_\epsilon})$. Así que $\forall n \in \mathbb{N}(n_\epsilon \leq n \rightarrow |s_n - b| \leq \frac{1}{n} < \epsilon)$. Por lo tanto, $\lim_{n\to\infty} s_n = b$ y $\lim_{n\to\infty} s_n = \lim_{n\to\infty} S_n$. ∎

Esta medida definida para el intervalo [0,1) se puede generalizar muy fácilmente al caso de la potencia de toda la recta real. La idea es intersectar el conjunto con todos los intervalos de tipo $[z, z+1]$ con $z \in \mathbb{Z}$ y trasladar cada intersección al intervalo $[0,1)$ para tomar la medida de Banach y después sumar las medidas de todas esas intersecciones trasladadas, como se puede ver con el siguiente corolario:

**Corolario 7 (Medida de Banach para la recta real)** *Existe* $\mu : P(\mathbb{R}^1) \rightarrow [0,\infty]$ *tal que:*

(1) A los intervalos les asigna su longitud. $\forall A = (a,b) \subset \mathbb{R}^1$, $\mu(A) = |a-b|$.
(2) Es invariante bajo traslaciones. $\forall A \subset \mathbb{R}^1$ y $\forall x \in \mathbb{R}$, $\mu(A_{trasx}) = \mu(A)$.
(3) Es finito aditiva. $\forall \{A_i\}_{i=1}^n [A_i \subset \mathbb{R}^1 \wedge (i \neq j \Rightarrow A_i \cap A_j = \emptyset) \Longrightarrow \mu(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n \mu(A_i)]$

**Demostración.** Sea $A \subset \mathbb{R}^1$. Definimos la función $\mu : P(\mathbb{R}^1) \rightarrow [0,\infty]$ de la siguiente manera:

(1) Si $A$ no está acotado, entonces $\mu(A) = \infty$[7].

(2) Si $A$ está acotado, sea $\{z_i\}_{i=1}^\infty$ una numeración de $\mathbb{Z}$ e $I_i = [z_i, z_i+1)$. Como $A$ es acotado existe $n \in \mathbb{N}$ tal que para toda $m > n$, $A \cap [z_m, z_m+1) = \emptyset$. Entonces dada $l \leq n$ tomamos $A_l = A \cap [z_l, z_l+1)$. Claramente $(A_l)_{tras(-z_l)} = \{a - z_l : a \in A_l\} \subset [0,1)$. Así que $\mu(A) = \sum_{l=1}^n \mu_B(A_l)_{tras(-z_l)}$.

[7] Esto lo podemos hacer porque no nos interesa que la medida sea $\sigma-$*aditiva*.

Es fácil ver que $\mu$, definida de esta manera, coincide con la longitud cuando se trata de intervalos, que es invariante bajo traslaciones y finito-aditiva. Aquí sólo demostraremos la tercera condición, por el papel que juega en nuestra discusión. Sea $A = \bigcup_{i=1}^{n} B_i$, tal que $B_i \cap B_j = \emptyset$ si $i \neq j$.

(1) $A$ es no acotado. Como $A$ es unión finita de conjuntos disjuntos, entonces existe $m < n$ tal que $B_m$ es no acotado. Así que por la forma en que se definió $\mu$ tenemos que $\mu(B_m) = \infty$. De modo que $\sum_{i=1}^{n} \mu(B_i) = \infty = A$.

(2) $A$ es acotado. Así que existe $k \in \mathbb{N}$ tal que para $m > k$, $A \cap [z_m, z_m + 1] = \emptyset$. Nótese que $\forall l \leq k$, $A_l = (\bigcup_{i=1}^{n} B_i) \cap [z_l, z_l + 1] = \bigcup_{i=1}^{n} (B_i \cap [z_l, z_l + 1])$. Además si denotamos con $B_{il}$ al conjunto $B_i \cap [z_l, z_l + 1]$ tenemos que $A_l = \bigcup_{i=1}^{n} (B_{il})$. Claramente $B_{il} \cap B_{jl} = \emptyset$, si $i \neq j$, pues por hipótesis $B_i \cap B_j = \emptyset$. De modo que usando la finito aditividad de $\mu_B$ tendríamos que $\mu_B(A_l)_{tras(-z_l)} = \sum_{i=1}^{n} \mu_B(B_{il})_{tras(-z_l)}$ y, por lo tanto, $\mu(A) = \sum_{l=1}^{k} \mu_B(A_l)_{tras(-z_l)} = \sum_{l=1}^{k} \sum_{i=1}^{n} \mu_B(B_{il})_{tras(-z_l)} =$

$\sum_{i=1}^{n} \sum_{l=1}^{k} \mu_B(B_{il})_{tras(-z_l)} = \sum_{i=1}^{n} \mu(B_i)$, pues $B_{il} = B_i \cap [z_l, z_l + 1]$. ■

La medida de Banach nos hace pensar que la condición de que la medida sea invariante bajo traslaciones no es la responsable de que falle lo que hemos llamado la hipótesis de la medida, pues abandonando sólo la condición de la $\sigma-$aditividad y preservando la invarianza bajo traslaciones se vio que la versión correspondiente (más débil) de la hipótesis de la medida se cumplía para $\mathbb{R}^1$, no sólo sin necesidad de negar que exista un buen orden para los reales, sino incluso haciendo un uso explícito del axioma de elección, a través del teorema de Hahn-Banach[8]. Más aún, este resultado señala a la $\sigma-$aditividad como responsable, en el sentido de ser condición necesaria, de la negación de la hipótesis de la medida .

## 2.2. Teorema de Banach-Tarski

Aunque la medida de Banach nos hace pensar que la invarianza bajo traslaciones no es responsable de la negación de la hipótesis de la medida, tenemos que considerarlo con más cuidado. Hasta ahora, por simplicidad, hemos planteado formalmente el problema de la medida para el caso de $\mathbb{R}^1$, pero en realidad puede pensarse para el caso general de $\mathbb{R}^n$.

[8] Como $ZF + Hahn - Banach \vdash HM_{-4}$ y $(AE \Longrightarrow Hahn - Banach)$, también se tiene que $ZFE \vdash HM_{-4}$.

Lo que nos estaríamos preguntando es si es posible generalizar la noción tamaño geométrico del $n-$producto cartesiano de intervalos[9]. Es decir, quisiéramos saber si existe una función definida en toda $\wp(\mathbb{R}^n)$ tal que: (1) coincida con el tamaño geométrico de los "intervalos" de $\mathbb{R}^n$, (2) sea invariante bajo las transformaciones rígidas de $\mathbb{R}^n$,(3) sea finito-aditiva, (4) sea $\sigma - aditiva$. Uno podría suponer que los resultados respecto al problema de la medida para $\mathbb{R}^1$ se pueden generalizar a $\mathbb{R}^n$. Sin embargo la medida de Banach junto con el famosísimo teorema de Banach-Tarski, que a continuación enunciaremos, hacen ver que algunos resultados sí dependen de la dimensión de que se trate.

**Definición 6** *Sea $G_3$ el grupo de isometrías en $\mathbb{R}^3$ y $A, B \subseteq \mathbb{R}^3$. $A$ es congruente con $B$ ($A \cong_{G_3} B$) si y sólo si existe $g \in G_3$ tal que $g(A) = B$ .*

**Teorema 8 (Banach-Tarski)** *Sea $B^3 = \{x \in \mathbb{R}^3 : |x| \leq 1\}$ entonces existen tres particiones finitas de $B^3$, $P = \{P_i\}_{i=1}^r, Q = \{Q_i\}_{i=1}^s$ y $T = \{T_i\}_{i=1}^{r+s}$ tales que $P_i \cong_{G_3} T_i$ y $Q_i \cong_{G_3} T_{i+r}$.*

Este teorema, que suele conocerse como la paradoja de Banach -Tarski, nos dice que es "posible" partir una bola unitaria en una cantidad finita de pedazos, que rotados y trasladados[10], generan dos copias de la bola original. Este resultado parece ser bastante contraintuitivo, sobre todo si por pedazos tenemos en mente poliedros o conjuntos conexos que es lo que correspondería a los pedazos materiales en los que podemos partir los objetos de nuestra realidad física. Sin embargo, la realidad matemática no siempre puede ser modelada en nuestra realidad física, como bien lo demuestra la construcción del conjunto de Cantor[11]. Así que el "posible" de nuestro teorema es simplemente teórico. No obstante, aún dentro de un contexto matemático este resultado sorprende, pues por analogía con el caso de $\mathbb{R}^1$ se esperaría que al abandonar la $\sigma - aditividad$ todos los conjuntos de $\mathbb{R}^3$ fueran medibles. Esto no sucede por el teorema de Banach-Tarski:

**Teorema 9 (Medida-Banach-Tarski)** *No existe $\mu : \wp(\mathbb{R}^3) \longrightarrow [0, \infty]$ tal que:*

(1) $\forall I = [a_1 . b_1] \times [a_2, b_2] \times [a_3, b_3] \subseteq \mathbb{R}^3, \mu(I) = \prod_{i=1}^{3} |a_i - b_i|$.
(2) $\forall A, B \subset \mathbb{R}^3 (A \cong_{G_3} B \Longrightarrow \mu(A) = \mu(B))$.
(3) $\forall A, B \subseteq \mathbb{R}^3 (A \cap B = \emptyset \Longrightarrow \mu(A \cup B) = \mu(A) + \mu(B)$.

**Demostración.** Supongamos que existe $\mu : \wp(\mathbb{R}^3) \longrightarrow [0, \infty]$ que cumple las condiciones (1) a (3). Tomemos las tres particiones finitas $P$, $Q$ y $T$ de $B^3$ dadas por el teorema de Banach-Tarski. Como se trata de particiones de $B^3$ tenemos que $\bigcup_{i=1}^{r} P_i = \bigcup_{i=1}^{s} Q_i = \bigcup_{i=1}^{r+s} T_i = B^3$. Así que $\mu(\bigcup_{i=1}^{r} P_i) = \mu(\bigcup_{i=1}^{s} Q_i) = \mu(\bigcup_{i=1}^{r+s} T_i) =$

[9] En $n = 2$ sería la superficie de un rectángulo, para $n = 3$ el volumen de un paralelepípedo, etc.

[10] Aunque en realidad sólo se necesitan rotaciones, como se verá más adelante.

[11] El conjunto de Cantor es una abstracción matemática que tampoco tiene realidad física, en el sentido de que su construcción no se puede llevar a cabo físicamente.

$\mu(B^3)$. Por otro lado, como $\mu$ es finito-aditiva y ($P$ partición $\Longrightarrow P_i \cap P_j = \emptyset$ si $i \neq j$), entonces $\sum_{i=1}^{r} \mu(P_i) = \mu(\bigcup_{i=1}^{r} P_i) = \mu(\bigcup_{i=1}^{r+s} T_i) = \sum_{i=1}^{r+s} \mu(T_i)$.

Además como $\forall i(1 \leq i \leq r)\ P_i \cong_{G_3} T_i$, por la invarianza bajo transformaciones rígidas (condición 2), $\forall i(1 \leq i \leq r)$, $\mu(P_i) = \mu(T_i)$. De modo que $\forall i(r+1 \leq i \leq r+s)$, $\mu(T_i) = 0$ y $\forall i(1 \leq i \leq s)$, $\mu(Q_i) = 0$, ya que $Q_i \cong_{G_3} T_{i+r}$. Así que $\mu(\bigcup_{i=1}^{s} Q_i) = \sum_{i=1}^{s} \mu(Q_i) = 0$ y, por lo tanto, $\mu(B^3) = 0$. Pero eso es imposible, ya que $A = \left[-\frac{1}{2}, \frac{1}{2}\right] \times \left[-\frac{1}{2}, \frac{1}{2}\right] \times \left[-\frac{1}{2}, \frac{1}{2}\right] \subset B^3$ y, por las condiciones (1) y (3), $\mu(A) = 1 \leq \mu(A) + \mu(B^3 \setminus A) = \mu(B^3)$. Por lo tanto, no existe $\mu : \wp(\mathbb{R}^3) \longrightarrow [0, \infty]$ que cumpla las condiciones dadas. ∎

De esto podemos concluir que la invarianza bajo transformaciones rígidas de $\mathbb{R}^3$, a diferencia de la invarianza bajo traslaciones sobre la recta real, tiene responsabilidad en la negación de la hipótesis de la medida finito-aditiva en dicha dimensión ($HM^3_{-4}$). En cambio, con relación a la $\sigma-$*aditividad* podemos decir que para el caso de $\mathbb{R}^3$ no se ganaría nada, en términos del problema de la medida, si dicha propiedad se abandona.

### 2.2.1. Análogo de Banach-Tarski en $\mathbb{R}^2$ y $\mathbb{R}^1$.

Si los pedazos en los que se descompone la esfera no son poliedros y, por el teorema que acabamos de ver, dada cualquier medida que extienda la noción de volumen sabemos que al menos algunos de ellos no son medibles, entonces ¿qué clase de objetos extraños son? Esto nos recuerda al conjunto de Vitali, sobre el cual nos hacíamos la misma pregunta. De modo que no debe sorprendernos que la demostración del teorema de Banach-Tarski tenga su versión análoga en $\mathbb{R}^1$y $\mathbb{R}^2$ (en donde, por supuesto, las particiones no serán finitas sino numerables). Más aún, veremos que los pedazos en los que se parte la esfera están conformados por un conjunto que, como Vitali, se obtiene mediante una cantidad $c$ de elecciones. En ese sentido, aún dentro de un marco puramente teórico no es del todo adecuado decir que es "posible" dar la partición finita de la esfera, pues el conjunto con el que se construyen los pedazos no se puede dar explícitamente. Por lo que sería más adecuado decir simplemente que dicha partición existe[12].

Desde luego que, puesto que los resultados respecto a la hipótesis de la medida finito-aditiva ($HM_{-4}$) dependen de la dimensión, no es posible decir tampoco que para el caso de $\mathbb{R}^3$ el único responsable es el axioma de elección. De hecho, veremos que la demostración del teorema de Banach-Tarski no sólo hace un uso importante de la hipótesis que sostiene que existe un buen orden para los reales, sino que también se basa en algunas propiedades del grupo de isometrías de $\mathbb{R}^3$, que no comparten el grupo de isometrías de $\mathbb{R}^1$ y $\mathbb{R}^2$.

[12]Lo mismo sucede con el buen orden de los reales, ya que muchas veces su existencia se expresa como: "los reales son bien ordenables". Esto hace pensar que es posible dar el orden explícitamente, cosa que no es cierta, pues de serlo se podría resolver el problema del continuo.

Sea $S^1 = \{x \in \mathbb{R}^2 : |x| = 1\}$ y $R_{\mathbb{Q}}$ el grupo de rotaciones racionales del plano módulo $2\pi$, con centro en el origen y con ángulo de rotación entre $-\pi$ y $\pi$ radianes. Podemos exponer el análogo del teorema de Banach-Tarski para el caso de $\mathbb{R}^2$ en los mismos términos que para el caso de $\mathbb{R}^3$:

**Teorema 10** *Existen tres particiones numerables de $S^1$, $P = \{P_i\}_{i\in\mathbb{N}}$, $Q = \{Q_i\}_{i\in\mathbb{N}}$ y $T = \{T_i\}_{i\in\mathbb{N}}$ tales que $P_i \cong_{R_{\mathbb{Q}}} T_{2i}$ y $Q_i \cong_{R_{\mathbb{Q}}} T_{2i+1}$.*

En el caso de $\mathbb{R}^3$ se dió el teorema de Banach-Tarski en estos términos (es decir, con tres particiones) para facilitar la demostración de la negación de $HM^3_{-,1}$. Sin embargo, como la idea de introducir los análogos de Banach-Tarski en $\mathbb{R}^2$ y $\mathbb{R}^1$ es clarificar el extraño resultado de $\mathbb{R}^3$, trabajaremos con una versión más intuitiva, aunque no por ello menos rigurosa, del teorema anterior:

**Teorema 11** *Existe una familia numerable $\{A_1, ..., A_n, ..., B_1, ..., B_n, ...\}$ de subconjuntos de $S^1$ ajenos dos a dos, tal que $S^1 = (\bigcup_{i\in\mathbb{N}} A_i) \cup (\bigcup_{j\in\mathbb{N}} B_j)$, $S^1 = \bigcup_{i\in\mathbb{N}} \sigma_i(A_i)$ y $S^1 = \bigcup_{j\in\mathbb{N}} \sigma_j(B_j)$ con $\sigma_i$ y $\sigma_j$ elementos de $R_{\mathbb{Q}}$.*

En otras palabras, este teorema nos dice que el círculo unitario se "puede" partir en una cantidad numerable de pedazos a partir de los cuales se pueden construir, bajo rotaciones, dos círculos del mismo tamaño. Además, si usamos la proyección radial tenemos el mismo resultado en el disco unitario sin el origen[13]. Nótese que para el caso de la esfera sólida, también por la proyección radial, bastaría con dar las particiones de la superficie esférica[14].

**Demostración.** Definimos la siguiente relación sobre $S^1$: $\forall x, y \in S^1, x \sim_{R_Q} y$ ssi existe $\rho \in R_{\mathbb{Q}}$ tal que $\rho(x) = y$. Como $R_{\mathbb{Q}}$ es grupo, contiene a la identidad, a los inversos de todos sus elementos y es cerrado bajo composición, por lo que $\sim_{R_Q}$es una relación de equivalencia. Si recordamos además que, dado un grupo G que actúa sobre un conjunto A[15] y dado $x$ un elemento de A, el conjunto $O_G(x) = \{g(x) : g \in G\}$ se conoce como una $G-$*órbita* de A, entonces tenemos que:

$$x \sim_{R_Q} y \;\text{ ssi }\; O_{R_Q}(x) = O_{R_Q}(y)$$

De modo que las clases de equivalencia de la partición $P$ que $\sim_{R_Q}$induce sobre $S^1$ son las órbitas bajo las rotaciones racionales con centro en el origen. Por otro lado, si $\theta_\rho$ denota el ángulo de rotación de $\rho$ entonces $R_{\mathbb{Q}} = \{\rho : \exists q \in \mathbb{Q} \cap (-1, 1)$ tal que $\theta_\rho = q\pi$ radianes$\}$. Como $\mathbb{Q}\cap(-1, 1)$ es numerable, $R_{\mathbb{Q}}$ también lo es. Sea $\{\rho_i\}_{i\in\mathbb{N}}$ una numeración de $R_{\mathbb{Q}}$. Eso quiere decir que cada clase de equivalencia es

[13] Para el disco unitario tomamos la familia numerable $\{A'_1, ..., A'_n, ..., B'_1, ..., B'_n ...\}$ donde $A'_i = \bigcup_{0<\alpha\leq 1} \alpha A_i$ y si pensamos en el plano con coordenadas polares $\alpha A_i = \{(\alpha r, \theta) : (r, \emptyset) \in A_i\}$

[14] En realidad se tendría, como para el caso de $\mathbb{R}^2$, una partición sobre la esfera sin el origen. En el Apéndice C se explica qué hay que hacer para que se tenga sobre toda la esfera.

[15] Recuérdese también que un grupo G actúa sobre un conjunto A si $\forall g \in G$, $g$ es una función de A en A y se cumple que $\forall g, h \in G; \forall x \in A\; g(h(x)) = gh(x)$ y $1(x) = x$, con 1 el elemento neutro de G.

numerable, pues por definición $O_{R_{\mathbb{Q}}}(x) = \{\ \rho_i(x) : \rho_i \in R_{\mathbb{Q}}\}$, y como $|S^1| = c$, entonces tenemos que en la partición $P$ hay $c$ clases de equivalencia. Por el axioma de elección tenemos $M$ un conjunto de representantes de las clases de equivalencia. Por otro lado, como las clases de equivalencia son las órbitas bajo $R_{\mathbb{Q}}$, que no tienen puntos fijos, y $M$ sólo tiene un miembro de cada órbita, entonces $\rho_i(M) \cap \rho_j(M) = \emptyset$ si $i \neq j$. Pues si existen $x, y \in M$ tales que $\rho_i(x) = \rho_j(y)$ entonces $x = \rho_i^{-1}\rho_j(y)$ y $x \in O_{R_{\mathbb{Q}}}(y)$. Pero como $y \in M$ y $M$ sólo tiene un elemento de cada órbita, entonces $x = y$. Eso quiere decir que $x = \rho_i^{-1}\rho_j(x)$. Así que $\rho_i^{-1}\rho_j$ es la identidad (ya que es la única rotación de $R_{\mathbb{Q}}$ que deja puntos fijos de $S^1$) y, por lo tanto, $\rho_i = \rho_j$ (es decir, $i = j$). Por otro lado, $S^1 = \bigcup_{x\in S^1} O_{R_{\mathbb{Q}}}(x) = \bigcup_{i\in\mathbb{N}} \rho_i[M]$. ya que $\forall x \in S^1\ \exists y \in M\ (x \in O_{R_{\mathbb{Q}}}(y))$. Es decir, $\forall x \in S^1 \exists \rho_i \in R_{\mathbb{Q}}\ (\rho_i(y) = x)$ y, por lo tanto, $x \in \bigcup_{i\in\mathbb{N}} \rho_i[M]$.

La otra contención es trivial, pues $M \subset S^1$ y $S^1$ es cerrado bajo $R_{\mathbb{Q}}$. Así que $P' = \{\rho_i(M) : \rho_i \in R_{\mathbb{Q}}\}$ es una partición numerable de $S^1$. Si ahora tomamos $A_i = \rho_{2i}(M)$ y $B_j = \rho_{2j+1}(M)$, entonces $\{A_1, ..., A_n, ..., B_1, ..., B_n, ...\}$ $= \{\rho_i(M) : \rho_i \in R_{\mathbb{Q}}\} = P'$ es una familia de conjuntos ajenos dos a dos de $S^1$, ya que P′ es partición. Si tomamos $\sigma_i = \rho_i\rho_{2i}^{-1}$, como $R_{\mathbb{Q}}$ es grupo, $\sigma_i$ pertenece a $R_{\mathbb{Q}}$. Nótese que $\sigma_i(A_i) = \sigma_i\rho_{2i}(M) = \rho_i(M)$, de modo que $\bigcup_{i\in\mathbb{N}} \sigma_i(A_i)$ $= \bigcup_{i\in\mathbb{N}} \rho_i(M) = S^1$. De manera análoga, si tomamos $\sigma_j = \rho_j\rho_{2j}^{-1}$ tendríamos que $\bigcup_{j\in\mathbb{N}} \sigma_j(B_j) = S^1$. Finalmente, ya habíamos visto que $\{A_1, ..., A_n, ..., B_1, ..., B_n, ...\}$ $= \{\rho_i(M) : \rho_i \in R_{\mathbb{Q}}\} = P'$, así que $\bigcup_{i\in\mathbb{N}} \rho_i(M) = \bigcup_{i\in\mathbb{N}} A_i \cup \bigcup_{j\in\mathbb{N}} B_j$. ■

El análogo de Banach-Tarski en $\mathbb{R}^1$ se desprende como corolario, mediante las siguientes definiciones:

**Definición 7** *Una traslación módulo 1 es una función $t_q : [0,1) \to [0,1)$ tal que $\forall x \in [0,1)$*

$$t_q(x) = x + q - [x + q]$$

*con $q \in \mathbb{Q} \cap (-1,1)$ y $[x+q] = \max\{m \in \mathbb{Z} : m \leq x + q\}$*[16].

**Observación 5** *Es fácil ver que el conjunto de traslaciones módulo 1, que denotaremos con $T_{\mathbb{Q}}$, es un grupo, excepto quizás por la existencia de inversos, que será lo único que demostraremos:*

Sea $t_q \in T_{\mathbb{Q}}$. Afirmamos que $t_q^{-1} = t_{-q}$. En otras palabras que $\forall x \in [0,1)$ , $t_{-q}t_q(x) = x$. Tenemos dos casos:

(1) $[x+q] = 1$. Entonces $t_q(x) = x + q - 1$ y $t_{-q}(t_q(x)) = x - 1 - [x - 1]$ . Pero $-1 \leq (x-1) < 0$, ya que $x \in [0,1)$. Así que $[x-1] = -1$ y $t_{-q}(t_q(x)) = x$.

(2) $[x+q] = 0$. Es trivial.

[16] Por ejemplo, es fácil ver que lo que definimos en la demostración del Teorema 4 como $V_q$ no es sino $V_{t_q}$. Es decir, $V$ trasladado por $q$ módulo 1.

Como corolario del Teorema 11 tenemos que:

**Corolario 12** *Existe una familia numerable* $\{A_1, ..., A_n, ..., B_1, ..., B_n, ...\}$ *de subconjuntos de* $I = [0,1)$ *ajenos dos a dos tal que* $I = \bigcup_{i\in\mathbb{N}} A_i \cup \bigcup_{j\in\mathbb{N}} B_j$ , $I = \bigcup_{i\in\mathbb{N}} \sigma_i(A_i)$ *y* $I = \bigcup_{j\in\mathbb{N}} \sigma_j(B_j)$ *con* $\sigma_i$ *y* $\sigma_j$ *elementos de* $T_{\mathbb{Q}}$.

**Demostración.** Basta dar una biyección entre $S^2$ y $[0,1)$ y un isomorfismo entre $R_{\mathbb{Q}}$ y $T_{\mathbb{Q}}$. Para la biyección tomamos $f : S^2 \longrightarrow [0,1)$ tal que $f((1,\theta)) = \frac{\theta}{2\pi}$ y si tomamos $g : R_{\mathbb{Q}} \longrightarrow \mathbb{Q} \cap (-1,1)$ tal que $g(\rho) = \frac{\theta_\rho}{2\pi}$, tenemos el isomorfismo $F : R_{\mathbb{Q}} \longrightarrow T_{\mathbb{Q}}$ dado por $F(\rho) = t_{g(\rho)}$. ∎

Este resultado no debería sorprendernos si es que nos sentimos tranquilos con el conjunto de Vitali, pues la demostración del Teorema 11 está basado en una generalización de la partición del intervalo [0,1) en una cantidad numerable de clases de equivalencia a partir del conjunto de Vitali. En ese sentido, $M$ es la versión correspondiente a Vitali en $\mathbb{R}^2$, pues ambos son conjuntos de representantes de $c$ clases de equivalencia. De hecho, para obtener a $M$ no es necesario el axioma de elección con toda su fuerza, sino que basta, como para Vitali, suponer que hay un buen orden para los reales. Es decir, si suponemos $BO_{\mathbb{R}}$ podemos pensar en $M$ como $V^2 = \{x \in \mathbb{R}^2 : x = \min_{<_{\mathbb{R}^2}} O_{R_Q}\}$[17]. Así que, aunque nosotros confirmamos que ZFE $\vdash$ Teorema 11, también se tiene que $ZF + BO_{\mathbb{R}} \vdash$ Teorema 11. Sin embargo, Vítali y $M$ no son análogos simplemente por ser conjuntos de representantes de $c$ clases de equivalencia, sino también, porque las clases de equivalencia de donde se obtienen son análogas.

Recuérdese que las clases de equivalencia de las que surge Vitali están dadas por la relación $\sim_v$ y nótese que $\forall x, y \in [0,1)$ :

$$x \sim_v y \iff \exists q \in \mathbb{Q}(y - x = q) \iff \exists q \in \mathbb{Q} \cap (-1,1)(y = q + x)$$

$$\iff {}^{18}\exists \tau \in T_{\mathbb{Q}}\ (\tau(x) = y) \iff y \in T_{\mathbb{Q}}(x).$$

Esto quiere decir que las clases de equivalencia inducidas por $\sim_v$ no son sino las $T_{\mathbb{Q}}$ órbitas de [0,1). De modo que la construcción de Vitali y de la partición numerable que induce sobre [0,1) podría plantearse en términos del grupo de isometrías en $\mathbb{R}^1$.

---

[17] Nótese que si un conjunto $A$ tiene un buen orden $<_A$, entonces $A \times A$ también lo tiene. El orden inducido es el orden lexicográfico: $(a,b) <_{A\times A} (c,d) \Leftrightarrow (a <_A b) \vee (a = b \wedge c <_A d)$. Por otro lado, cabe señalar que si quisiéramos ser más formales deberíamos escribir $O_{R_Q}(y)$ para alguna $y \in S^1$, pero lo dejamos como $O_{R_Q}$ para simplifcar la notación.

[18] Veamos el regreso de este bicondicional que no parece nada obvio. Sea $\tau \in T_{\mathbb{Q}}$ $(\tau(x) = y)$. Entonces existe $q' \in \mathbb{Q} \cap (-1,1)$ tal que $y = x + q' - [x + q']$ . Si $[x + q'] = 0$ entonces tomamos $q = q'$. Si $[x + q'] = 1$ entonces tomamos $q = q' - 1$. En cuyo caso sólo falta ver que $q \in \mathbb{Q} \cap (-1,1)$. Como $q' \in \mathbb{Q} \cap (-1,1)$, si $0 < q'$ entonces $-1 < q' - 1 < q' < 1$, así que $q \in \mathbb{Q} \cap (-1,1)$. El caso $q' \leq 0$ no se puede dar, pues $q' \leq 0 \Rightarrow (q' - 1 \leq -1) \Rightarrow (x + (q' - 1) \leq x - 1) \Rightarrow (y \leq x - 1)$ y como $0 < x < 1$, entonces $q' \leq 0 \Longrightarrow y < 0$!(pues $0 < y < 1$).

En la Sección 1.2 el conjunto de Vitali se había definido como $V = \{x : x = \min_{<_{\mathbb{R}}} C_i\}$, pero acabamos de ver que $F_{\sim_r} = \{C_i\}_{i\in I} = \{O_{T_{\mathbb{Q}}}(x) : x \in [0,1)\}$. Así que el conjunto de Vitali también se puede describir como $V = \{x : x = \min_{<_{\mathbb{R}}} O_{T_{\mathbb{Q}}}\}$.

Dado que la discusión del problema de la medida se ha planteado en términos de $\mathbb{R}^1$ no nos detendremos a dar la demostración del teorema de Banach-Tarski con lujo de detalles. Sin embargo, a partir de las demostraciones de los análogos de Banach-Tarski en $\mathbb{R}^1$ y $\mathbb{R}^2$ podremos dar un esbozo general de la demostración del teorema para el caso de $\mathbb{R}^3$, con lo que quedará claro que:

1) Los pedazos en los que se parte la esfera no son sino rotaciones de la versión del conjunto de Vitali en $\mathbb{R}^3$.

2) Para este caso también basta suponer que hay un buen orden para los reales.

3) La demostración está basada en propiedades que tienen que ver con teoría de grupos, en particular con el grupo de isometrías. Con lo que se verá por qué Vitali no sirve para negar la hipótesis de la medida finito-aditiva en $\mathbb{R}^1$, pero su análogo sí para negarla en $\mathbb{R}^3$.

### 2.2.2. Demostración de Banach-Tarski: análogo de Vitali en $\mathbb{R}^3$.

Hemos señalado ya que la demostración del teorema de Banach-Tarski se basa en una partición finita de la superficie esférica (es decir, en un partición finita de $S^2 = \{x \in \mathbb{R}^3 : |x| = 1\}$) y que dicha partición se construye a partir del análogo de Vitali en $\mathbb{R}^3$. Para ello se toma, como para los casos de $\mathbb{R}^1$ y $\mathbb{R}^2$, un subgrupo particular de isometrías que es numerable. De hecho, se trata de un subgrupo de rotaciones generado a partir de dos rotaciones particulares $\tau$ y $\sigma$[19]. De modo que, si $G_{\tau\sigma}$ denota a dicho grupo y $SO_3$ al grupo de isometrías en $\mathbb{R}^3$, lo que estamos diciendo es que $G_{\tau\sigma}$ es subgrupo de $SO_3$ y que $|G_{\tau\sigma}| = \aleph_0$. Sea $\{\rho_i\}_{i\in\mathbb{N}}$ una numeración de $G_{\tau\sigma}$. Como $G_{\tau\sigma}$ es un grupo que actúa sobre $S^2$ podemos pensar, una vez más, en las $G_{\tau\sigma}$−órbitas como clases de equivalencia sobre $S^2$ y $P = \{O_{G_{\tau\sigma}}(x) : x \in S^2\}$ como la partición correspondiente. Nótese que $P$ tiene $c$ clases de equivalencia, ya que cada órbita es a lo más numerable (pues $G_{\tau\sigma}$ lo es) y $|\bigcup P| = |S^2| = c$.

Lo siguiente sería obtener un conjunto $M$ de representantes de cada clase de equivalencia. Pare ello nos servimos, una vez más, del axioma de elección en su versión $BO_{\mathbb{R}}$ y puesto que, si existe un buen orden para los reales $<_{\mathbb{R}}$, se puede obtener un buen orden $<_{\mathbb{R}^3}$ para $\mathbb{R}^3$, entonces podemos tomar $M = \{x \in S^2 : x = \min_{<_{\mathbb{R}^3}} O_{G_{\tau\sigma}}\}$. Como $G_{\tau\sigma}$ es numerable entonces $P' = \{\rho_i[M] : \rho_i \in G_{\tau\sigma}\}$ es una familia numerable.

[19] Wagon [24] (p.15) da las rotaciones $\tau$ y $\sigma$ explícitamente.

Hasta aquí la construcción es completamente análoga al caso de $\mathbb{R}^2$. Lo siguiente sería tratar de agrupar esos pedazos de manera que en realidad se tenga una cantidad finita. Sin embargo, antes debemos enfrentar otra dificultad: los elementos de $P'$ no son disjuntos y, por lo tanto, $P'$ no es una partición de $S^2$. Esto es porque las rotaciones de $G_{\tau\sigma}$, a diferencia de los elementos de $T_{\mathbb{Q}}$ y $R_{\mathbb{Q}}$, dejan fijos dos puntos fijos. Específicamente dejan fijos a los puntos de la intersección del eje de rotación con la superficie esférica. Esto es problemático, pues dadas $\rho_s \in G_{\tau\sigma}$ distinta de la identidad y $x \in S^2$ tales que $\rho_s(x) = x$, como $x \in S^2$, existen $y \in M$ y $\rho_j \in G_{\tau\sigma}$ tales que $\rho_j(y) = x$. De modo que $\rho_s\rho_j(y) = x$ y como $G_{\tau\sigma}$ es grupo, entonces existe $\rho_i \in G_{\tau\sigma}$ tal que $\rho_i = \rho_s\rho_j$. Pero además $\rho_i \neq \rho_j$, ya que $\rho_s$ distinta de la identidad. De modo que, $\rho_i[M] \cap \rho_j[M] \neq \emptyset$ con $\rho_i \neq \rho_j$.

Esto se puede resolver si tomamos $S^2 \setminus D$ , con $D = \{x \in \mathbb{R}^3 : \exists \rho_i \in G_{\tau\sigma}$ tal que $\rho_i(x) = x\}$. Nótese que como $G_{\tau\sigma}$ es numerable y cada rotación deja dos puntos fijos, $D$ es numerable. En el (Apéndice C)se explicará porqué una partición de $S^2 \setminus D$ induce una partición adecuada en $S^2$. Por lo pronto, hay que repetir con $S^2 \setminus D$ lo que se hizo con $S^2$. Es decir, $\forall x \in S^2 \setminus D$ sea $O'_{G_{\tau\sigma}}(x) = O_{G_{\tau\sigma}}(x) \cap S^2 \setminus D$[20].

Sea $Q = \{ O'_{G_{\tau\sigma}}(x) : x \in S^2 \setminus D\}$. Como $P$ es una partición de $S^2$ y $Q = P \cap S^2 \setminus D$, tenemos que $Q$ es una partición de $S^2 \setminus D$. Así que tomaremos el conjunto de representantes que induce la existencia de un buen orden de los reales. Es decir, tomamos $V^3 = \{x \in S^2 \setminus D : x = \min_{<_{\mathbb{R}^3}} O'_{G_{\tau\sigma}} \}$. De modo que, en este caso, $P' = \{\rho_i[V^3] : \rho_i \in G_{\tau\sigma}\}$ sí es una partición de $S^2 \setminus D$, aunque todavía nos queda el problema de que es numerable[21].

Por otro lado, por la forma en la que se eligen $\tau$ y $\sigma$, se puede ver que $G_{\tau\sigma}$ tiene cuatro subconjuntos ajenos $R_1, R_2, R_3$ y $R_4$ tales que $R_1 \cup \tau R_3 = G_{\tau\sigma}$ y $R_2 \cup \sigma R_4 = G_{\tau\sigma}$. Con esta propiedad se completa la demostración, ya que se toman $A_1 = \bigcup_{\rho_i \in R_1} \rho_i[V^3], A_2 = \bigcup_{\rho_i \in R_3} \rho_i[V^3], B_1 = \bigcup_{\rho_i \in R_2} \rho_i[V^3]$ y $B_2 = \bigcup_{\rho_i \in R_4} \rho_i[V^3]$ que claramente son ajenos, pues están formados por los elementos de la partición $P'$. Además como $R_1 \cup \tau R_3 = G_{\tau\sigma}$ entonces $A_1 \cup \tau A_3 = \bigcup_{\rho_i \in G_{\tau\sigma}} \rho_i[V^3] = \bigcup P' = S^2 \setminus D$, pues $P'$ es partición de $S^2 \setminus D$. Lo mismo sucede con $B_1$ y $B_2$. Así que $S^2 \setminus D$ se puede partir en cuatro pedazos (en la que cada uno no es sino unión de rotaciones del análogo de Vitali en $\mathbb{R}^3$) con los cuales se pueden formar dos copias, bajo rotaciones, de $S^2 \setminus D$.

[20] No podemos tomar todas las órbitas, ya que $S^2 \setminus D$ no necesariamente es cerrado bajo $G_{\tau\sigma}$.

[21] La partición numerable que se ha obtenido hasta este punto también sirve para demostrar que en $\mathbb{R}^3$ se niega la hipótesis de la medida. La demostración es igual a la de $\mathbb{R}^1$con el conjunto de Vitali.

El hecho de que $G_{\tau\sigma}$, a pesar de ser numerable como lo son $R_{\mathbb{Q}}$ y $T_{\mathbb{Q}}$, tenga dos subconjuntos propios ($R_1$ y $R_3$) tales que, dejando invariante a uno y aplicando una sola rotación al otro, se vuelva a obtener todo $G_{\tau\sigma}$ es una propiedad especialmente fuerte, ya que para el caso de $\mathbb{R}^1$ para poder recuperar a todo $T_{\mathbb{Q}}$ a partir de la familia $\{\rho_{2i}\}_{i\in\mathbb{N}}$ era necesario usar una traslación particular ($\sigma_i = \rho_i\rho_{2i}^{-1}$ )para cada uno de sus elementos.

Nótese que, tanto en los casos de $\mathbb{R}^1$y $\mathbb{R}^2$ como de $\mathbb{R}^3$, considerar los trasformaciones rígidas desde la teoría de grupos es de gran utilidad, ya que la sola definición de grupo nos garantiza la cerradura bajo composición y la existencia de inversos, que es lo único que se emplea para las particiones numerables. Sin embargo, para el caso de $\mathbb{R}^3$ se tiene una propiedad adicional que marca la diferencia con $\mathbb{R}^1$y $\mathbb{R}^2$. Se trata de una propiedad que puede plantearse en los siguientes términos:

**Definición 8** *Un grupo $G$ es paradójico si y sólo si existe una familia finita $\{G_i\}_{i=0}^{2n}$ de subconjuntos de $G$ ajenos dos a dos y una colección de elementos $\{g_i\}_{i=0}^{2n}$ de $G$ tales que $G = \bigcup_{i=0}^{n} g_{2i}[G_{2i}]$ y $G = \bigcup_{i=0}^{n-1} g_{2i+1}[G_{2i+1}]$*[22].

Desde luego que la denominación "grupo paradójico" señala ya la presencia de un comportamiento extraño, pero la definición no parece, al menos de entrada, ser contraintuitiva. Incluso en casos particulares como el grupo $G_{\tau\sigma}$. Es, más bien, cuando se piensa en los conjuntos sobre los que actúa un grupo que esta definición se transforma en afirmaciones especialmente sorprendentes, como lo es el teorema de Banach-Tarski. La propiedad de un conjunto de ser paradójico bajo la acción de un grupo se puede formalizar de la siguiente manera:

**Definición 9** *Sea $G$ un grupo que actúa sobre un conjunto $X$ . $E \subseteq X$ es G-paradójico si y sólo si existen $A_1, ..., A_m$, $B_1, ..., B_n$ subconjuntos ajenos de $E$ y $g_1, ...g_m, h_1, ..., h_n$ elementos de $G$ tales que $E = \bigcup_{i=0}^{m} g_i[A_i]$ y $E = \bigcup_{i=0}^{n} h_i[B_i]$.*

Cabe señalar, que si un conjunto es G-paradójico no necesariamente quiere decir que G es un grupo paradójico. Sin embargo, en el caso particular del teorema de Banach-Tarski, $S^2$ es $G_{\tau\sigma}$−paradójico porque $G_{\tau\sigma}$ es un grupo paradójico. En ese sentido, es interesante señalar que es el axioma de elección el que permitió la herencia de la propiedad "paradójica", dando un resultado verdaderamente contraintuitivo. De modo que, aunque el resultado de Banach-Tarski se base en propiedades del grupo de isometrías, éstas no servirían de nada si no se acepta la posibilidad de elegir sobre una familia de $c$ conjuntos no vacíos[23].

---

[22] $g_{2i}[G_{2i}] = \{g_{2i}h : h \in G_{2i}\}$ y $g_{2i+1}[G_{2i+1}] = \{g_{2i+1}h : h \in G_{2i+1}\}$

[23] Esto entra en consonancia con los resultados de Solovay, pues aunque todas las demostraciones de la negación de la hipótesis de la medida que daremos hacen uso de alguna de las propiedades de la medida, éstas no son condiciones suficientes. Mientras que, bajo ciertas hipótesis teórico conjuntistas (ver Sección 5.3), el axioma de elección es condición necesaria.

El estudio de las descomposiciones paradójicas, aunque actualmente está inmerso en la teoría de grupos, tuvo su origen en los diversos intentos por clarificar y responder el problema de la medida. Hay quienes consideran que el conjunto de Vitali es " el primer caso en el que se usaron descomposiciones paradójicas" (Wagon [24] p.7) , aunque se tratara de una cantidad numerable de piezas. En ese sentido, se podría tener una versión más laxa de la definición de que un conjunto sea paradójico bajo la acción de un grupo, si se permite una familia numerable de subconjuntos, en lugar de una finita. En cuyo caso, como ya se había señalado, $S^2$ también sería $G_{\tau\sigma}$ descomponible simplemente porque $G_{\tau\sigma}$ es un grupo numerable. Sin embargo, conviene mantener la definición original, porque con ella se ve por qué si se abandona la $\sigma$ – *aditividad* no se tienen resultados a-dimensionales. En resumen, lo que podemos concluir con lo visto hasta ahora es que la dimensión de los espacios euclideanos modifica el papel que juega la $\sigma$-aditividad en la existencia de conjuntos no medibles.

| Hipótesis | | Resultado | Hipótesis |
|---|---|---|---|
| Dimensión | Medida | Conjuntos no medibles | Adicionales |
| $\mathbb{R}^1$ | $\sigma$ – *aditiva* Invariante | Vitali | $BO_{\mathbb{R}}$ |
| $\mathbb{R}^3$ | $\sigma$ – *aditiva* Invariante | Análogo de Vitali | $BO_{\mathbb{R}}$ |
| $\mathbb{R}^1$ | Finito-aditiva Invariante | Ninguno (Medida de Banach) | Hahn-Banach (Lema de Zorn) |
| $\mathbb{R}^3$ | Finito-aditiva Invariante | Análogo de Vitali | Banach-Tarski (Descomposiciones paradójicas y $BO_{\mathbb{R}}$) |

# Capítulo 3

# Invarianza bajo transformaciones rígidas

El caso de $\mathbb{R}^1$ indica que basta con abandonar la $\sigma-aditividad$ para dar una respuesta positiva al problema de la medida en dicha dimensión. Sin embargo, el caso de $\mathbb{R}^3$ muestra que no se ganaría nada con hacerlo, ya que ahí la $\sigma-$ *aditividad* no es condición necesaria para tener conjuntos no medibles. Eso nos conduce a considerar la invarianza bajo traslaciones, que es la otra condición que se le pide a una extensión de $lh$. Si se piensa el problema para cualquier dimensión, lo que nos estamos preguntando es ¿Qué pasaría si se abandona la invarianza bajo transformaciones rígidas? o de manera más específica ¿Es la invarianza bajo transformaciones rígidas condición necesaria para negar la hipótesis de la medida?

Si la respuesta a esta última pregunta es afirmativa, entonces tendríamos que la invarianza bajo transformaciones rígidas es realmente culpable de la negación de la hipótesis de la medida, frente a lo cual habría que evaluar con detenimiento cuáles serían los costos de abandonarla. Sin embargo, se verá que dicha propiedad no es la responsable. Para ello introduciremos la medida de Lebesgue y demostraremos que la pregunta sobre cuál es el dominio de la medida de Lebesgue es equivalente al problema de la medida; no sólo porque la medida de Lebesgue cumple con las condiciones (1) a (4), de manera que si su dominio fuera $\wp(\mathbb{R}^1)$ entonces se cumpliría la hipótesis de la medida, sino porque cualquier medida que cumpla las condiciones (1) a (4) es en realidad una extensión de la medida de Lebesgue. Finalmente se dará una demostración de la existencia de conjuntos no medibles sin usar la invarianza bajo traslaciones.

## 3.1. Medida de Lebesgue.

La existencia del conjunto de Vitali niega la hipótesis de la medida, pero ¿Quiere decir esto que si aceptamos el axioma de elección entonces es imposible extender la noción de tamaño geométrico a conjuntos distintos de la unión de intervalos? Desde luego que no y la construcción de la medida de Lebesgue es un ejemplo de ello. Más aún, veremos que la medida de Lebesgue extiende la noción de tamaño geométrico a conjuntos como el de Cantor, con el que introdujimos el problema de la medida. En lo que sigue detallaremos la construcción de la medida de Lebesgue dada por Oxtoby [19], señalando desde ahora los puntos en los que se hace uso de axioma de elección, así como la versión de la que se trate.

**Definición 10** *Dado un conjunto $A \subset \mathbb{R}^1$, decimos que $C$ es una cubierta abierta de $A$ si y sólo si $C = \{I_i\}_{i\in F}$, con $I_i \in I_{\mathbb{R}} = \{(a,b) : a, b \in \mathbb{R}\}$ y $A \subset \bigcup_{i\in F} I_i$.*

**Observación 6** *Aunque en la definición no se especifica la cantidad de intervalos que conforman una cubierta abierta, como $\mathbb{R}^1$ es un espacio separable, a lo largo de toda esta sección se trabajará con cubiertas a lo más numerables.*

Las cubiertas abiertas son la clave para construir una función que aproxime el tamaño geométrico de un conjunto.

**Definición 11** *Sea $\mu^* : \wp(\mathbb{R}^1) \to [0,\infty]$ tal que para $A \subset \mathbb{R}^1$ :*

$$\mu^*(A) = \inf\left\{\sum_{n\in\mathbb{N}} |a_n - b_n| : A \subset \bigcup_{n\in\mathbb{N}} (a_n, b_n)\right\}$$

Nótese que $\mu^*$ está bien definida, ya que $C = \{(-n,n)\}_{n\in\mathbb{N}}$ es una cubierta abierta de cualquier subconjunto de reales y, por lo tanto, el conjunto sobre el que se toma el ínfimo es distinto del vacío. Además, el rango de $\mu^*$ efectivamente es $[0,\infty]$, pues la longitud de un intervalo siempre es un valor no negativo. La función $\mu^*$ se conoce como la medida exterior inducida por $lh$[1], ya que $\mu^*$ está definida a partir de la medida de los "vestidos que tapan al conjunto". En ese sentido, el ínfimo del tamaño de las cubiertas de un conjunto podría pensarse como "el traje hecho a su medida". Desafortunadamente no podemos garantizar que para todos los conjuntos ese ínfimo sea el traje perfecto. De modo que $\mu^*$, aunque está definida para todos los conjuntos de reales, sólo nos refiere el tamaño geométrico de un conjunto si se cumplen ciertas condiciones, bajo las cuales diremos que un conjunto es $\mu^*-$medible.

**Definición 12** *Sea $A \subset \mathbb{R}^1$. $A$ es $\mu^*$-medible si y sólo si dada $\epsilon > 0$ existen $F$ cerrado y $G$ abierto tales que $F \subset A \subset G$ y $\mu^*(G - F) < \epsilon$.*

[1] Nótese que se trata de una generalización del concepto de contenido exterior, ya que en este caso las cubiertas pueden ser numerables.

La definición original de un conjunto $\mu^* -$*medible* es bastante menos intuitiva, como bien lo señala Folland [2]. En cambio, la que hemos dado puede pensarse como un indicador de que es posible encontrar cubiertas que se ajusten cada vez mejor al "cuerpo del conjunto"[3]; ya que se pueden encontrar trajes hechos a partir de los "retazos" topológicos (es decir, de los intervalos abiertos y sus complementos) que cubran al conjunto por dentro y por fuera y cuya diferencia en tamaño sea mínima. En ese sentido podríamos decir que para los conjuntos $\mu^*$-medibles existe un traje que es casi su piel, mientras que para los conjuntos que no cumplen esa condición no es posible confeccionar trajes, con los materiales dados, que se ajusten a su "cuerpo"; de manera que, aunque existan trajes que los cubran, estos están lejos de referirnos su tamaño geométrico.

**Notación 2** $\mathfrak{M} = \{A \subset \mathbb{R}^1 : A \text{ es } \mu^*\text{-medibles}\}$.

A continuación demostraremos una serie de resultados que caracterizan a $\mathfrak{M}$ y que demuestran que $\mu^* |_{\mathfrak{M}}$ cumple con las condiciones (1) a (4) de la hipótesis de la medida.

**Lema 13** *$\mu^*(\emptyset) = 0$. Además si $A = \{a\}$, con $a \in \mathbb{R}$, entonces $\mu^*(A) = 0$.*

**Demostración.** Sea $C_n = \{(\frac{-1}{n}, \frac{1}{n})\}$. Está claro que $\forall n \in \mathbb{N}$, $C_n$ es cubierta del $\emptyset$ y además que $\inf\{lh(\frac{-1}{n}, \frac{1}{n}) : n \in \mathbb{N}\} = 0$, así que $\mu^*(\emptyset) = 0$. Para la segunda parte sea $A = \{a\}$ y sea $\epsilon > 0$, entonces tomamos la cubierta $C_\epsilon = \{(a - \frac{\epsilon}{2}, a + \frac{\epsilon}{2})\}$ para ver que $\mu^*(A) \leq \epsilon$. ∎

**Teorema 14** *Sea $I = [a, b] \subseteq \mathbb{R}^1$, entonces $\mu^*(I) = |a - b|$.*

**Demostración.** Por un lado, $\mu^*(I) \leq |a - b|$, pues si tomamos la sucesión de cubiertas $\{C_n\}_{n\in\mathbb{N}}$ tal que $\forall n \in \mathbb{N}$, $C_n = \{(a - \frac{1}{n}, b + \frac{1}{n})\}$ está claro que $\mu^*([a, b]) \leq \inf\{|(a - \frac{1}{n}) - (b + \frac{1}{n})| : n \in \mathbb{N}\} = |a - b|$. Para ver la otra desigualdad, sea $\epsilon > 0$ y $C_\epsilon = \{(a_n, b_n)\}_{n\in\mathbb{N}}$ una cubierta de $I$ tal que $\sum\limits_{n\in\mathbb{N}} |a_n - b_n| \leq \mu^*(I) + \epsilon$. Como $I \subset \mathcal{U} = \bigcup\limits_{n\in\mathbb{N}} (a_n, b_n)$ e $I$ es, por el teorema de Heine-Borel, compacto, entonces existe $\mathcal{K} = \{(a_{n_i}, b_{n_i})\}_{i=o}^{k} \subset C_\epsilon$ una subcubierta finita de $I$. Por otro lado, para $i \leq k$ tomamos $I_i = (a_{n_i}, b_{n_i}) \cap I$. Nótese que $I_i$ es un intervalo tal que $lh(I_i) \leq |a_{n_i} - b_{n_i}|$ o es el vacío y que $I = \bigcup\limits_{i=0}^{k} I_i$. De modo que:

$$|a - b| = lh(I) \leq \sum_{i=0}^{k} lh(a_{n_i}, b_{n_i}) = \sum_{i=0}^{k} |a_{n_i} - b_{n_i}| \leq \sum_{n\in\mathbb{N}} |a_n - b_n| \leq \mu^*(I) + \epsilon.$$

Esta última desigualdad es válida para toda $\epsilon > 0$, así que $|b - a| \leq \mu^*(I)$. ∎

[2] La definición de Lebesgue dice: $A \subset \mathbb{R}^1$ acotado es $\mu^*$-medible ssi $\mu^*(A) + \mu^*((a, b) \setminus A) = |b - a|$. La que se maneja aquí (Oxtoby [19]) también difiere de la de Folland, que dice: $A \subset \mathbb{R}^1$ es $\mu^*$-medible ssi dado $E \subset \mathbb{R}^1$, $\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$. Halmos ([11] p.44) argumenta que esta última no sólo facilita algunas demostraciones, sino que resulta natural si se consideran las propiedades aditivas que genera. Sin embargo, la de Oxtoby parece ser más adecuada para la discusión del problema de la medida en la recta real, ya que da una idea muy clara sobre qué tipo de conjuntos son medibles.

[3] Ya que los "trajes" internos y externos se parecen más y más.

**Teorema 15 ($\mu^*$ es invariante bajo traslaciones)** *Si $A \in \mathfrak{M}$ y $x \in \mathbb{R}$, entonces $A_{trasx} = \{a + x : a \in A\} \in \mathfrak{M}$ y $\mu^*(A_{trasx}) = \mu^*(A)$.*

**Demostración.** Primero demostraremos que $\mu^*(A_{trasx}) = \mu^*(A)$. Como ya sabemos que la longitud de los intervalos es invariante bajo traslaciones, basta demostrar que el conjunto de las cubiertas de $A_{trasx}$ es igual al conjunto de las cubiertas de $A$ trasladadas por $x$. Si $\mathcal{C} = \{(a_n, b_n)\}_{n\in\mathbb{N}}$, entonces definimos $\mathcal{C}_{trasx} = \{(a_n, b_n)_{trasx}\}_{n\in\mathbb{N}}$. Sea $(C_A)_{trasx} = \{\mathcal{C}_{trasx} : \mathcal{C}$ es cubierta abierta de $A\}$ y sea $C_{A_{trasx}} = \{\mathcal{C} : \mathcal{C}$ es cubierta abierta de $A_{trasx}\}$. Hay que demostrar que $C_{A_{trasx}} = (C_A)_{trasx}$.

a) $C_{A_{trasx}} \supseteq (C_A)_{trasx}$. Sea $\mathcal{C} = \{(a_n, b_n)\}_{n\in\mathbb{N}}$ una cubierta de $A$. Como $A \subset \bigcup\limits_{n\in\mathbb{N}} (a_n, b_n)$, entonces $A_{trasx} = \{a + x : a \in A\} \subset \bigcup\limits_{n\in\mathbb{N}} (a_n, b_n)_{trasx}$. Así que $\mathcal{C}_{trasx} = \{(a_n, b_n)_{trasx}\}_{n\in\mathbb{N}}$ es cubierta abierta de $A_{trasx}$.

b) $C_{A_{trasx}} \subseteq (C_A)_{trasx}$. La demostración de esta contención es análoga a la anterior, pues dada $\mathcal{C} \in C_{A_{trasx}}$, hay que tomar $\mathcal{C}_{tras-x}$, que se puede ver que es cubierta de $A$.

Por último, hay que ver que $A_{trasx} \in \mathfrak{M}$. Sea $\epsilon > 0$. Como $A \in \mathfrak{M}$ existen $F$ cerrado y $G$ abierto tales que $F \subset A \subset G$ y $\mu^*(F-G) < \epsilon$. Claramente $F_{trasx} \subset A_{trasx} \subset G_{trasx}$ y por el resultado anterior $\mu^*(F - G) = \mu^*((F - G)_{trasx})$. Además $(F - G)_{trasx} = (F_{trasx} - G_{trasx})$. Así que $\mu^*(F_{trasx} - G_{trasx}) < \epsilon$ y, por lo tanto, $A_{trasx} \in \mathfrak{M}$. ■

Para demostrar la $\sigma - aditividad$ es necesario probar varios resultados preliminares.

**Lema 16 ($\mu^*$ es monótona)** *Si $A \subset B$ entonces $\mu^*(A) \leq \mu^*(B)$.*

**Demostración.** Esto es porque toda cubierta abierta de B es cubierta abierta de A. Es decir, como

$$\left\{\sum_{n\in\mathbb{N}} |a_n - b_n| : B \subset \bigcup_{n\in\mathbb{N}} (a_n, b_n)\right\} \subset \left\{\sum_{n\in\mathbb{N}} |a_n - b_n| : A \subset \bigcup_{n\in\mathbb{N}} (a_n, b_n)\right\}$$

entonces $\mu^*(A) \leq \mu^*(B)$. ■

**Lema 17 ($\mu^*$ es subaditiva)** *Sea $A = \bigcup\limits_{j\in\mathbb{N}} A_j$, entonces $\mu^*(A) \leq \sum\limits_{j\in\mathbb{N}} \mu^*(A_j)$.*

**Demostración.** Dada $\epsilon > 0$, para cada $j \in \mathbb{N}$ sea $\mathcal{C}_j = \{(a_{ij}, b_{ij})\}_{i\in\mathbb{N}}$ una cubierta de $A_j$ tal que $\sum\limits_{i\in\mathbb{N}} |a_{ij} - b_{ij}| \leq \mu^*(A_j) + \frac{\epsilon}{2^j}$. Esto quiere decir que $\mathcal{C} = \bigcup\limits_{j\in\mathbb{N}} \mathcal{C}_j$ es una cubierta de $A$, así que $\mu^*(A) \leq \sum\limits_{j\in\mathbb{N}} \sum\limits_{i\in\mathbb{N}} |a_{ij} - b_{ij}| \leq \sum\limits_{j\in\mathbb{N}} \left(\mu^*(A_j) + \frac{\epsilon}{2^j}\right) = \sum\limits_{j\in\mathbb{N}} \mu^*(A_j) + \epsilon$. Como esto es válido para cualquier $\epsilon > 0$, tenemos que $\mu^*(A) \leq \sum\limits_{j\in\mathbb{N}} \mu^*(A_j)$. ■

**Observación 7** *Si* $A = \bigcup_{n\in\mathbb{N}} \{a_n\}$ *entonces* $\mu^*(A) = 0$, *ya que por el lema anterior* $\mu^*(A) \leq \sum_{n\in\mathbb{N}} \mu^*(\{a_n\})$ *y, por el Lema 13,* $\forall n \in \mathbb{N}$, $\mu^*(\{a_n\}) = 0$. *De modo que* $\mu^*(A) = 0$. *En otras palabras, todo conjunto numerable tiene medida exterior cero.*

**Lema 18** *Sea* $\{A_i\}_{i=0}^n$ *una sucesión finita de conjuntos cerrados, acotados y ajenos dos a dos. Si* $A = \bigcup_{i=0}^n A_i$ *entonces* $\mu^*(A) = \sum_{i=0}^n \mu^*(A_i)$[1].

**Demostración.** La demostración se hará mediante inducción sobre el número de conjuntos ajenos:

1) Caso base: $n = 0$. No hay nada que demostrar pues $A = A_0$.

2) Paso inductivo. Sea $A = \bigcup_{i=0}^n A_i$ y $A_{n+1}$ tales que $A_l \cap A_j = \emptyset$ si $l \neq j$ y supongamos que $\mu^*(A) = \sum_{i=0}^n \mu^*(A_i)$. Hay que demostrar que $\mu^*(A \cup A_{n+1}) = \sum_{i=0}^{n+1} \mu^*(A_i)$. Para ello basta probar que si $B_1$ y $B_2$ conjuntos cerrados y acotados tales que $B_1 \cap B_2 = \emptyset$, entonces $\mu^*(B_1 \cup B_2) = \mu^*(B_1) + \mu^*(B_2)$, pues podemos tomar $B_1 = \bigcup_{i=0}^n (A_i)$ y $B_2 = A_{n+1}$ y usar la hipótesis de inducción.

**Observación 8** *Si* $B_1$ *y* $B_2$ *son cerrados y* $B_1 \cap B_2 = \emptyset$, *entonces existe* $\delta > 0$ *tal que para toda* $I = (a,b)[|I| \leq \delta \Longrightarrow (B_1 \cap I \neq \emptyset \rightarrow B_2 \cap I = \emptyset) \wedge (B_2 \cap I \neq \emptyset \rightarrow B_1 \cap I = \emptyset)]$.

Dada $\epsilon > 0$ existe una cubierta abierta $\mathcal{C} = \{(a_n, b_n)\}_{n\in\mathbb{N}}$ de $(B_1 \cup B_2)$ tal que $\sum_{n\in\mathbb{N}} |a_n - b_n| \leq \mu^*(B_1 \cup B_2) + \frac{\epsilon}{2}$. Más aún, existe una cubierta $\mathcal{C}_\delta = \{(c_n, d_n)\}_{n\in\mathbb{N}}$ de $(B_1 \cup B_2)$ tal que $\forall n \in \mathbb{N}$, $|c_n - d_n| \leq \delta$ y $\sum_{n\in\mathbb{N}} |c_n - d_n| \leq \sum_{n\in\mathbb{N}} |a_n - b_n| + \frac{\epsilon}{2}$[5]. Así que $\sum_{n\in\mathbb{N}} |c_n - d_n| \leq \mu^*(B_1 \cup B_2) + \epsilon$. Sean $\mathcal{C}_{B_1} = \{(c_n, d_n) \in \mathcal{C}_\delta : (c_n, d_n) \cap B_1 \neq \emptyset\}$ y $\mathcal{C}_{B_2} = \{(c_n, d_n) \in \mathcal{C}_\delta : (c_n, d_n) \cap B_2 \neq \emptyset\}$. Como $\mathcal{C}_\delta$ es cubierta de $(B_1 \cup B_2)$ y $\forall n \in \mathbb{N}$, $|c_n - d_n| \leq \delta$, por la Observación 8, tenemos que $\mathcal{C}_{B_1}$ y $\mathcal{C}_{B_2}$ son cubiertas ajenas de $B_1$ y $B_2$ respectivamente. Así que $\mu^*(B_1) + \mu^*(B_2) \leq \sum_{(c_n,d_n)\in\mathcal{C}_{A_1}} |c_n - d_n| + \sum_{(c_n,d_n)\in\mathcal{C}_{A_2}} |c_n - d_n| \leq \sum_{n\in\mathbb{N}} |c_n - d_n| \leq$

[1] Nótese que este lema demuestra la finito- aditividad de $\mu^*$ para los conjuntos cerrados y acotados. Este resultado servirá como base para demostrar la $\sigma$−aditividad de $\mu^*$ en general. De hecho, éste es un buen ejemplo de cómo se van generalizando los resultados para la construcción de la medida de Lebesgue.

[5] Sea $(a, b)$ un intervalo cualquiera. Sea $m$ el mayor entero de $\frac{2|a-b|}{\delta}$ y $\xi = \min\{\delta, \frac{\epsilon}{2}\}$, entonces tomamos $P_\delta = \{(a + n\frac{\delta}{2}, a + (n+1)\frac{\delta}{2}) : 0 \leq n \leq m-1\} \cup \{(a + m\frac{\delta}{2} - \frac{\xi}{2}, a + m\frac{\delta}{2} + \frac{\xi}{2})\}$. Claramente $P_\delta$ es una cubierta de $(a, b)$ tal que los intervalos que la conforman tienen longitud menor igual que $\delta$ y $\mu^*(P_\delta) \leq |a - b| + \frac{\epsilon}{2}$.

Esta construcción se puede hacer con todos los intervalos $(a_n, b_n)$ que componen la cubierta $C$, de modo que $\mu^*(P_{\delta n}) \leq |a_n - b_n| + \frac{\epsilon}{2^n}$.

$\mu^*(B_1 \cup B_2) + \epsilon$. Pero como esta desigualdad se cumple para toda $\epsilon > 0$, tenemos qué $\mu^*(B_1) + \mu^*(B_2) \leq \mu^*(B_1 \cup B_2)$. La otra desigualdad se desprende inmediatamente de la subaditividad de $\mu^*$ (ver Lema 17). ■

**Lema 19** *Sea $G$ un conjunto abierto y acotado, entonces dada $\epsilon > 0$ existe $F$ un subconjunto cerrado de $G$ tal que $\mu^*(G) - \epsilon < \mu^*(F)$.*

**Demostración.** Como todo conjunto abierto de reales es unión numerable de intervalos abiertos disjuntos[6] tenemos que $G = \bigcup\limits_{n\in\mathbb{N}} (a_n, b_n)$, con $(a_i, b_i) \cap (a_j, b_j) = \emptyset$, si $i \neq j$. Claramente $\{(a_n, b_n)\}_{n\in\mathbb{N}}$ es una cubierta abierta de G, así que $\mu^*(G) \leq \sum\limits_{n\in\mathbb{N}} |a_n - b_n|$. Sea $m \in \mathbb{N}$ tal que $\mu^*(G) - \frac{\epsilon}{2} < \sum\limits_{n=0}^{m} |a_n - b_n|$[7]. Para toda $n < m$ sea $J_n \subset (a_n, b_n)$ un intervalo cerrado tal que $|a_n - b_n| - \frac{\epsilon}{2n} < |J_n|$. Nótese que $J_i \cap J_j = \emptyset$ si $i \neq j$ y que $F = \bigcup\limits_{n=0}^{m} J_n$ es un subconjunto cerrado de G, por ser unión finita de cerrados. Por otro lado, por el Lema 14, $\mu^*(J_n) = |J_n|$. Además, como $J_n$ es cerrado y acotado, por el Lema 18, $\mu^*(F) = \sum\limits_{n=0}^{m} |J_n|$. Por lo tanto, $\sum\limits_{n=0}^{m} |a_n - b_n| - \frac{\epsilon}{2} < \mu^*(F)$. Además $\mu^*(G) - \frac{\epsilon}{2} < \sum\limits_{n=0}^{m} |a_n - b_n|$ . Así que, juntando estos dos últimos resultados, tenemos que $\mu^*(G) - \epsilon < \mu^*(F)$. ■

**Lema 20** *Sea $F$ un subconjunto cerrado de un conjunto abierto y acotado $G$, entonces $\mu^*(G - F) = \mu^*(G) - \mu^*(F)$.*

**Demostración.** Como $F \subset G$ y $G$ es acotado, entonces $\mu^*(F) < \infty$. Así que $\mu^*(G - F) = \mu^*(G) - \mu^*(F) \iff \mu^*(G - F) + \mu^*(F) = \mu^*(G)$. Demostraremos esta última igualdad.

(1) $\mu^*(G - F) + \mu^*(F) \leq \mu^*(G)$. Sea $\epsilon > 0$. Como $G - F$ es un conjunto abierto y acotado, pues $G$ y $F^c$ lo son y $G - F = G \cap F^c$, entonces por el lema anterior sabemos que existe $F'$ un conjunto cerrado tal que $F' \subset G - F$ y $\mu^*(G - F) - \epsilon < \mu^*(F')$, entonces $\mu^*(G - F) + \mu^*(F) - \epsilon < \mu^*(F') + \mu^*(F)$. Por otro lado, como $F'$ y $F$ son cerrados ajenos, por el Lema 18, tenemos que $\mu^*(F') + \mu^*(F) = \mu^*(F' \cup F)$. Como $(F' \cup F) \subset G$ y $\mu^*$ es monótona (ver Lema 16) entonces $\mu^*(F') + \mu^*(F) = \mu^*(F' \cup F) \leq \mu^*(G)$. Así que dada $\epsilon > 0$, $\mu^*(G - F) + \mu^*(F) - \epsilon \leq \mu^*(G)$ y, por lo tanto, $\mu^*(G - F) + \mu^*(F) - \epsilon \leq \mu^*(G)$.

(2) $\mu^*(G) \leq \mu^*(G - F) + \mu^*(F)$. Esta desigualdad se desprende inmediatamente de la subaditividad de $\mu^*$ (ver Lema 17), ya que $G = (G - F) \cup F$. ■

[6]La demostración de este resultado es muy sencilla y puede consultarse en Folland [9] (p.12).

[7]De lo contrario, $\sum\limits_{n=0}^{\infty} |a_n - b_n| \leq \mu^*(G) - \frac{\epsilon}{2} < \mu^*(G)$!

El siguiente teorema muestra con mayor claridad la idea de que los conjuntos $\mu^*$-medibles son aquellos conjuntos para los cuales es posible confeccionar un traje (con los retazos que se tienen) que se ajuste tanto como se quiera al conjunto y que nos refiriera su tamaño del mismo modo que las tallas nos refieren el tamaño de las personas.

**Teorema 21** *Sea $A$ un conjunto acotado. Si dada $\epsilon > 0$ existe $F \subset A$ tal que $F$ es cerrado y $\mu^*(A) - \epsilon < \mu^*(F)$ entonces $A$ es $\mu^* -$ medible.*

**Demostración.** Sea $\epsilon > 0$ y $F \subset A$ tal que $F$ es cerrado y $\mu^*(A) - \frac{\epsilon}{2} < \mu^*(F)$. Como A es acotado $\mu^*(A) < \infty$ , entonces existe una cubierta $C = \{(a_n, b_n)\}_{n\in\mathbb{N}}$ de $A$ tal que $|a_n - b_n| < 1$ y $\sum_{n\in\mathbb{N}} |a_n - b_n| < \mu^*(A) + \frac{\epsilon}{2}$[8]. Sea $B = \{(a_n, b_n) \in C : (a_n, b_n) \cap A \neq \emptyset\}$ y sea $G = \bigcup B$. Nótese que $F \subset A \subset G$, con G abierto y además acotado[9]. Así que por el Lema 20 $\mu^*(G - F) = \mu^*(G) - \mu^*(F)$. Además por la monotonía y la subaditividad de $\mu^*$ tenemos que $\mu^*(G) \leq \sum_{n\in\mathbb{N}} |a_n - b_n|$ y, por lo tanto, $\mu^*(G - F) = \mu^*(G) - \mu^*(F) \leq \sum_{n\in\mathbb{N}} |a_n . b_n| - \mu^*(F) < \mu^*(A) + \frac{\epsilon}{2} - (\mu^*(A) - \frac{\epsilon}{2}) < \epsilon$. De modo que $A$ es $\mu^*-$medible. ■

Este lema, que puede pensarse como un criterio de $\mu^*$- medibilidad, es interesante porque muestra que los conjuntos que. en términos de $\mu^*$, se "parecen mucho" a conjuntos cerrados deben ser $\mu^*-$medibles. En particular nos garantiza que los intervalos cerrados son $\mu^* -$ *medibles* como se ve en el siguiente corolario.

**Corolario 22** *Si $I = [a, b]$, entonces $I \in \mathfrak{M}$.*

**Demostración.** Sabemos, por el Lema 14 que si $I = [a, b]$, entonces $\mu^*(I) = |a - b|$. Así que basta tomar a $F = I$ para poder aplicar el teorema anterior y tener que $I \in \mathfrak{M}$. Pues dada $\epsilon > 0$, siempre se tiene que $\mu^*(I) - \epsilon < \mu^*(I)$. ■

Este teorema además apunta hacia un aspecto central en la discusión del problema de la medida, a saber: que la hipótesis de la medida se traduce en una propiedad de "buen comportamiento" de los conjuntos que conforman la potencia de los reales, como se verá con mayor claridad cuando nos adentremos en los resultados de Solovay.

**Lema 23** *Sea $A = \bigcup_{i\in\mathbb{N}} A_i$ (con $A_i \in \mathfrak{M}$ y $A_i \cap A_j = \emptyset$ si $i \neq j$) un conjunto acotado, entonces $A$ es $\mu^*-$medible y $\mu^*(A) = \sum_{i\in\mathbb{N}} \mu^*(A_i)$.*

---

[8] Para garantizar la desigualdad estricta es necesario que el conjunto sea acotado.

[9] Como $\forall y \in G\ \exists x \in A(\ |x - y| < 1)$, ya que $|a_n - b_n| < 1)$, entonces $G$ es acotado porque A lo es.

**Demostración.** Como $A$ es acotado existe $I_A = (a,b)$ tal que $A \subset I_A$. Nótese que $A_i$ también es acotado. Como por hipótesis $A_i$ es medible, entonces dada $\epsilon > 0$ existen $F_i$ cerrado y $G_i$ abierto tales que $F_i \subset A_i \subset G_i$ y $\mu^*(G_i - F_i) < \epsilon$. Más aún, podemos asegurar que existe un conjunto $G_i'$ acotado que cumple con lo mismo que $G_i$, pues tomamos $G_i' = G_i \cap I_A$ que es abierto por ser intersección finita de abiertos y como $G_i' \subset G_i$ entonces $G_i' - F_i \subset G_i - F_i$. Así que, por la monotonía de $\mu^*$, tenemos que $\mu^*(G_i' - F_i) \leq \mu^*(G_i - F_i) < \epsilon$.

La ventaja de tomar $G_i'$ es que, al ser un abierto acotado tal que $F_i \subset A_i \subset G_i \cap I_A = G_i'$, podemos usar el Lema 20 para ver que $\mu^*(G_i') - \mu^*(F_i) = \mu^*(G_i' - F_i) < \epsilon$. Usando una vez más la monotonía de $\mu^*$ tenemos que $\mu^*(A_i) < \mu^*(G_i')$. Por lo tanto, $\mu^*(A_i) - \mu^*(F_i) < \mu^*(G_i') - \mu^*(F_i) < \epsilon$.

Si en particular tomamos $\frac{\epsilon}{2^i}$, con $\epsilon > 0$, tenemos un resultado de gran utilidad para nuestra demostración. Dada $i \in \mathbb{N}$, existe un conjunto cerrado $F_i \subset A_i$ tal que $\mu^*(A_i) - \frac{\epsilon}{2^i} < \mu^*(F_i)$. Además, como por hipótesis $A = \bigcup\limits_{i\in\mathbb{N}} A_i$, usando la subaditividad de $\mu^*$ tenemos que $\mu^*(A) \leq \sum\limits_{i\in\mathbb{N}} \mu^*(F_i)$, de donde se sigue que existe $k \in \mathbb{N}$ tal que:

$$\mu^*(A) - \frac{\epsilon}{2} < \sum_{i=1}^{k} \mu^*(A_i).$$

Nótese que $k$ existe, pues si $\sum\limits_{i\in\mathbb{N}} \mu^*(A_i) \leq \mu^*(A) - \frac{\epsilon}{2}$, como $\mu^*(A) \leq \sum\limits_{i\in\mathbb{N}} \mu^*(A_i)$, entonces $\mu^*(A) \leq \mu^*(A) - \frac{\epsilon}{2}$. Pero, como $\epsilon > 0$, eso querría decir que $\mu^*(A) < \mu^*(A) - \frac{\epsilon}{2} < \mu^*(A)$! Sea $F = \bigcup\limits_{i=1}^{k} F_i$. Nótese que F es cerrado y acotado por ser unión finita de cerrados y acotados. Como $F_i \cap F_j = \emptyset$, podemos usar el Lema 18, que nos garantiza la finito aditividad de $\mu^*$ para el caso de conjuntos cerrados, junto con lo que hemos señalado antes, para ver que:

$$\mu^*(A) - \epsilon < \sum_{i=1}^{k} \mu^*(A_i) - \frac{\epsilon}{2} < \sum_{i=1}^{k} \mu^*(A_i) - \frac{\epsilon}{2^i} < \sum_{i=1}^{k} \mu^*(F_i) = \mu^*(F).$$

De modo que, por el Teorema 21, $A$ es medible. Por otro lado, $\forall n \in \mathbb{N}$, $\sum\limits_{i=1}^{n} \mu^*(A_i) - \frac{\epsilon}{2^i} < \sum\limits_{i=1}^{n} \mu^*(F_i)$ y $\sum\limits_{i=1}^{n} \mu^*(A_i) - \frac{\epsilon}{2} < \sum\limits_{i=1}^{n} \mu^*(A_i) - \frac{\epsilon}{2^i}$. Así que $\forall n \in \mathbb{N}$:

$$\sum_{i=1}^{n} \mu^*(A_i) < \sum_{i=1}^{n} \mu^*(F_i) + \frac{\epsilon}{2} = \mu^*(\bigcup_{i=1}^{n} F_i) + \frac{\epsilon}{2} < \mu^*(A) + \frac{\epsilon}{2}.$$

Cuando $n \to \infty$ y $\epsilon \to 0$ tenemos que $\sum\limits_{i\in\mathbb{N}} \mu^*(A_i) \leq \mu^*(A)$. Por otro lado, la subaditividad (Lema 17) nos garantiza que $\mu^*(A) \leq \sum\limits_{i\in\mathbb{N}} \mu^*(A_i)$. De modo que $\sum\limits_{i\in\mathbb{N}} \mu^*(A_i) = \mu^*(A)$. ■

Con este lema ya es muy fácil demostrar la *σ−aditividad* para el caso general. Esta forma de demostrar resultados dentro de teoría de la medida no debe sorprendernos, ya que los conjuntos verdaderamente conflictivos (en términos del problema de la medida) son los conjuntos acotados, como se puede ver con la construcción de la medida de Banach. Por otro lado, tampoco debería de sorprendernos que la $\sigma - aditividad$ sea la propiedad más difícil de demostrar, ya que, en el caso de $\mathbb{R}^1$ es condición necesaria para la negación de la hipótesis de la medida.

**Teorema 24** **($\mu^*$ es $\sigma - aditiva$ sobre $\mathfrak{M}$)** *Sea* $A = \bigcup_{i\in\mathbb{N}} A_i$ *(con* $A_i \in \mathfrak{M}$ *y* $A_i \cap A_j = \varnothing$ *si* $i \neq j$*), entonces* $A \in \mathfrak{M}$ *y* $\mu^*(A) = \sum_{i\in\mathbb{N}} \mu^*(A_i)$.

Para dar la demostración de este teorema es necesario introducir un lema que además nos servirá para demostrar que $\mathfrak{M}$ es una $\sigma$−álgebra:

**Lema 25** *Si* $A, B \in \mathfrak{M}$ *entonces* $A \cap B \in \mathfrak{M}$.

**Demostración.** Sea $\epsilon > 0$. Dado que $A$ y $B \in \mathfrak{M}$, existen $F_1$ y $F_2$ conjuntos cerrados y $G_1$ y $G_2$ conjuntos abiertos tales que $F_1 \subset A \subset G_1$, $F_2 \subset B \subset G_2$, $\mu^*(G_1 - F_1) < \frac{\epsilon}{2}$ y $\mu^*(G_2 - F_2) < \frac{\epsilon}{2}$. Sean $F = F_1 \cap F_2$ y $G = G_1 \cap G_2$. Nótese que $F$ y $G$ son cerrados y abiertos respectivamente, por ser intersección finita de conjuntos del mismo tipo. Además $F \subset A \cap B \subset G$ y $G - F = G_1 \cap G_2 - F_1 \cap F_2 = G_1 \cap G_2 \cap (F_1^c \cup F_2^c) = (G_1 \cap G_2 \cap F_1^c) \cup (G_1 \cap G_2 \cap F_2^c) \subset (G_1 - F_1) \cup (G_2 - F_2)$. Así que usando que $\mu^*$ es monótona y subaditiva tenemos que $\mu^*(G - F) \leq \mu^*(G_1 - F_1) + \mu^*(G_2 - F_2) < \epsilon$. ■

**Demostración.** (Del Teorema 24) Sea $\{z_j\}_{j=1}^{\infty}$ una numeración de $\mathbb{Z}$ y $I_j = [z_j, z_j + 1]$. Para cada $j \in \mathbb{N}$ tomamos $A_{ij} = A_i \cap [z_j, z_j + 1]$, que es un conjunto medible por ser intersección de dos medibles (Corolario 22 y Lema 25) y acotado. Nótese además que para $j$ fija, si $i \neq l$ entonces $A_{ij} \cap A_{lj} = \emptyset$, ya que $A_{ij} \subset A_i$ y $A_{lj} \subset A_l$ y, por hipótesis, $A_i \cap A_l = \emptyset$ si $i \neq l$. Sea $B_j = \bigcup_{i\in\mathbb{N}} A_{ij}$. Nótese que $B_j$ es acotado (pues $A_{ij} \subset [z_j, z_j + 1]$) y medible, por los dos últimos lemas. De modo que dada $\epsilon > 0$, existen $F_j$ cerrado y $G_j$ abierto y acotado tales que $F_j \subset B_j \subset G_j$ y $\mu^*(G_j - F_j) = \mu^*(G_j) - \mu^*(F_j) < \frac{\epsilon}{2^j}$. Sea $F = \bigcup_{j\in\mathbb{N}} F_j$ y $G = \bigcup_{j\in\mathbb{N}} G_j$. $G$ es abierto por ser unión de abiertos. Hay que ver que $F$ es cerrado. Sea $x$ un punto de acumulación de $F$ entonces existe una sucesión $\{x_n\}_{n=0}^{\infty} \subset F = \bigcup_{j\in\mathbb{N}} F_j$ tal que $\lim_{n\to\infty} x_n = x$. Como toda sucesión de reales convergente es acotada y $F_j \subset [z_j, z_j + 1]$ entonces existe $n_0 \in \mathbb{N}$ tal que $\{x_n\}_{n=0}^{\infty} \subset \bigcup_{j=1}^{n_0} F_j$. Nótese además que $x \in \bigcup_{j=1}^{n_0} F_j \subset F$, ya que la unión finita de cerrados es un conjunto cerrado y $\lim_{n\to\infty} x_n = x$. Por lo tanto, $F$ es cerrado. Por otro lado, cabe señalar que $F \subset A \subset G$ y que $G - F = \bigcup_{j\in\mathbb{N}} G - F_j \subset \bigcup_{j\in\mathbb{N}} G_j - F_j$. Usando la subaditividad y la monotonía de $\mu^*$, $\mu^*(G - F) \leq \sum_{j\in\mathbb{N}} \mu^*(G_j - F_j) < \epsilon$. Por lo tanto, A es $\mu^*$−medible.

Para ver que $\mu^*(A) = \sum_{i\in\mathbb{N}} \mu^*(A_i)$ sólo hay que demostrar que $\sum_{i\in\mathbb{N}} \mu^*(A_i) \leq \mu^*(A)$, ya que la otra desigualdad se desprende inmediatamente de la subaditividad de $\mu^*$. Como $A_i = \bigcup_{j\in\mathbb{N}} A_{ij}$, entonces $\mu^*(A_i) \leq \sum_{j\in\mathbb{N}} \mu^*(A_{ij})$. Así que, usando la monotonía, la subaditividad de $\mu^*$y el Lema 23 tenemos que $\sum_{i\in\mathbb{N}} \mu^*(A_i) \leq \sum_{j\in\mathbb{N}} \sum_{i\in\mathbb{N}} \mu^*(A_{ij}) = \sum_{j\in\mathbb{N}} \mu^*(B_j)$. Por otro lado, como $F_j \subset B_j \subset G_j$, con $G_j$ abierto y acotado, usando el Lema 20 y la subaditividad de $\mu^*$, tenemos que $\mu^*(B_j) \leq \mu^*(G_j) = \mu^*(G_j - F_j) + \mu^*(F_j)$. De modo que $\sum_{j\in\mathbb{N}} \mu^*(B_j) \leq \sum_{j\in\mathbb{N}} \mu^*(G_j - F_j) + \sum_{j\in\mathbb{N}} \mu^*(F_j) \leq \epsilon + \sum_{j\in\mathbb{N}} \mu^*(F_j) = \epsilon + \mu^*(F)$. Poniendo todos estos resultados juntos llegamos a que dada $\epsilon > 0$, $\sum_{i\in\mathbb{N}} \mu^*(A_i) \leq \sum_{j\in\mathbb{N}} \mu^*(B_j) \leq \epsilon + \mu^*(F) \leq \epsilon + \mu^*(A)$. La última desigualdad se desprende de que $F \subset A$ y de la monotonía de $\mu^*$. Finalmente si tomamos $\epsilon \to 0$ tenemos que $\sum_{i\in\mathbb{N}} \mu^*(A_i) \leq \mu^*(A)$. ■

A continuación damos un teorema que garantiza propiedades de $\mathfrak{M}$ importantes para teoría de la medida[10].

**Teorema 26 ($\mathfrak{M}$ es una $\sigma$−álgebra)** *El conjunto de los subconjuntos de reales $\mu^*$−medibles ($\mathfrak{M}$) forma una $\sigma$−álgebra.*

**Demostración.** Por el Teorema 24 se tiene que si $A = \bigcup_{i\in\mathbb{N}} A_i$, con $A_i \in \mathfrak{M}$, entonces $A$ es $\mu^*$−medible. Así que sólo falta verificar que si $A \in \mathfrak{M}$, entonces $A^c \in \mathfrak{M}$. Sea $A \in \mathfrak{M}$. Dada $\epsilon > 0$, existen $F$ cerrado y $G$ abierto tales que $F \subset A \subset G$ y $\mu^*(G - F) < \epsilon$. Si tomamos $G' = F^c$ y $F' = G^c$, tenemos que $G'$ es abierto, $F'$ es cerrado y $F' \subset A^c \subset G'$. Además $G - F = G \cap F^c = (F')^c \cap G' = G' - F'$. De modo que $\mu^*(G' - F') = \mu^*(G - F) < \epsilon$. ■

Esta forma de caracterizar a $\mathfrak{M}$ resulta de gran utilidad, pues aunque no conozcamos a todos los elementos de $\mathfrak{M}$, sabemos que se trata de un conjunto cerrado bajo las operaciones conjuntistas de unión arbitraria y complemento. En particular, tenemos que si $B \subset \mathfrak{M}$, entonces la $\sigma$-álgebra generada por $B$ está contenida en $\mathfrak{M}$[11]. Esto nos permite ver que el dominio de $\mu^*$ contiene propiamente a $I_{\mathbb{R}}$, ya que si tomamos $B = I_{\mathbb{R}}$ entonces la $\sigma$−álgebra generada por $B$ (conocida como la $\sigma$− álgebra de Borel y denotada por $\mathfrak{B}_{\mathbb{R}}$) está contenida en $\mathfrak{M}$, pues veremos que $I_{\mathbb{R}} \subset \mathfrak{M}$. Además $I_{\mathbb{R}} \subsetneq \mathfrak{B}_{\mathbb{R}}$, ya que los conjuntos $G_\delta$ (es decir, los conjuntos que son intersección arbitrarias de abiertos) pertenece a la $\sigma$−álgebra de Borel[12] y, desde luego, que hay conjuntos $G_\delta$ que no son intervalos.

[10] En particular, nos permitirán demostrar que la medida de Lebesgue es una extensión significativa de $lh$, pues se verá que su dominio es equipotente con la potencia de los reales.

[11] La $\sigma$− álgebra generada por un conjunto es la mínima $\sigma$− álgebra que lo contiene.

[12] La intersección arbitraria de abiertos se puede ver como el complemento de la unión de los complementos de los abiertos involucrados. Recuérdese además que una $\sigma$−álgebra es cerrada bajo uniones arbitrarias y complementos.

Por otro lado, como ya mencionamos antes, la propiedad de ser $\mu^*$-medible es garantía de que el valor que otorga $\mu^*$ nos refiere el "tamaño geométrico" de un conjunto. Así que $\mu^*$ restringida a $\mathfrak{M}$ es una buena extensión de $lh$. Esa restricción es justamente lo que se conoce como la medida de Lebesgue.

**Definición 13** *La medida de Lebesgue es* $\mu^* |_{\mathfrak{M}}$ .

**Notación 3** $\mu_L$ *denota a la medida de Lebesgue (i.e.* $\mu_L = \mu^* |_{\mathfrak{M}}$*)*

**Teorema 27** *La medida de Lebesgue es extensión de lh.*

**Demostración.** Como $dom(\mu^*) = \wp(\mathbb{R}^1)$ y por definición $\mathfrak{M} \subset dom(\mu^*)$, entonces $\mathfrak{M} = dom(\mu_L) \subset \wp(\mathbb{R}^1)$. El Teorema 15 y el Teorema 24 demuestran que $\mu_L$ cumple con las condiciones (2) a (4) de la hipótesis de la medida. Así que queda por demostrar que $\forall (a,b) \subset \mathbb{R}^1, (a,b) \in \mathfrak{M}$ y $\mu^*((a,b)) = |a-b|$. $\mu^*((a,b)) = |a-b|$, se demuestra exactamente igual que como se hizo para los intervalos cerrados. Por el Corolario 22, $[a,b] \in \mathfrak{M}$, $[a,a] = \{a\} \in \mathfrak{M}$ y $[b,b] = \{b\} \in \mathfrak{M}$. Así que usando el hecho de que $\mathfrak{M}$ es una $\sigma-$álgebra (Teorema 26) tenemos que $(a,b) = [a,b] - \{a,b\} = [a,b] \cap \{a,b\}^c \in \mathfrak{M}$. ∎

## 3.2. Unicidad de la medida de Lebesgue

El hecho de que el dominio de la medida de Lebesgue contenga a los borelianos da una respuesta positiva, aunque parcial, al problema de la medida, pues al menos nos garantiza la existencia de extensiones de $lh$. Sin embargo, a continuación presentaremos una serie de resultados que colocan a la medida de Lebesgue en el centro del problema de la medida, pues demostraremos que cualquier extensión de $lh$ tiene que coincidir con la medida de Lebesgue en $\mathfrak{M}$. En particular, tenemos que cualquier extensión de $lh$ debe otorgarle a los borelianos su medida de Lebesgue.

Este resultado, que se puede entender como la unicidad de la medida de Lebesgue, coincide con la idea de que el tamaño geométrico es único y, por lo tanto, confirma que la medida de Lebesgue **realmente** es la forma natural de extender la noción de longitud, área, volumen, etc. Antes de demostrarlo, presentaremos un lema que nos será de gran utilidad.

**Lema 28** *Sea* $\upsilon$ *una extensión de lh y* $\mathfrak{L}$ *una* $\sigma-$*álgebra tal que* $\mathfrak{L} \subseteq dom(\upsilon)$ *entonces* $\upsilon |_{\mathfrak{L}}$ *es:*

(a) Monótona. Si $A, B \in \mathfrak{L}$ y $A \subset B$ entonces $\upsilon(A) \leq \upsilon(B)$.

(b) Subaditiva. Si $\{A_n\}_{n=1}^{\infty} \subset \mathfrak{L}$ y $A = \bigcup_{n\in\mathbb{N}} A_n$ entonces $\upsilon(A) \leq \sum_{n\in\mathbb{N}} \upsilon(A_n)$.

(c) Tiene límites inferiores. Si $A = \bigcup_{n\in\mathbb{N}} A_n$ con $A_n \subset A_{n+1} \forall n \in \mathbb{N}$, entonces $\upsilon(A) = \lim_{n\to\infty} \upsilon(A_n)$[13].

[13] Nótese que las primeras dos propiedades ya se habían demostrado para el caso particular de la medida de Lebesgue. Hubiéramos podido presentar estos resultados más generales antes y ahorrarnos demostraciones. Sin embargo, las demostraciones que se hacen sobre la medida de Lebesgue permiten tener una idea más clara de cómo se comporta esa medida particular.

**Demostración.** (a) Sean $A, B \in \mathfrak{L}$ tales que $A \subset B$. Como $\mathfrak{L}$ es una $\sigma-$álgebra, tenemos que $B - A = B \cap A^c \in \mathfrak{L}$. Además, $v$ es finito aditiva (pues por hipótesis es extensión de $lh$). Así que $v(B) = v(B - A) + v(A) \geq v(A)$.

(b) Sea $\{A_n\}_{n=0}^{\infty} \subset \mathfrak{L}$ y $A = \bigcup_{n \in \mathbb{N}} A_n$. Lo primero que haremos es ajenizar a los $A_n$'s recursivamente. Sea $B_0 = A_0$ y $B_n = A_n - \bigcup_{i=0}^{n-1} B_i$ para $n > 0$. Supongamos, sin pérdida de generalidad, que $l < j$, entonces $B_l \subset \bigcup_{i=1}^{j} B_i$. Sin embargo, por definición, $B_j = A_j - \bigcup_{i=0}^{j-1} B_i$. Así que $A = \bigcup_{n \in \mathbb{N}} B_n$, con $B_l \cap B_j = \emptyset$ si $l \neq j$. Además, como $B_n \subset A_n$, utilizando el inciso (a) tenemos que $v(B_n) \leq v(A_n)$. Así que $\sum_{n \in \mathbb{N}} v(B_n) \leq \sum_{n \in \mathbb{N}} v(A_n)$.

Por otro lado, como $v$ es extensión de $lh$, en particular es $\sigma-$aditiva, así que $v(\bigcup_{n \in \mathbb{N}} B_n) = \sum_{n \in \mathbb{N}} v(B_n) \leq \sum_{n \in \mathbb{N}} v(A_n)$. Pero como ya habíamos visto que $A = \bigcup_{n \in \mathbb{N}} B_n$, lo que en realidad tenemos es que $v(A) \leq \sum_{n \in \mathbb{N}} v(A_n)$.

(c) Lo primero es ajenizar los conjuntos como en el inciso anterior de manera que $B_0 = A_0$ y $B_n = A_n - \bigcup_{i=0}^{n-1} B_i$ para $n > 0$. Cabe señalar que $A_n = \bigcup_{i=0}^{n} B_i$ con $B_l \cap B_j = \emptyset$ si $l \neq j$. Así que usando la finito aditividad de $v$ tenemos que $v(A_n) = \sum_{i=0}^{n} v(B_i)$. Esto junto con la $\sigma-$ *aditividad* pemite ver que $v(A) = \sum_{n \in \mathbb{N}} v(B_n) = \text{lím}_{n \to \infty} \sum_{i=1}^{n} v(B_i) = \text{lím}_{n \to \infty} v(A_n)$. ■

**Teorema 29** *Si $v$ es una extensión de $lh$ tal que $\mathfrak{M} \subset dom(v)$, entonces $v|_{\mathfrak{M}} = \mu_L$.*

**Demostración.** Sea $A \in \mathfrak{M}$. Como $v$ es extensión de $lh$ tenemos que para $(p, q)$ $v((p, q)) = \mu_L((p, q)) = |p - q|$. Esto quiere decir que dada $A \in \mathfrak{M}$, si A es abierto entonces $v(A) = \mu_L(A)$. Los intervalos abiertos con extremos racionales forman una base para la topología, así que todo abierto es unión a lo más numerable de intervalos disjuntos, abiertos y con extremos racionales. De modo que usando la $\sigma-$aditividad y que tanto $v$ como $\mu_L$ son extensiones de $lh$ tenemos la igualdad para el caso de $A$ abierto. Supongamos que $A$ no es abierto, pero es acotado. Demostraremos primero que $v(A) \leq \mu_L(A)$. Sea $\mathcal{C} = \{(a_i, b_i)\}_{i \in \mathbb{N}}$ una cubierta abierta de A. Como $\mathfrak{M}$ es una $\sigma$-álgebra y $\forall i \in \mathbb{N}$, $(a_i, b_i) \in \mathfrak{M}$, entonces $\bigcup_{i \in \mathbb{N}} (a_i, b_i) = \mathcal{U} \in \mathfrak{M}$. Pero además $\mathcal{U}$ es abierto, así que por lo dicho antes $v(\mathcal{U}) = \mu_L(\mathcal{U})$ y utilizando el inciso (a) del lema anterior tenemos que $v(A) \leq v(\mathcal{U}) = \mu_L(\mathcal{U})$. Como eso sucede para toda cubierta abierta de $A$ tenemos, por la definición de la medida de Lebesgue (como el ínfimo de las medidas de la unión de las cubiertas abiertas) que $v(A) \leq \mu_L(A)$. Por otro lado, como $A$ es acotado, entonces $\mu_L(A) < \infty$. Así que dada $\epsilon > 0$ existe $\mathcal{C}$ una cubierta de A tal que $\mu_L(\mathcal{U}) < \mu_L(A) + \epsilon$. Como $\mathcal{U}$, $A \in \mathfrak{M}$ y $\mathfrak{M}$ es una

$\sigma$−álgebra tenemos que $\mathcal{U} - A = \mathcal{U} \cap A^c \in \mathfrak{M}$ y $\mu_L(\mathcal{U} - A) < \epsilon$. Usando que la igualdad se cumple para los conjuntos abiertos y la finito aditividad de $v$ tenemos:

$$\mu_L(A) \leq \mu_L(\mathcal{U}) = v(\mathcal{U}) = v(A) + v(\mathcal{U} - A) \leq v(A) + \mu_L(\mathcal{U} - A) < v(A) + \epsilon.$$

Pero eso sucede para toda $\epsilon > 0$, de modo que $\mu_L(A) \leq v(A)$.

Para el caso general, demostraremos directamente la igualdad, basándonos en los conjuntos acotados. Como $\mathbb{R}^1 = \bigcup_{i \in \mathbb{N}} [z_i, z_i + 1)$ con $z_i$ que pertenece a una numeración de $\mathbb{Z}$[14],entonces dada $B \in \mathfrak{M}$, sea $B_i = B \cap [z_i, z_i + 1)$. Claramente $B_i \in \mathfrak{M}$ y $B_i \subset [z_i, z_i + 1)$. Así que usando la subaditividad de la medida de Lebesgue tenemos que $\forall i \in \mathbb{N}$, $\mu_L(B_i) \leq 1$ y, por el resultado anterior, que $\mu_L(B_i) = v(B_i)$. Por otro lado, $B = \bigcup_{i \in \mathbb{N}} B_i$ con $B_i \cap B_j = \emptyset$ (si $i \neq j$), así que usando la $\sigma$−aditividad, tanto de $v$ como de $\mu_L$, tenemos finalmente que:

$$v(B) = \sum_{i \in \mathbb{N}} v(B_i) = \sum_{i \in \mathbb{N}} \mu_L(B_i) = \mu_L(B)$$

Por lo tanto, $\forall B \in \mathfrak{M}$, $v(B) = \mu_L(B)$. ■

Este resultado nos permite dar un paso importante dentro del problema de la medida, en el sentido de que nos garantiza la existencia de una única forma de hablar del tamaño geométrico, al menos de todos los conjuntos borelianos. Sin embargo, podemos decir, sin exagerar, que la contribución de la medida de Lebesgue a la solución del problema de la medida es "proporcional" al tamaño de su dominio. En este sentido, es posible demostrar, aunque aquí no lo haremos, que hay solamente $c$ conjuntos de Borel[15]. Es decir, $|\mathfrak{B}_{\mathbb{R}}| = c < 2^c = |\wp(\mathbb{R})|$. Esto no parece ser muy alentador, pues si el dominio de la medida de Lebesgue estuviera conformado sólo por los borelianos, entonces no podríamos, con los resultados que tenemos, garantizar que hay una única forma de medir el tamaño geométrico de una cantidad importante[16] de subconjuntos de reales. Sin embargo, a continuación daremos un resultado que no sólo demuestra que $|\mathfrak{M}| = |dom(\mu_L)| = 2^c$, sino que conjuntos, como el de Cantor, con el que introdujimos inicialmente el problema de la medida, pertenecen a los conjuntos Lebesgue medibles. Esto evidentemente hace aún más atractiva la medida de Lebesgue.

[14] La idea de trabajar con conjuntos acotados primero para después generalizarlo a todos los conjuntos ya se había trabajado con la medida de Banach, en la que también se utilizó una numeración de $\mathbb{Z}$ para generar una cubierta de intervalos disjuntos de $\mathbb{R}^1$ (ver Corolario 7).

[15] Devlin da una demostración de esto especialmente elegante (ver Devlin [7] p. 101).

[16] Tantos como toda la potencia de los reales.

## 3.3. $2^c$ conjuntos medibles

Recuérdese que el conjunto de Cantor se obtiene de partir sucesivamente cada intervalo en tres y remover el subintervalo abierto intermedio, es decir:

Sea $I = [0,1]$, $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$, $C_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$, . . ., $C_n = [0, \frac{1}{3^n}] \cup [\frac{2}{3^n}, \frac{3}{3^{n-1}}] \cup [\frac{2}{3^{n-1}}, \frac{7}{3^n}] \cup \ldots \cup [\frac{3^n-1}{3^n}, 1]$. El conjunto de Cantor, que denotaremos con $C$, es la intersección de todos los $C_n$. En ese sentido, como $C = \bigcap\limits_{n \in \mathbb{N}} C_n$, $C$ pertenece a la $\sigma-$álgebra de Borel y, por lo tanto, es Lebesgue medible. Además como para toda $n \in \mathbb{N}$, $C \subset C_n$ entonces $\mu_L(C) \leq \mu(C_n)$. Por otro lado, usando la finito aditividad de la medida de Lebesgue tenemos que para $n \in \mathbb{N}$, $\mu_L(C_n) = \left(\frac{2}{3}\right)^n$ y como $\lim_{n\to\infty} \left(\frac{2}{3}\right)^n = 0$, entonces $\mu_L(C) = 0$[17]. Nótese que el conjunto de Cantor es equipotente con los reales, ya que $\forall x \in C$ $(x = \sum\limits_{i \in \mathbb{N}} \frac{a_i}{3^i}$,con $a_i = 0$ ó $a_i = 2)$. Así que podemos tomar $f : C \to [0,1]$ tal que $\forall x \in C(f(x) = \sum\limits_{i \in \mathbb{N}} \frac{a_i}{2^i})$.Claramente $rang(f) \subset [0,1]$. Sólo falta ver que es suprayectiva. Pero eso se desprende inmediatamente al tomar la expansión binaria de los números que están en el intervalo [0,1].

**Observación 9** *Como $|C| = c$ entonces tenemos que hay $2^c$ subconjuntos del conjunto de Cantor.*

Una vez que hemos recordado la construcción del conjunto de Cantor y algunas de sus propiedades podemos entonces demostrar el siguiente teorema:

**Teorema 30** $|dom(\mu_L)| = 2^c$.

**Demostración.** Como $|\wp(C)| = 2^c$ entonces basta demostrar que $\wp(C) \subset dom(\mu_L)$. Sea $A \subset C$, entonces por la monotonía de $\mu^*$(Lema 16 ) tenemos que $\mu^*(A) \leq \mu^*(C) = 0$. Así que dada $\epsilon > 0$ existe $C = \{(a_i, b_i)\}_{i \in \mathbb{N}}$ una cubierta abierta de A tal que $\sum\limits_{i \in \mathbb{N}} |a_i - b_i| < \epsilon$. Si $\mathcal{U} = \bigcup\limits_{i \in \mathbb{N}} (a_i, b_i)$ tenemos que $\mu^*(\mathcal{U}) < \epsilon$. Tomamos entonces $G = \mathcal{U}$ y $F = \emptyset$. Nótese que $G$ y $F$ son conjuntos abierto y cerrado respectivamente que cumplen que $F \subset A \subset G$ y que $\mu^*(G - F) = \mu^*(G) < \epsilon$[18]. Por lo tanto, $A \in \mathfrak{M}$ y $\wp(C) \subset dom(\mu_L)$. ■

Desde una perspectiva conjuntista el hecho de que se tengan $2^c$ conjuntos Lebesgue medibles puede ser muy alentador, aunque cabe la posibilidad de que también se tengan $2^c$ conjuntos no Lebesgue medibles. De hecho, como la medida de Lebesgue cumple con las condiciones (1) a (4) de la hipótesis de la medida y como $ZF + BO_{\mathbb{R}} \vdash \neg HM$, entonces $ZF + BO_{\mathbb{R}} \vdash dom(\mu_L) \neq \wp(\mathbb{R}^1)$.

[17] Es posible construir conjuntos tipo Cantor que tengan medida positiva (ver Folland.[9] p. 41). Esto muestra que los conceptos topológicos, de cardinalidad y de medida son muy distintos, pues hay conjuntos densos en ninguna parte que tienen medida positiva y conjuntos con la cardinalidad de los reales que tienen medida cero.

[18] Nótese que esta demostración sirve para un resultado más general: *Todo subconjunto de un conjunto de medida cero es Lebesgue medible y tiene medida cero.* Por esta propiedad se dice que la medida de Lebesgue es completa.

Sin embargo, desde la perspectiva del análisis se puede ser un poco más optimista[19], pues bastante significativo resulta que los conjuntos con los que se pensó el problema de la medida, como el de Cantor, sean Lebesgue medibles.Lo que interesa no es cuántos conjuntos no medibles hay, sino si es posible establecer porqué no son medibles. En ese sentido, cabe notar que para el caso de la medida de Lebesgue es posible concluir, a partir de la existencia de un buen orden para los reales, algo más que $dom(\mu_L) \neq \wp(\mathbb{R}^1)$: se puede afirmar que el conjunto de Vitali, en particular, no es Lebesgue medible. En la demostración del Teorema 4, que se hizo por reducción al absurdo, partimos de la hipótesis de que $dom(\mu) = \wp(\mathbb{R}^1)$ y utilizando la $\sigma-$aditividad y la invarianza bajo traslaciones llegamos a una contradicción. Para el caso de la medida de Lebesgue, basta suponer que $V \in dom(\mu_L)$, pues como $dom(\mu_L)$ es una $\sigma-$álgebra eso quiere decir que todos los conjuntos que se utilizan para generar la contradicción ($V^p_{<1} = V_{trasp} \cap [0, 1-p)$ y $V^p_{>1} = V_{trasp} \cap [1-p, 2)$, $(V^p_{>1})_{trasp}$, $(V^p_{<1})_{tras1-p}$) son Lebesgue medibles. Así que tenemos un ejemplo de un conjunto no medible. Sin embargo, como ya se ha hecho notar, no es mucho lo que podemos decir de este conjunto, sino simplemente que existe si es que existe un buen orden para los reales.

El teorema de la unicidad de la medida de Lebesgue (Teorema 29) coloca el punto de partida un poco más adelante, pues la pregunta ya no es si es posible extender la noción de longitud, sino si es posible y bajo qué condiciones se puede extender la medida de Lebesgue. Es decir, el problema de la medida es equivalente a la pregunta: ¿Existe $v : \wp(\mathbb{R}^1) \to [0, \infty]$ tal que

(1) Si $A \in \mathfrak{M}$, entonces $v(A) = \mu_L(A)$.
(2) $v$ es invariante bajo traslaciones.
(3) $v$ es finito-aditiva.
(4) $v$ es $\sigma-$aditiva?

Esta pregunta es verdaderamente equivalente al problema de la medida en el sentido de que si la denotamos por $PM_L$ y a su respuesta afirmativa como $HM_L$, entonces, por el teorema de unicidad de la medida de Lebesgue y las propiedades que cumple dicha medida, tenemos que $HM \iff HM_L$. En cambio, si denotamos por $HM_{dom(\mu_L)}$a la hipótesis "$dom(\mu_L) = \wp(\mathbb{R}^1)$" por lo pronto sólo podemos garantizar que $HM_{dom(\mu_L)} \implies HM$. El regreso no parece ser necesariamente cierto. Sin embargo, dadas las ventajas que hemos señalado de la medida de Lebesgue, parece más sensato enfrentar el problema de la medida, tratando de buscar hipótesis bajo las cuales se pueda cumplir que "$dom(\mu_L) = \wp(\mathbb{R}^1)$". Es decir, lo que en realidad quisiéramos saber es qué es lo que produce que haya conjuntos no Lebesgue medibles. ¿Es acaso alguna de las propiedades (1)-(4) que cumple la medida de Lebesgue, o es una propiedad de los conjuntos no medibles, o una combinación de ambas?

[19] Desde la perspectiva del análisis es más importante saber qué clase de conjuntos son los no medibles que la cantidad de ellos que hay.

Ya habíamos señalado, a partir de la demostración de Vitali, que para la negación de la hipótesis de la medida se plantean tres sospechosos: la invarianza bajo traslaciones, la $\sigma-$aditividad y la existencia de un buen orden para los reales. Sin embargo, una vez acotado el problema de la medida al problema del dominio de la medida de Lebesgue, podemos descartar a la invarianza bajo traslaciones, pues a continuación demostraremos, sin usar que la medida de Lebesgue es invariante bajo traslaciones, que hay conjuntos no Lebesgue medibles. Cabe mencionar que aunque la demostración que daremos se refiere al caso de $\mathbb{R}^1$ se puede generalizar al caso de $\mathbb{R}^n$, de manera que efectivamente la invarianza bajo traslaciones resulta ser inocente al margen de la dimensión en que se trabaje.

## 3.4. Conjuntos de Bernstein

Lo que haremos ahora es introducir un tipo de conjuntos, llamados conjuntos de Bernstein, que comparten con el conjunto de Vitali no sólo el hecho de no ser Lebesgue medibles, sino también que su existencia está garantizada por el axioma de elección, en particular, por la hipótesis de que hay un buen orden para los reales[20]. El hecho de que los ejemplos de conjuntos no medibles que presentamos sean conjuntos que no conocemos explícitamente nos lleva a preguntarnos si es posible encontrar conjuntos explícitos que no sean Lebesgue medibles. Sin embargo, dejaremos esta pregunta pendiente, puesto que lo que ahora interesa es ver que la propiedad de la invarianza bajo traslaciones no es la responsable de que existan conjuntos no medibles.

**Definición 14** *Un conjunto $B \subset \mathbb{R}$ es de Bernstein si y sólo si para todo conjunto de reales $C$, cerrado y no numerable, se tiene que $B \cap C \neq \emptyset$ y $B^c \cap C \neq \emptyset$.*

El conjunto de Bernstein que daremos, como el que dimos de Vitali, se basa, al menos indirectamente, en una cantidad $c$ de elecciones, puesto que para formarlo se elige un elemento de cada miembro de una familia conformada por $c$ conjuntos no vacíos. Sin embargo, a diferencia de Vitali, todos los conjuntos que pertenecen a esa familia tienen, a su vez, una cantidad $c$ de elementos[21].

Sea $\mathcal{F}$ la familia de todos los subconjuntos de la recta real que son cerrados y no numerables. Lo primero que haremos, en el mismo tenor que con Vitali, será demostrar que hay $c$ subconjuntos con esas propiedades :

**Lema 31** *Si $\mathcal{F} = \{C \subset \mathbb{R}^1 : \bar{C} = C \wedge \aleph_0 < |C|\}$ entonces $|\mathcal{F}| = c$.*

**Demostración.** Sabemos que hay una cantidad numerable de intervalos abiertos, pues es posible identificarlos con los pares ordenados de $\mathbb{Q} \times \mathbb{Q}$, donde

[20] O como señala Oxtoby: "It is based on the possibility of well ordering a set of power $c$". (Oxtoby [19] p.23).

[21] En el caso de Vitali cada conjunto de la familia (i.e. cada clase de equivalencia) es numerable.

claramente $|\mathbb{Q} \times \mathbb{Q}| = \aleph_0$, pero además dichos intervalos forman una base para la topología usual de la recta real. Eso quiere decir que todo conjunto abierto es a lo más unión numerable de intervalos abiertos con extremos racionales y, por lo tanto, que se puede identificar con el subconjunto de naturales que sirven como índices de los intervalos que lo conforman. De modo que, si $A$ es la familia de los subconjuntos abiertos de reales, entonces $|A| \leq |\wp(\mathbb{N})| = c$. Por otro lado, como el complemento de todo conjunto cerrado es único y abierto tenemos que hay a lo más $c$ conjuntos cerrados. Es decir, si $\mathcal{H}$ es la familia de conjuntos cerrados, entonces $|\mathcal{H}| \leq |A| \leq c$. Y como $\mathcal{F} \subset \mathcal{H}$, entonces $|\mathcal{F}| \leq c$. La otra desigualdad se sigue del hecho de que hay $c$ intervalos cerrados con interior no vacío[22]. ■

Lo siguiente será demostrar que cada uno de los conjuntos de $\mathcal{F}$ contiene, a su vez, una cantidad $c$ de elementos. Este resultado se desprenderá de un teorema, que en una versión más débil suele conocerse como el teorema de Cantor-Bendixson y que es importante no sólo porque nos permite dar un conjunto de Bernstein, como veremos a continuación, sino porque garantiza que la hipótesis del continuo se cumple para los conjuntos de reales $G_\delta$ (o en su versión débil, para los conjuntos cerrados). Esto podría significar un gran avance dentro del problema del continuo, pues si se pudiera demostrar que cualquier subconjunto de reales es numerable o contiene un conjunto $G_\delta$ no numerable (o en la versión más débil: contiene un conjunto cerrado no numerable, en cuyo caso se dice que tiene la propiedad del conjunto perfecto), entonces se demostraría la hipótesis del continuo. Más adelante, cuando nos refiramos a los resultados de Solovay, nos detendremos con más detalle en esto. Sin embargo, podemos adelantar que, suponiendo el axioma de elección, hay conjuntos que no tienen la propiedad del conjunto perfecto, en particular que no la tienen los conjuntos de Bernstein. Esto confirma el carácter conflictivo del axioma de elección no sólo en el terreno de teoría de la medida, sino de la teoría de conjuntos.

**Teorema 32 (Cantor-Bendixson Fuerte)** *Todo conjunto de reales $G_\delta$ no numerable contiene un conjunto cerrado de cardinalidad de los reales*[23].

**Demostración.** Sea $A \subset \mathbb{R}^1$ un conjunto no numerable tal que $A = \bigcap_{n\in\mathbb{N}} A_n$ con $A_n$ abierto $\forall n \in \mathbb{N}$. La demostración del teorema consiste en la construcción de un conjunto tipo Cantor a partir de $A$. Sea $A'$ el conjunto de los puntos de acumulación de A y $A^o = \{x \in A \cap A' : \forall (a,b) \subset \mathbb{R}^1 (x \in (a,b) \rightarrow |(a,b) \cap A| > \aleph_0\}$, que se conoce como el conjunto de los puntos de condensación de A.

[22] Simplemente tómese los intervalos de la forma $[0,x]$, con $0<x<1$.

[23] El teorema de Cantor-Bendixson dice: *Todo conjunto cerrado y no vacío de reales contiene un conjunto cerrado y sin puntos aislados* (i.e.contiene un conjunto perfecto). Esto junto con otro teorema de Cantor: (*Todo conjunto perfecto tiene la cardinalidad de los reales*) nos garantiza la hipótesis del continuo para los cerrados. (ver Kanamori [16], p.133). Esto sería suficiente para dar el conjunto de Bernstein. Sin embargo, aquí presentamos una versión más fuerte del teorema de Cantor-Bendixson, para demostrar algunas propiedades relevantes de los conjuntos de Bernstein.

Afirmamos que:

a) $A^o \neq \emptyset$.

b) $A^o \subseteq (A^o)'$. Es decir, todos los puntos de $A^o$ son de acumulación.

El inciso (a) se demostrará por reducción al absurdo. Supongamos que $A^o = \emptyset$, eso quiere decir que $\forall x \in A\ \exists (a_x, b_x) \subset \mathbb{R}^1 [x \in (a_x, b_x) \wedge |(a_x, b_x) \cap A| \leq \aleph_0]$. De modo que $\bigcup_{x \in A} (a_x, b_x)$ es una cubierta abierta de $A$, pero como toda cubierta abierta de un conjunto de reales tiene una subcubierta numerable[24], tenemos que $A \subset \bigcup_{i \in \mathbb{N}} (a_{x_i}, b_{x_i})$, con $|(a_{x_i}, b_{x_i}) \cap A| \leq \aleph_0$. Así que $|A| \leq \sum_{i \in \mathbb{N}} |(a_{x_i}, b_{xi}) \cap A| \leq \aleph_0 \bullet \aleph_0 = \aleph_0$! (pues. por hipótesis, A es no numerable). De modo que $A^o \neq \emptyset$. El inciso (b) también se demostrará por reducción al absurdo: sea $x \in A^o$ un punto aislado, entonces existe $(a, b) \subset \mathbb{R}^1$ tal que $(a, b) \cap A^o = \{x\}$. Es decir, $\forall y \in (a, b) \cap A [y \neq x \implies y \in (A^o)^c]$, así que $\forall y \in (a, b) \cap A [y \neq x \implies \exists (a_y, b_y)(y \in (a_y, b_y) \wedge |(a_y, b_y) \cap A| \leq \aleph_0]$. Usando el mismo argumento que para el inciso (a) tendríamos que $|(a, b) \cap A| \leq \aleph_0$! (pues $x \in (a, b) \cap A^o \implies |(a, b) \cap A| > \aleph_0$, por definición de $A^o$). Por lo tanto, $A^o \subseteq (A^o)'$.

Una vez demostradas estas dos propiedades de $A^o$ podemos construir el conjunto tipo Cantor para concluir la demostración del teorema. Construiremos de manera recursiva intervalos con propiedades particulares. Ilustraremos el procedimiento con el caso base. Tomamos $I(0)$ e $I(1)$ intervalos cerrados (no unitarios) tales que:

(i) Sean ajenos. $I(0) \cap I(1) = \emptyset$.

(ii) Contengan puntos de condensación de A. $I(0) \cap A^o \neq \emptyset$ e $I(1) \cap A^o \neq \emptyset$.

(iii) Estén contenidos en el primer abierto de la intersección de abiertos que conforman a A. $I(0) \subset A_1$e $I(1) \subset A_1$.

(iv) Tengan una longitud determinada. $lh(I(0) < \frac{1}{3}$ y $lh(I(1) < \frac{1}{3}$.

Estos intervalos se pueden tomar ya que $A^o$ es distinto del vacío y no tiene puntos aislados, de modo que existe $x \in A^o$ y, si tomamos $I = (a, b)$ tal que $x \in I$ , existe $y \neq x \in I \cap A^o$. Sean $\delta = \frac{\text{mín}\{\frac{1}{3}, |x-y|\}}{2}$ , $I_x = (x - \delta, x + \delta)$ e $I_y = (y - \delta, y + \delta)$. Nótese que $I_x$ e $I_y$ son ajenos y su longitud es menor que $\frac{1}{3}$. Por otro lado, como $A^o \subset A \subset A_1$ y $A_1$ es abierto entonces existen $a_x, b_x \in \mathbb{R}$ tales que $a_x < x < b_x$ y $[a_x, b_x] \subset I_x \cap A_1$. Lo mismo sucede para $y$. Así que tomamos $I(0) = [a_x, b_x]$ e $I(1) = [a_y, b_y]$.

Este procedimiento se puede generalizar. Supongamos que para $n \in \mathbb{N}$ tenemos $2^n$ intervalos cerrados (no unitarios), cada uno asociado a una $n$-eada de ceros y unos tales que:

(i) Si $I(a_1, ..., a_n)$ e $I(b_1, ..., b_n)$ son tales que $a_k \neq b_k$ para alguna $k \leq n$, entonces $I(a_1, ..., a_n) \cap I(b_1, ..., b_n) = \emptyset$.

(ii)$I(a_1, ..., a_n) \cap A^o \neq \emptyset$.

(iii)$I(a_1, ..., a_n) \subset A_n \cap I(a_1, ..., a_{n-1})$.

(iv)$lh(I(a_1, ..., a_n)) < \frac{1}{3^n}$.

[24] Esto se desprende de un teorema general de topología que nos dice: *Dado $X$ un espacio topológico con base numerable, para toda cubierta abierta de un conjunto en $X$ existe una subcubierta numerable* (ver Kelly [17] p. 49.)

Entonces para $n+1$ tenemos $2^{n+1}$ intervalos que cumplen los incisos (i) a (iv) correspondientes a $n+1$. Para demostrarlo se aplica el caso base a cada uno de los intervalos del caso $n$. Es decir, de cada uno de los intervalos dados se extraen dos intervalos que cumplen con las condiciones por las mismas razones por las que se cumplía en el caso base. Nótese que esta construcción es completamente análoga a la del conjunto de Cantor.

Está claro que $C_n = \bigcup I(a_1, ..., a_n)$ es compacto, ya que es unión finita de $(2^n)$ intervalos cerrados. Además, por la propiedad (iii) tenemos que para toda $n \in \mathbb{N}$, $C_n \subset A_n$. Sea $C = \bigcap_{n\in\mathbb{N}} C_n$. $C$ es cerrado por ser intersección de cerrados y $C \subset A$, pues $C = \bigcap_{n\in\mathbb{N}} C_n \subset \bigcap_{n\in\mathbb{N}} A_n = A$. Afirmamos que existe una función biyectiva entre el intervalo $[0,1]$ y los elementos de $C$. Sea $x \in [0,1]$ entonces pensamos en la expansión binaria de $x$ como una función $s_x : \mathbb{N} \longrightarrow \{0,1\}$. De modo que dada $x \in [0,1]$ tomamos la sucesión de intervalos $\{I_n^x\}_{n=1}^{\infty}$ tal que $I_n^x = I(s_x(1), ..., s_x(n))$. Por la forma en la que construimos los intervalos, tenemos que $I_{n+1} \subset I_n$ y que $\lim_{n\to\infty} lh(I_n) = 0$. Esto quiere decir que tenemos una sucesión de intervalos anidados cuya longitud tiende a cero, así que su intersección es un punto. Es decir, existe $c \in \mathbb{R}$ tal que $\bigcap_{n\in\mathbb{N}} I_n^x = \{c\}$. Como $\bigcap_{n\in\mathbb{N}} I_n^x \subset \bigcap_{n\in\mathbb{N}} C_n$, entonces $c \in C$. Sea $g : [0,1] \to C$ tal que:

$$g(x) = \bigcap_{n\in\mathbb{N}} I_n^x$$

Nótese que $g$ es inyectiva, ya que si $x \neq y \Longrightarrow s_x \neq s_y \Longrightarrow \bigcap_{n\in\mathbb{N}} I_n^x \neq \bigcap_{n\in\mathbb{N}} I_n^x$. Para ver que $g$ es suprayectiva basta ver que todos los elementos de $C$ tienen una expresión binaria. Sea $c \in C$. Entonces, dada $n \in \mathbb{N}$, existe $s_n$ una sucesión de tamaño $n$ de ceros y unos tal que $c \in I(s_n)$. Afirmamos que $\forall m, n \in \mathbb{N}[c \in I(s_n) \cap I(s_m) \wedge m < n \Longrightarrow s_m \subset s_n]$. De modo que $s_c = \bigcup_{n\in\mathbb{N}} s_n$ es una sucesión infinita de ceros y unos que representa a c, ya que $c \in \bigcap_{n\in\mathbb{N}} I(s_n)$. Supongamos que existen $m < n$ y $l < m$ tales que $s_m(l) \neq s_n(l)$. En ese caso, por la propiedad (i) de la construcción de los intervalos tendríamos que $I(s_m) \cap I(s_n) = \emptyset$! (pues, por hipótesis, $c \in I(s_n) \cap I(s_m)$). Por lo tanto, $s_m \subset s_n$ y $s_c = \bigcup_{n\in\mathbb{N}} s_n$ es una sucesión infinita de ceros y unos. De modo que existe $x \in [0,1]$ tal que $s_x = s_c$. En otras palabras, existe $x \in [0,1]$ tal que $g(x) = \bigcap_{n\in\mathbb{N}} I_n^x = \bigcap_{n\in\mathbb{N}} I(s_n) = c$. ■

**Corolario 33** *Todo conjunto $G_\delta$ de reales es contable o tiene la cardinalidad de los reales*[25].

[25] En otras palabras, la hipótesis del continuo se cumple para los conjuntos $G_\delta$.

**Corolario 34** *Todo conjunto cerrado no numerable tiene la cardinalidad de los reales.*

**Demostración.** Por el teorema anterior basta demostrar que todo conjunto cerrado es un conjunto $G_\delta$. Sea A un conjunto cerrado, entonces dados $x \in A$ y $n \in \mathbb{N}$, tomamos $I_x^n = (x - \frac{1}{n}, x + \frac{1}{n})$. Nótese que $A_n = \bigcup_{x \in A} I_x^n$ es abierto, pues los $I_x^n$ lo son. Afirmamos que $A = \bigcap_{n \in \mathbb{N}} A_n$. Es muy fácil ver que $A \subset \bigcap_{n \in \mathbb{N}} A_n$, pues dada $x \in A$ para toda $n \in \mathbb{N}$, $x \in I_x^n \subset A_n$ . Para la otra contención, basta demostrar que $\bigcap_{n \in \mathbb{N}} A_n \subset A'$, pues por hipótesis A es cerrado. Sea $x \in \bigcap_{n \in \mathbb{N}} A_n$, entonces dada $n \in \mathbb{N}$ existe $x_n \in A$ tal que $|x - x_n| < \frac{1}{n}$. De modo que $\lim_{n \to \infty} x_n = x$ y, por lo tanto, $x \in A'$. ■

Nótese que, con este corolario, hemos demostrado que los conjuntos cerrados no numerables tienen la misma cantidad de elementos que el conjunto de reales, al margen de si la cardinalidad de este último conjunto es un ordinal o no. Es decir, hemos simplemente demostrado que todo subconjunto de reales cerrado no numerable es equipotente con los reales. Sin embargo, si suponemos que existe un buen orden para los reales este resultado no sólo nos garantiza que los conjuntos sobre los que tomamos los elementos para construir el conjunto de Bernstein son no vacíos, sino que además heredan un buen orden que servirá para la "construcción" del conjunto de Bernstein.

**Lema 35** *Supongamos que existe un buen orden para los reales, entonces existe un conjunto de Bernstein.*

**Demostración.** Sea $\mathcal{F}$ la familia de los conjuntos cerrados no numerables. Como estamos suponiendo que existe un buen orden para los reales, es legítimo pensar en el cardinal de los reales como un ordinal[26]. Así que $|\mathcal{F}| = c$ y, por lo tanto, los elementos de $\mathcal{F}$ se pueden indexar con los ordinales estrictamente menores que c[27]. Es decir, $\mathcal{F} = \{C_\alpha : \alpha \in c\}$ donde para todo $\alpha < c$, $C_\alpha \subset \mathbb{R}^1$ es cerrado y no numerable.

Por otro lado, denotaremos, una vez más, con $<_\mathbb{R}$ al buen orden supuesto de los reales. Entonces elegimos recursivamente dos elementos distintos de cada miembro de la familia $\mathcal{F}$. Sea $q_1 = \min <_\mathbb{R} C_1$ y $p_1 = \min <_\mathbb{R} C_1 - \{q_1\}$. En general, $\forall \alpha < c$ tomamos $q_\alpha = \min <_\mathbb{R} C_\alpha - (\{q_\beta : \beta < \alpha\} \cup \{p_\beta : \beta < \alpha\})$ y $p_\alpha = \min <_\mathbb{R} C_\alpha - (\{q_{\beta,} : \beta \leq \alpha\} \cup \{p_\beta : \beta < \alpha\})$. Nótese que estos mínimos siempre existen, pues por el corolario del teorema de Cantor-Bendixson Fuerte, tenemos que $\forall \alpha < c$, $|C_\alpha| = c$. De modo que $C_\alpha - (\{q_{\beta,} : \beta \leq \alpha\} \cup \{p_\beta : \beta < \alpha\}) \neq \emptyset$.

[26] En la Sección 1.1 se habló de las dos formas de considerar la cardinalidad: como equipotencia o mediante ordinales. También se señaló que la segunda presupone el principio del buen orden.

[27] Recuérdese que $c = \min\{\alpha \in OR : \alpha \sim \mathbb{R}\}$.

Afirmamos que $B = \{q_\alpha : \alpha < c\}$ es un conjunto de Bersntein. Por un lado, tenemos que $\forall\alpha < c$ $B \cap C_\alpha = \{q_\alpha\}$. Así que $B$ efectivamente intersecta a todos los conjuntos cerrados no numerables. Por otro, por la forma en que tomamos a $B$ tenemos que $\{p_\alpha : \alpha < c\} \subset B^c$, pero dicho conjunto también intersecta a todos los cerrados no numerables, ya que $\{p_\alpha : \alpha < c\} \cap C_\alpha = \{p_\alpha\}$. Así que $B^c \cap C_\alpha \neq \emptyset$. Con lo que queda demostrado que $B$ es un conjunto de Bernstein. ■

Nótese que el conjunto de Bernstein se forma usando el buen orden de los reales en dos niveles: para indexar a los conjuntos cerrados y no numerables y para definir dos elementos distintos de cada uno de ellos, de los cuales uno va a ser parte del conjunto. Cabe señalar, además, que el conjunto de Bernstein, como el de Vitali, no está dado explícitamente, ya que se basa en la existencia de un buen orden para los reales que nunca puede estar dado explícitamente[28].

A continuación veremos que es posible caracterizar la medida de Lebesgue de un conjunto a partir de la medida de Lebesgue de los conjuntos compactos que contiene, lo que nos será de gran utilidad para demostrar que los conjuntos de Bernstein no son Lebesgue medibles:

**Lema 36** **($\mu_L$ es regular)** *Si $A \in \mathfrak{M}$, entonces* $\mu_L(A) = \sup\{\mu_L(F) : F \subset A \wedge F$ *es compacto*$\}$.

**Demostración.** Primero cabe señalar que, puesto que $\mathfrak{M}$ es una $\sigma$-álgebra, los conjuntos compactos (que en particular son cerrados por el teorema de Heine-Borel) son Lebesgue medibles. De modo que el conjunto sobre el que se toma el supremo está bien definido. Por otro lado, gracias a la subaditividad de $\mu_L$, si $F \subset A$ , entonces $\mu_L(F) \leq \mu_L(A)$. Así que $\sup\{\mu_L(F) : F \subset A \wedge F$ es compacto$\} \leq \mu_L(A)$. Para la otra desigualdad, tomaremos una sucesión de conjuntos acotada por A. Sea $A_n = A \cap (-n, n)$. Obsérvese que si $m < n$, entonces $A_m \subset A_n \subset A$. Eso quiere decir que estamos en el caso (c) del Lema 28 y, por lo tanto, $\lim_{n\to\infty}\mu_L(A_n) = \mu_L(A)$. Tenemos dos casos:

(1) $A$ es acotado. Como $\mu_L(A) < \infty$, entonces dada $\epsilon > 0$ existe $m \in \mathbb{N}$ tal que $\mu_L(A) - \frac{\epsilon}{2} < \mu_L(A_m)$. Por otro lado, como $A_m$ es acotado y medible existen $G_m$ abierto y acotado y $F_m$ cerrado tales que $F_m \subset A_m \subset G_m$ y $\mu_L(G_m - F_m) < \frac{\epsilon}{2}$. Como $G_m$ es abierto y acotado entonces por el Lema 20 $\mu_L(G_m) - \mu_L(F_m) < \frac{\epsilon}{2}$ y por el Lema 16 $\mu_L(A_m) - \mu_L(F_m) < \frac{\epsilon}{2}$. Por lo tanto, existe $F_m$ cerrado y acotado tal que $F_m \subset A_m$ y $\mu_L(A_m) - \frac{\epsilon}{2} < \mu_L(F_m)$. Juntando todos los resultado tenemos que para toda $\epsilon > 0$, existe $F$ compacto $(F_m)$ tal que $\mu_L(A) - \epsilon < \mu_L(F)$. Así que $\mu_L(A) \leq \sup\{\mu_L(F) : F \subset A$ y $F$ es compacto$\}$.

[28] Nótese que para la construcción del conjunto de Bernstein en el fondo se utilizan **c** elecciones dependientes (pues el conjunto sobre el que se va eligiendo cambia según la elección que se haya hecho antes), a diferencia del conjunto de Vitali que simplemente usa una cantidad **c** de elecciones.

(2) Si $A$ no es acotado. Sea $I_{2j} = (j, j+1)$ e $I_{2j+1} = (-j, -j+1)$. Entonces tomamos $A_n = A \cap I_n$ y $\frac{\epsilon}{2^n}$, de modo que por el inciso anterior, para toda $n$ existe $F_n$ compacto tal que $F_n \subset A_n$ y $\mu_L(A_n) - \frac{\epsilon}{2^n} < \mu_L(F_n)$. Sea $C_k = \bigcup_{n=0}^{k} F_n$. Nótese que $C_k$ es compacto (por ser unión finita de compactos) y $C_k \subset A$. Además $\sum_{n=0}^{k}(\mu_L(A_n) - \frac{\epsilon}{2^n}) < \sum_{n=0}^{k} \mu_L(F_k)$ y como $I_n \cap I_m = \emptyset$ (si $n \neq m$), entonces $\mu_L(\bigcup_{n=0}^{k} A_n) - \epsilon < \mu_L(C_k)$. Finalmente $\mu_L(A) - \epsilon = \text{lím}_{n\to\infty} \mu_L(\bigcup_{n=0}^{k} A_n) - \epsilon \leq$ $\text{lím}_{n\to\infty} \mu_L(C_k) = \sup\{\mu_L(C_k) : k \in \mathbb{N}\} \leq \sup\{\mu_L(F) : F \subset A$ y $F$ es compacto$\}$. ■

**Teorema 37** *Sea $B$ un conjunto de Bernstein, entonces $B$ no es Lebesgue medible.*

**Demostración.** Supongamos que $B$ es Lebesgue medible, como para todo $C$ conjunto cerrado no numerable $B^c \cap C \neq \emptyset$, entonces $\forall F \subset B[F = \bar{F} \implies |F| \leq \aleph_0]$. Así que, por el Corolario 7, $\forall F \subset B[F = \bar{F} \implies \mu_L(F) = 0]$. Como todo compacto es cerrado, por el lema anterior tendríamos entonces que $\mu_L(B) = 0$. Por otro lado, como estamos suponiendo que $B \in \mathfrak{M} = dom(\mu_L)$, que ya se vio forma una $\sigma-$álgebra, entonces $B^c$ también sería Lebesgue medible y aplicando exactamente el mismo argumento que para $B$ concluiríamos que $\mu_L(B^c) = 0$. Utilizando la finito aditividad de $\mu_L$ tendríamos que $\mu_L(\mathbb{R}^1) = \mu_L(B) + \mu_L(B^c) = 0 + 0 = 0$! (pues $\mu_L(\mathbb{R}^1) = \infty$). Por lo tanto, los conjuntos de Bernstein no son Lebesgue medibles[29]. ■

Cabe señalar que los argumentos que hemos utilizado para ver que el conjunto de Bernstein no es Lebesgue medible se aplican a cualquier medida que extienda a $lh$ y sea regular.

Los conjuntos de Bernstein se conocen también como *conjuntos universalmente no medibles* (ver Stromberg [22]) y aunque aquí no lo hemos probado, permiten ver que todo conjunto de medida positiva contiene un conjunto no medible (Oxtoby [19] p.24).

Podemos resumir nuestros resultados de la siguiente manera:

$$\begin{array}{l} \text{ZF} + \text{BO}_{\mathbb{R}} \vdash \exists B(B \text{ es un conjunto de Bernstein}) \\ \text{ZF} \quad \vdash \quad \forall B[B \text{ es un conjunto de Bernstein} \implies B \notin dom(\mu_L)] \\ \text{ZF} + \text{BO}_{\mathbb{R}} \vdash dom(\mu_L) \neq \wp(\mathbb{R}^1). \end{array}$$

Así que hemos dado una demostración de la imposibilidad de que se cumpla la hipótesis de la medida que no emplea la invarianza bajo traslaciones que se puede generalizar a todas las dimensiones.

---

[29] Nótese que la invarianza bajo traslación no se utilizó, mientras que la $\sigma-$aditividad fue pieza clave para probar la regularidad de la medida de Lebesgue (Lema 36).

Con el conjunto de Cantor pudimos concluir que hay $2^c$ conjuntos medibles, pues dicho conjunto tiene la cardinalidad de los reales y todos sus subconjuntos son Lebesgue medibles. Por otro lado, hemos señalado ya que nada impide que a su vez se tengan $2^c$ conjuntos no medibles. Esto lo probaremos de manera análoga a como se trabajó con el conjunto de Cantor, aunque teniendo como base el conjunto de Bernstein que se construyó en el Lema 35. Sin embargo, la argumentación será un poco menos directa. Primero se demostrará que el conjunto de Bernstein tiene al menos tantos subconjuntos no medibles como medibles. Sea $B$ un conjunto de Bernstein y sean $\mathrm{B}_m = \{A \subset B : A \in \mathfrak{M}\}$ y $\mathrm{B}_m^c = \{A \subset B : A \notin \mathfrak{M}\}$. Es decir, $\mathrm{B}_m$ es la familia de los subconjuntos de Bernstein Lebesgue medibles y $\mathrm{B}_m^c$ la de los no medibles, entonces tenemos el siguiente resultado:

**Teorema 38** $|B_m| \leq |B_m^c|$

**Demostración.** Sea $A \in B_m$, entonces $B - A$ no es Lebesgue medible, pues de otro modo, como $A \subset B \Longrightarrow B = (B - A) \cup A$, entonces $B$ sería medible! De modo que $B - A \in B_m^c$. Definimos la función $g : B_m \longrightarrow B_m^c$ de la siguiente manera: $\forall A \in B_m$,

$$g(A) = B - A$$

Esta función es inyectiva, pues si $A,\ A' \in B_m$ son tales que $g(A) = g(A')$ entonces $B - A = B - A'$. Como $A \subset B$ y $A' \subset B$ entonces $A = A'$. Por lo tanto, $g$ es inyectiva y $|B_m| \leq |B_m^c|$. ■

**Corolario 39** *Hay $2^c$ conjuntos no medibles.*

**Demostración.** Como $|\wp(B)| = 2^c$ y $\wp(B) = B_m \cup B_m^c$, usando aritmética cardinal y el teorema anterior, tenemos que $|\wp(B)| = \max\{|B_m|, |B_m^c|\} = |B_m^c| = 2^c$. ■

Como la demostración de este teorema se basa en el resultado $\mathrm{ZF} + \mathrm{BO}_{\mathbb{R}} \vdash \exists B(B$ es un conjunto de Bernstein) (Lema 35), entonces está claro que suponiendo que existe un buen orden para los reales hay muchos más conjuntos no medibles que simplemente el conjunto de Vitali y el de Bernstein[30]. Así que los argumentos de cardinalidad respecto a los conjuntos medibles y no medibles no ayudan a tener un panorama más claro sobre el problema de la medida. Por otro lado, este resultado se parece al de Cantor en el sentido de que se tienen $2^c$ conjuntos no medibles, pero que en realidad son todos subconjuntos de un conjunto particular, que en este caso es el de Bernstein. Lo interesante de eso es que, aunque tenemos $2^c$ conjuntos no medibles, ninguno de ellos está dado explícitamente, sino que todos comparten que su existencia está garantizada por el axioma de elección en su versión $BO_{\mathbb{R}}$.

[30] De hecho, no sólo hay una gran cantidad de conjuntos no medibles, sino también una gran variedad. Halmos ([11] p. 70) presenta un ejemplo muy interesante de un conjunto no medible tal que la medida exterior de su intersección con cualquier conjunto medible es 1, pero la medida interior (ver definición en [11] p. 58) es 0.

# Capítulo 4

# Cardinalidad y medida

## 4.1. Teoremas de Ulam

Como ya hemos dicho antes, la demostración de la negación de la hipótesis de la medida mediante los conjuntos de Bernstein no hace uso de la invarianza bajo traslaciones y se puede generalizar a cualquier dimensión. Sin embargo, es importante señalar que se basa en propiedades topológicas de $\mathbb{R}^1$, como se ve con el uso del Lema 36. A continuación veremos un par de demostraciones de la negación de la hipótesis de la medida que no sólo no hacen uso de la invarianza bajo traslaciones, ni de propiedades topológicas, sino que establecen un vínculo importante entre propiedades de cardinalidad y de medida. Los resultados que presentamos muestran que el problema de la medida poco tiene que ver con las propiedades euclideanas de la hipótesis de la medida y que, bajo hipótesis conjuntistas particulares, el problema tiene un carácter más general. En el sentido de que para aquellos conjuntos que comparten ciertas propiedades de cardinalidad no se cumple una variante de la hipótesis generalizada de la medida. Para ello, sin embargo, necesitamos introducir primero un resultado que además de ser de gran utilidad en lo que sigue, pone de manifiesto, una vez más, el papel central que ocupan los conjuntos acotados y de medida positiva:

**Lema 40** *Dado un conjunto $A$ de medida finita, toda familia de subconjuntos de $A$ con medida positiva y ajenos dos a dos es a lo más numerable.*

**Demostración.** Sea $A \in \mathfrak{M}$ tal que $\mu_L(A) < \infty$. Supongamos que existe $\mathcal{F} = \{A_i \subset A : \mu_L(A_i) > 0 \wedge A_j \cap A_i = \emptyset \text{ si } i \neq j\}$ tal que $|\mathcal{F}| > \aleph_0$. Sea $B_n = \{A_i \in \mathcal{F} : \mu_L(A_i) > \frac{1}{n}\}$. Nótese que $\mathcal{F} = \bigcup\limits_{n \in \mathbb{N}} B_n$, $B_n \subset B_{n+1}$. Por lo tanto, $|\mathcal{F}| = \sup\{|B_n| : n \in \mathbb{N}\}$. Así que existe $n_0 \in \mathbb{N}$ tal que $|B_{n_0}| > \aleph_0$. De lo contrario, $|\mathcal{F}| \leq \aleph_0$! Sea $\{A_j\}_{j \in \mathbb{N}}$ una sucesión numerable de elementos de $B_{n_0}$. Como $B_{n_0} \subset \mathcal{F}$, entonces los elementos de la sucesión son ajenos dos a dos y, por lo tanto, $\mu_L(\bigcup\limits_{j \in \mathbb{N}} A_j) = \sum\limits_{j \in \mathbb{N}} \mu_L(A_j) > \sum\limits_{j \in \mathbb{N}} \frac{1}{n_0} = \infty$. Pero $\bigcup\limits_{j \in \mathbb{N}} A_j \subset \bigcup B_{n_0} \subset A$, así que $\mu_L(A) = \infty$! Pues, por hipótesis, A tiene medida finita. ∎

**Observación 10** *Nótese que este resultado se puede generalizar a cualquier medida, ya que en la demostración sólo se utilizó la $\sigma$ − aditividad, propiedad que comparten, por definición, todas las medidas (ver Definición 2).*

### 4.1.1. Hipótesis del continuo y el problema de la medida

El primer resultado está vinculado con la hipótesis del continuo, así que vale la pena recordar que $2^{\aleph_0} = |\mathbb{R}|$ y que dicha hipótesis establece que $2^{\aleph_0} = \aleph_1$, donde $\aleph_1$ es el primer cardinal no numerable o el cardinal sucesor de $\aleph_0$[1]. La hipótesis del continuo ha sido bastante controversial y algunos matemáticos, como Gödel y Cohen[2], creen que es falsa. Sin embargo, como se ha dicho antes, estos resultados son importantes por el vínculo, central para los resultados de Solovay, que se establece entre cardinalidad y medida.

**Lema 41** $HC \Longrightarrow BO_{\mathbb{R}}$[3].

**Demostración.** Como por hipótesis $|\mathbb{R}| = |(0,1)| = \aleph_1$, eso quiere decir que existe $g : (0,1) \longrightarrow \aleph_1$ tal que g es biyectiva. Entonces definimos la siguiente relación de orden sobre el intervalo $(0,1)$ :

$$\forall x, y \in (0,1)\ [x <_g y \quad \text{ssi} \quad g(x) \in g(y)].$$

Con esta definición de orden para los reales, tendríamos que $\langle (0,1), <_g \rangle \cong \langle \aleph_1, \in \rangle$. Recuérdese además que los ordinales están bien ordenados mediante la relación de pertenencia y, como la propiedad de ser buen orden se preserva bajo isomorfismos, entonces $\langle \mathbb{R}, <_g \rangle$ es un buen orden. ■

**Teorema 42 (Ulam-Hipótesis del Continuo)** *Si la hipótesis del continuo es verdadera entonces $dom(\mu_L) \neq \wp(\mathbb{R}^1)$.*

**Demostración.** Lo haremos por una cadena larga de reducciones al absurdo. Supongamos la hipótesis del continuo y supongamos que $dom(\mu_L) = \wp(\mathbb{R}^1)$. Trabajaremos sobre el intervalo $I = (0,1)$, ya que $|I| = |\mathbb{R}|$, pero $\mu_L$ restringida a $I$ es una medida finita. Usando la notación y el contenido del lema anterior, $\forall y \in (0,1)$ sea $M_y = \{x \in (0,1) : x <_g y\}$. Se puede pensar a $M_y$ como el segmento inicial determinado por $y$, según el orden $<_g$. Nótese que $<_g$ es el orden de $\aleph_1$ calcado sobre $\mathbb{R}$, de modo que no sólo es un orden discreto, sino que $\langle M_y, <_g \rangle \cong \langle g(y), \in \rangle$, con $g(y) \in \aleph_1$. Así que por la definición de $\aleph_1$, tenemos que $|M_y| = |g(y)| \leq \aleph_0 = |\mathbb{N}|$ y, por lo tanto, existe $f_y$ una función inyectiva entre $M_y$ y $\mathbb{N}$. Para $n \in \mathbb{N}$ y $t \in (0,1)$ definimos los siguientes conjuntos:

$$A_t^n = \{y \in (0,1) : t \in M_y \wedge f_y(t) = n\}.$$

Nótese que si dejamos fija a $n$, entonces $\forall t, s \in (0,1)[s \neq t \Longrightarrow A_t^n \cap A_s^n = \emptyset]$, pues $y \in A_t^n \cap A_s^n \Longrightarrow f_y(t) = n = f_y(s)$ y como $f_y$ es inyectiva, entonces $y \in A_t^n \cap A_s^n \Longrightarrow t = s$.

[1] Por definición, $\aleph_1 = \{\alpha \in OR : |\alpha| \leq \aleph_0\}$.
[2] Ver Campero [6] p. 96.
[3] De hecho, se tiene un resultado más fuerte: $HGC \Rightarrow PBO$ (ver José Alfredo Amor [1]).

Podemos pensar los conjuntos $A_t^n$ como las entradas de una matriz infinita de $\aleph_0$ renglones por $\aleph_1$ columnas, en la que además, por lo dicho antes, todos los conjuntos de cada renglón son ajenos dos a dos:

$$\begin{array}{ccccc} A_{t_1}^1 & A_{t_2}^1 & \dots & A_{t_n}^1 & \dots \\ A_{t_1}^2 & A_{t_2}^2 & \dots & A_{t_n}^2 & \dots \\ \dots & \dots & \dots & \dots & \dots \\ A_{t_1}^m & A_{t_2}^m & \dots & A_{t_n}^m & \dots \\ \dots & \dots & \dots & \dots & \dots \end{array}$$

Cabe notar que para $t \in (0,1)$, si $C_t$ denota la unión de la columna correspondiente, entonces $C_t = \bigcup_{n\in\mathbb{N}} A_t^n = (0,1) - M_t \cup \{t\}$. Esto es porque dada $s \in (0,1)$: $s \in C_t \iff \exists m \in \mathbb{N}(s \in A_t^m) \iff \exists m \in \mathbb{N}(t \in M_s \wedge f_s(t) = m) \iff t \in M_s = dom f_s \iff t <_g s \iff s \not\leq_g t \iff s \neq t \wedge s \notin M_t$.

Por lo tanto, $C_t = (0,1) - M_t \cup \{t\}$.

Para toda $t \in (0,1)$, como $|M_t| \leq \aleph_0$, $|M_t \cup \{t\}| \leq \aleph_0$. Así que lo anterior se traduce en: la *unión* de *toda columna de la matriz es igual al intervalo (0,1) salvo por un conjunto numerable*.

Por otro lado, como la medida de Lebesgue es finito-aditiva entonces dada $t \in (0,1)$, $\mu_L(C_t) = \mu_L((0,1)) - \mu_L(M_t \cup \{t\}) = 1$. pues todo conjunto numerable tiene medida 0 (ver Lema 7).

Lo siguiente será demostrar que existe $t \in (0,1)$ tal que $\forall n \in \mathbb{N}$, $\mu_L(A_t^n) = 0$. Es decir, que hay una columna que tiene como entradas sólo conjuntos de medida cero. Sea $A_{MP} = \{A_t^n : t \in (0,1) \wedge n \in \mathbb{N} \wedge \mu_L(A_t^n) \neq 0\}$ la familia de las entradas de la matriz que tienen medida positiva. Supongamos, contrario a lo que queremos demostrar, que toda columna tiene al menos una entrada con medida de Lebesgue positiva (i.e.$\forall t \in (0,1)\ \exists n \in \mathbb{N}\ \mu_L(A_t^n) \neq 0$). Como tenemos $\aleph_1$ columnas, lo anterior quiere decir que en la matriz hay $\aleph_1$ entradas con medida de Lebesgue positiva (i.e. $|A_{MP}| = \aleph_1$).

Por otro lado, como hay una cantidad numerable de renglones, entonces hay al menos un renglón que tiene $\aleph_1$ entradas de medida positiva. Sin embargo, hemos visto que las entradas de todo renglón son subconjuntos del intervalo (0,1) ajenos dos a dos, así que estaríamos contradiciendo el Lema 40 pues la medida del intervalo es finita.

Estas contradicciones provienen de suponer que $\forall t \in (0,1)\ \exists n \in \mathbb{N}\ (\mu_L(A_t^n) \neq 0)$. De modo que efectivamente hay una columna que tiene como entradas sólo conjuntos de medida cero (i.e. $\exists t_\alpha \in (0,1) \forall n \in \mathbb{N}\ [\mu_L(A_{t_\alpha}^n) = 0]$).

Afirmamos que $\mu(C_{t_\alpha}) = 0$, pues $C_{t_\alpha} = \bigcup_{n\in\mathbb{N}} A_{t_\alpha}^n$ con $\mu_L(A_{t_\alpha}^n) = 0$ y aunque no podemos aplicar directamente la $\sigma-$aditividad, ya que las entradas de una columna no son necesariamente ajenas, podemos llevar a cabo el proceso de ajenización que ya hemos usado antes[4] y tener que $C_{t_\alpha} = \bigcup_{n\in\mathbb{N}} B_{t_\alpha}^n$, con $B_{t_\alpha}^n \cap B_{t_\alpha}^m = \emptyset$ y $B_{t_\alpha}^n \subset A_{t_\alpha}^n$. De modo que usando que la medida de Lebesgue es monótona tendríamos que $\mu_L(B_{t_\alpha}^n) = 0$. Ahora sí, por la $\sigma - aditividad$, llegaríamos a que $\mu_L(C_{t_\alpha}) = 0$.

[4] Ver demostración del inciso (c) del Lema 28.

Esto contradice, a su vez, el hecho establecido, al principio de la demostración, de que $\forall t \in (0,1)$, $\mu_L(C_t) = 1$. Con lo que llegamos a la última contradicción que se desprendía de suponer que $dom(\mu_L) = \wp(\mathbb{R}^1)$. Por lo tanto, $HC \Longrightarrow dom(\mu_L) \neq \wp(\mathbb{R}^1)$. ■

Como el inverso del Lema 41 no se cumple[5], queda claro que suponer la hipótesis del continuo es algo más fuerte que suponer simplemente que existe un buen orden para los reales, como se había hecho hasta ahora. Sin embargo, aún en esta demostración, la existencia de un buen orden para los reales, que queda contenida en la hipótesis del continuo, juega un papel central. Recuérdese que los conjuntos $A_t^n$ se definieron a partir del orden particular, heredado de $\aleph_1$, que tendrían los reales si se asume la hipótesis del continuo.

Por otro lado, como en esta demostración no sólo no se utilizó la invarianza bajo traslaciones, sino que tampoco se usó el hecho de que la medida de Lebesgue fuera extensión de la noción de longitud (salvo porque la medida del intervalo (0,1) es finita), queda claro que este resultado tiene una forma más general conectada al problema generalizado de la medida[6]:

**Teorema 43** *Dada $X$ tal que $|X| = \aleph_1$ y $\mu : \wp(X) \to \mathbb{R}(\forall x \in X, \mu(\{x\}) = 0 \wedge \mu(X) < \infty)$ entonces $\mu(X) = 0$.*

La demostración es completamente análoga a la dada para el caso de la medida de Lebesgue[7], lo que explica que se pida que la medida sea finita y que en los unitarios sea cero [8], ya que estas son las únicas dos propiedades de la medida de Lebesgue que se emplean en la demostración del Teorema 42.

Es probable que llame la atención que no se haya utilizado este resultado para ver la inocencia de la invarianza bajo traslaciones, en lugar del dado por los conjuntos de Bernstein, que implicó un arduo camino en la construcción de la medida de Lebesgue. Sin embargo, habría que enfatizar que se hizo de este modo por tres razones: (1) porque, como ya hemos señalado antes, suponer que se cumpla la hipótesis del continuo es algo más fuerte que simplemente suponer la existencia de un buen orden para los reales, (2) porque los conjuntos de Bernstein son otro ejemplo de conjuntos no Lebesgue medibles que, junto con Vitali, hacen pensar que los conjuntos no medibles se caracterizan por desprenderse del axioma de elección, (3) porque lo que más nos interesa de este resultado no es la inocencia de la invarianza bajo traslaciones, sino resaltar cómo las hipótesis de cardinalidad influyen en resultados de medida.

[5] Con forcing se puede demostrar que ZFC + $\neg HC$ es consistente (ver Capítulo 7 de Kunen [18]). En particular, $BO_{\mathbb{R}}$ es consistente con $\neg HC$. Así que $BO_{\mathbb{R}} \not\Rightarrow HC$.

[6] Recuérdese que una medida sólo debe cumplir que su dominio sea una $\sigma-$álgebra y que sea $\sigma-$aditiva (ver definición 2).

[7] Aquí la omitiremos, pero puede consultarse en Oxtoby [19] p. 25.

[8] Por la $\sigma-$*aditividad* basta con que una medida valga 0 en los unitarios, para que cualquier conjunto numerable tenga medida 0.

### 4.1.2. Cardinales inaccesibles y el problema de la medida

A continuación se dará un resultado más fuerte, por ser más general, sobre la relación entre cardinalidad y medida. Cabe señalar desde ahora que la demostración, aunque es parecida a la del Teorema 42, introduce nuevos elementos, como la noción de cardinal inaccesible (que aparecerá más tarde con los resultados de Solovay) o la generalización de la $\sigma-$ *aditividad* de una medida. En ese sentido, es importante mantener presente la noción generalizada de medida que se dió en el Capítulo 1, pues la demostración del teorema que veremos se basa, entre otras cosas, en la construcción de una medida para un cardinal inaccesible $\kappa$[9].

Sea $CAR$ la clase de los cardinales infinitos.

**Definición 15** *Un cardinal $\kappa$ es regular si y sólo si no existe $\lambda \in CAR$ tal que $\lambda < \kappa$ , $\kappa = \bigcup_{\eta<\lambda} \xi_\eta$ y $|\xi_\eta| < \kappa \ \forall \eta < \lambda$* [10].

**Definición 16** *Un cardinal $\kappa$ es inaccesible si y sólo si $\kappa$ es un cardinal no numerable, límite y regular*[11].

Dadas estas definiciones podemos enunciar el teorema generalizado de Ulam como:

**Teorema 44 (Ulam-Cardinales Inaccesibles)** *Si no existe $\kappa$ cardinal inaccesible tal que $\kappa \leq 2^{\aleph_0}$, entonces $dom(\mu_L) \neq \wp(\mathbb{R}^1)$.*

Lo que se hará es demostrar la contrapuesta[12]. Es decir, se demostrará que si $dom(\mu_L) = \wp(\mathbb{R}^1)$ entonces existe $\kappa$ cardinal inaccesible tal que $\kappa \leq 2^{\aleph_0}$. La idea de la demostración es: primero, utilizar la equipotencia entre $2^{\aleph_0}$ y (0,1) para definir una medida sobre $\wp(2^{\aleph_0})$ a partir de la medida de Lebesgue; segundo, obtener un cardinal $\kappa$ que cumpla con ciertas propiedades relacionadas con la medida definida y, tercero, demostrar que dicho cardinal es inaccesible.

Antes de adentrarnos en la demostración daremos una definición que caracteriza a las medidas de acuerdo con ciertas propiedades aditivas:

**Definición 17** *Sea $\kappa$ un cardinal y $\mu$ una medida. $\mu$ es $\kappa-$aditiva si y sólo si para toda $A$ tal que $A = \bigcup_{\eta<\lambda} A_\eta$, con $\lambda < \kappa$ y $\mu(A_\eta) = 0$ para toda $\eta < \lambda$, se tiene que $\mu(A) = 0$.*

---

[9] De hecho, éste será un ejemplo de la utilidad de la definición de medida para espacios que no necesariamente son euclideanos (ver Definición 2).

[10] Es decir, $\kappa$ es regular si y sólo si no puede ser expresado como unión de menos que $\kappa$ conjuntos, cada uno de cardinalidad menor que $\kappa$. Por ejemplo, cualquier cardinal sucesor es regular (ver Hernández Hernández [13] p. 270).

[11] $\aleph_0$ es un cardinal límite regular. Así que se pide que el cardinal sea no numerable para discutir la existencia o no de otros cardinales regulares límite.

[12] Que en realidad es como se suele expresar este resultado. Sin embargo, la forma en que aquí se planteó sirve como introducción al trabajo de Solovay, pues invita a preguntarse qué pasa con el problema de la medida si hay un cardinal inaccesible menor igual que $2^{\aleph_0}$.

**Lema 45** *Toda medida es $\aleph_1 - aditiva$.*

**Demostración.** Sea $A = \bigcup_{n\in\mathbb{N}} A_n$ con $\mu(A_n) = 0$. Si los $A_n$ no son ajenos, se pueden ajenizar de la manera canónica para tener que $A = \bigcup_{n\in\mathbb{N}} B_n$ con los $B_n$ ajenos y $\mu(B_n) = 0$. Entonces se usa la $\sigma-aditividad$ de la medida para concluir que $\mu(A) = 0$. ∎

**Observación 11** *Nótese que la $\sigma$ – aditividad y la $\aleph_1$ – aditividad no son propiedades equivalentes. Una función definida en una $\sigma-$álgebra puede ser $\aleph_1$ – aditiva, pero no $\sigma$ – aditiva. Por ejemplo, tómese la función $\mu$ igual a la medida de Lebesgue excepto por el valor que le da a $(-\infty,\infty)$, que digamos sea 1. Claramente $\mu$ no es $\sigma-$aditiva, pero sí es $\aleph_1-$aditiva, ya que la recta no se puede expresar como unión numerable de subconjuntos de medida de Lebesgue igual a cero.*

**Observación 12** *La $\kappa-$aditividad se refiere a algo así como la $\lambda-$aditividad de los conjuntos de medida cero para toda $\lambda < \kappa$ y, a diferencia de la $\sigma$ – aditividad, no pide que los conjuntos de la unión sean disjuntos.*

**Demostración.** ( **Del teorema 44**) Supongamos que $dom(\mu_L) = \wp(\mathbb{R}^1)$. Sea g una biyección entre $2^{\aleph_0}$ y el intervalo (0,1), entonces definimos $\mu : \wp(2^{\aleph_0}) \longrightarrow [0,1]$ tal que $\forall B \subseteq 2^{\aleph_0}$ :

$$\mu(B) = \mu_L(g[B]).$$

Se puede ver fácilmente que $\mu$ es una medida, pues como está definida en toda la potencia de $2^{\aleph_0}$, que es una $\sigma-$álgebra, y como además $\mu(2^{\aleph_0}) = 1$, sólo hay que verificar que es $\sigma-$aditiva. Sea $B = \bigcup_{n\in\mathbb{N}} B_n$ con $B_n \cap B_m = \emptyset$ si $m \neq n$. Como g es biyectiva entonces $g[B_n] \cap g[B_m] = \emptyset$, así que usando la $\sigma-$aditividad de la medida de Lebesgue tendríamos que $\mu_L(\bigcup_{n\in\mathbb{N}} g[B_n]) = \sum_{n\in\mathbb{N}} \mu_L\,(g[B_n])$. Pero está claro que $\bigcup_{n\in\mathbb{N}} g[B_n] = g[\bigcup_{n\in\mathbb{N}} B_n]$, así que $\mu_L(g[\bigcup_{n\in\mathbb{N}} B_n] = \sum_{n\in\mathbb{N}} \mu_L\,(g[B_n])$ y, por lo tanto, $\mu(B) = \sum_{n\in\mathbb{N}} \mu\,(B_n)$.

Tomemos ahora la siguiente colección:

$$C = \{\kappa \in CAR : \mu \text{ es } \kappa - aditiva\}.$$

Lo primero que hay que señalar es que $C$ es no vacía, pues como $\mu$ es medida entonces es $\aleph_1 - aditiva$ (Lema 45).

Por otro lado, si $(2^{\aleph_0})^+$ denota el cardinal sucesor de $2^{\aleph_0}$, tenemos que $C$ está acotada por $(2^{\aleph_0})^+$. Esto es porque $2^{\aleph_0} = \bigcup_{\alpha\in 2^{\aleph_0}} \{\alpha\}$ y $\forall\alpha \in 2^{\aleph_0}$ $\mu(\{\alpha\}) = \mu_L(g(\alpha)) = 0$, pero $\mu(2^{\aleph_0}) = \mu_L((0,1)) = 1$. Cabe señalar que $\forall\kappa \in CAR$, si $(2^{\aleph_0})^+ \leq \kappa$ entonces $\mu$ no es $\kappa - aditiva$, pues de otro modo $\mu$ sería $(2^{\aleph_0})^+ -$ *aditiva*. Así que $C \subseteq (2^{\aleph_0})^+$ . Esto quiere decir que $C$ es un conjunto de cardinales y no una clase. Por lo tanto, $\sup C = \bigcup C \in CAR$[13].

[13] Es muy fácil ver que la unión de cualquier conjunto de cardinales es cardinal.

Sea $\kappa_0 = \sup C$. Nótese que, como $\aleph_1 \in C$ y $C \subsetneq (2^{\aleph_0})^+$, $\aleph_1 \leq \kappa_0 \leq 2^{\aleph_0}$.

Además $\kappa_0 \in C$ . Si $\kappa_0$ es un cardinal sucesor[14], entonces trivialmente $\kappa_0$ pertenece a $C$, ya que si $\kappa_o = \lambda^+$ y $\kappa_0 \notin C$ entonces $\lambda$ es cota superior de $C$, contradiciendo que $\kappa_0 = \sup C$. Si es un cardinal límite, dada $\lambda < \kappa_0$ existe $\kappa \in C$ tal que $\lambda < \kappa$. En ese caso dada $A = \bigcup_{\eta<\lambda} A_\eta$, si $\mu(A_\eta) = 0$ entonces $\mu(A) = 0$, pues como $\kappa \in C$ entonces $\mu$ es $\kappa -$ *aditiva*. Así que $\mu$ es $\kappa_0 -$ *aditiva* y, por lo tanto, $\kappa_0 \in C$.

Afirmamos que $\kappa_o$ es un cardinal inaccesible. Para demostrarlo se seguirá un camino largo.

Lo primero es que, como $\kappa_0$ es el máximo de $C$, existe $A \subset 2^{\aleph_0}$ tal que $A = \bigcup_{\eta<\kappa_0} A_\eta$, $\mu(A_\eta) = 0$ y $\mu(A) > 0$, pues de otro modo $\mu$ sería $\kappa_0^+ -$ *aditiva* (con $\kappa_0^+$ el cardinal sucesor de $\kappa_0$) contradiciendo la maximalidad de $\kappa_0$. Podemos además suponer, sin pérdida de generalidad, que los $A_\eta$ son ajenos. Sea $f : A \longrightarrow \kappa_0$ tal que $\forall a \in A$:

$$f(a) = \beta \iff a \in A_\beta.$$

Nótese que $f$ está bien definida, pues $A = \bigcup_{\eta<k_0} A_\eta$, con $A_\beta \cap A_\gamma = \emptyset$ si $\beta \neq \gamma$, de modo que $\forall a \in A$ $\exists! \beta < \kappa_0$ tal que $a \in A_\beta$. Lo siguiente será dar una medida para los subconjuntos de $\kappa_0$ a partir de $\mu$. Dicha medida nos permitirá, entre otras cosas, demostrar la regularidad de $\kappa_0$. Sea $\sigma : \wp(\kappa_0) \longrightarrow [0,1]$ tal que $\forall P \subset \kappa_0$:

$$\sigma(P) = \frac{\mu(f^{-1}[P])}{\mu(A)}.$$

Nótese que $\sigma$ está bien definida pues $\mu(A) > 0$ y $\forall P \subset \kappa_0$, $f^{-1}[P]$[15] $\subsetneq A \subsetneq 2^{\aleph_0}$ y $\mu$ está definida en $2^{\aleph_0}$ . Por otro lado, $\sigma$ es medida gracias a que $\mu$ lo es, pues la propiedad de ser medida se preserva bajo multiplicación de escalares[16]. No incluimos la demostración de la $\sigma$ aditividad porque es análoga a la demostración, que a continuación veremos, de que $\sigma$ hereda la $\kappa_0 -$ *aditividad* de $\mu$[17]. Sea $\lambda < \kappa_0$ y $B = \bigcup_{\eta<k_0} B_\eta \subset \kappa_0$ tal que $\forall \eta < k_0$, $\sigma(B_\eta) = 0$. Por definición de $\sigma$, lo anterior quiere decir que $\forall \eta < k_0$, $\mu(f^{-1}[B_\eta]) = 0$ y como $\mu$ es $\kappa_0 -$ *aditiva* entonces $\mu(\bigcup_{\eta<k_0} f^{-1}[B_\eta]) = 0$. Además $\bigcup_{\eta<k_0} f^{-1}[B_\eta] = f^{-1}\left[\bigcup_{\eta<k_0} B_\eta\right]$. Así que $\sigma(B) = 0$ y $\sigma$ también es $\kappa_0 -$ *aditiva*. Por otro lado, obsérvese que:

(1) $\forall \xi < \kappa_0$, como $\xi < \kappa_0 \iff \xi \in \kappa_0$ y todo cardinal es transitivo por ser un ordinal, $\xi < \kappa_0 \Longrightarrow \xi \subset \kappa_0$.

---

[14] Más adelante veremos que en realidad no puede serlo.

[15] Puede ser que $f^{-1}[P] = \emptyset$, pero $\emptyset \in \wp(2^{\aleph_0})$ que es la $\sigma$ álgebra sobre la que está definida $\mu$. Cabe señalar que en un sentido más intuitivo $\sigma$ "mide" a $P \cap \operatorname{Im} f$.

[16] En este caso se multiplica por el escalar $\frac{1}{\mu(A)}$ para que la medida del total sea 1 (i.e $\sigma(\kappa_0) = 1$).

[17] Para tener la $\sigma -$ *aditividad*, lo único que hay que verificar es que si $B_\alpha \cap B_\gamma = \emptyset$, con $\alpha \neq \beta$, entonces $f^{-1}(B_\alpha) \cap f^{-1}(B_\gamma) = \emptyset$. Pero eso se sigue inmediatamente del hecho de que $f$ es función.

(2) $\forall \xi < \kappa_0$, $\sigma(\xi) = 0$. Nótese que $\forall \phi \in \xi$, $f^{-1}[\{\phi\}] \subset A_\phi$ con $\mu(A_\phi) = 0$. De modo que $\forall \phi \in \xi$, $\sigma(\{\phi\}) = 0$. Y como $\xi = \bigcup\limits_{\phi \in \xi} \{\phi\}$, usamos la $\kappa_0$ -aditivad de $\sigma$ para concluir que $\sigma(\xi) = 0$[18].

(3) $\kappa_0$ es cardinal regular. Supongamos que no lo es, entonces existe $\lambda < \kappa_0$ tal que $\kappa_0 = \bigcup\limits_{\eta<\lambda} \kappa_\eta$, con $\kappa_\eta < \kappa_0$. Por el inciso (1) eso quiere decir que existe $\lambda < \kappa_0$ tal que $\kappa_0 = \bigcup\limits_{\eta<\lambda} \kappa_\eta$, con $\sigma(\kappa_\eta) = 0$. Así que usando la $\kappa_0$ aditividad de $\sigma$ tendríamos que $\sigma(\kappa_0) = 0$! (pues $\sigma(\kappa_0) = 1$). Por lo tanto, $\kappa_0$ es regular[19].

(4) $\kappa_0$ es cardinal límite. Supongamos que es cardinal sucesor, entonces existe $\lambda \in CAR$ tal que $\kappa_0 = \lambda^+$. Es decir, $\forall\ \xi \in OR(\xi < \kappa_0 \Longrightarrow |\xi| \leq \lambda)$. Así que dada $\xi < \kappa_0$, sea $f_\xi$ una función inyectiva entre $\xi$ y $\lambda$. A partir de esas funciones definimos, para cada $\alpha < \kappa_0$ y para $\eta < \lambda$, los siguientes conjuntos:

$$B^\eta_\alpha = \{\xi \in \kappa_0 : \alpha < \xi \wedge f_\xi(\alpha) = \eta\}$$

Obsérvese que si tomamos una $\eta$ fija entonces $\forall \alpha, \beta < \kappa_0$ $(\alpha \neq \beta \Longrightarrow B^\eta_\alpha \cap B^\eta_\beta = \emptyset)$, pues como $f_\xi$ es inyectiva, $\xi \in B^\eta_\alpha \cap B^\eta_\beta \Longrightarrow f_\xi(\alpha) = \eta = f_\xi(\beta) \Longrightarrow \alpha = \beta$. De modo que los conjuntos $B^\eta_\alpha$ pueden pensarse, una vez más, como las entradas de una matriz infinita de $\lambda\times$ $\kappa_0$ cuyos renglones contienen conjuntos que son ajenos dos a dos:

$$\begin{array}{ccccc}
B^1_1 & B^1_2 & \ldots & B^1_\alpha & \ldots \\
B^2_1 & B^2_2 & \ldots & B^2_\alpha & \ldots \\
\ldots & \ldots & \ldots. & \ldots & \ldots. \\
B^\eta_1 & B^\eta_2 & \ldots & B^\eta_\alpha & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots
\end{array}$$

Lo siguiente será ver que la unión de cada columna es igual a $\kappa_0$ excepto por una cantidad a lo más $\lambda$ de elementos. Es decir, $C_\alpha$ denota la unión de la columna determinada por $\alpha < \kappa_0$ (i.e. $C_\alpha = \bigcup\limits_{\eta<\lambda} B^\eta_\alpha$), entonces $|\kappa_0 - C_\alpha| \leq \lambda$[20]. Para ver esto, tomemos $\alpha < \kappa_0$ fija. Entonces $\forall \xi < \kappa_0$ $[\alpha < \xi \Longrightarrow \alpha \in \xi \Longrightarrow \alpha \in dom f_\xi \Longrightarrow \exists \eta < \lambda(f_\xi(\alpha) = \eta \Longrightarrow \xi \in B^\eta_\alpha)]$. Por lo tanto, $\forall \xi < \kappa_0[\alpha < \xi \Longrightarrow \xi \in C_\alpha]$. De donde podemos concluir que $\kappa_0 - C_\alpha \subseteq \{\xi : \xi \leq \alpha\} = \alpha + 1$[21]. Como $\kappa_0$ es cardinal, en

[18] Nótese que si $A$ es acotado, entonces existe $\alpha \in \kappa_o$ tal que $A \subset \dot{\alpha}$. Así que $\sigma(A) \leq \sigma(\alpha) = 0$. Esto quiere decir que todo subconjunto de $\kappa_0$ acotado tiene $\sigma$ medida cero.

[19] Nótese que si $\kappa_0$ es regular, entonces cualquier subconjunto no acotado tiene cardinalidad $\kappa_0$. Así que cualquier conjunto de $\sigma-$medida positiva, como no puede ser acotado, debe tener cardinalidad $\kappa_0$. Esto coincide con el caso de la medida de Lebesgue, en la que suponiendo axioma de elección, se puede demostrar que si $\mu_L(A) > 0$, entonces $|A| = c$. Ya que si $\mu_L(A) > 0$ entonces por el Lema 36 existe $F$ cerrado tal que $F \subset A$. De modo que, por el corolario34, $c = |F| \leq |A|$.

[20] Lo cual querría decir que $|C_\alpha| = \lambda^+ = \kappa_0$.

[21] $\alpha + 1$ es el ordinal sucesor de $\alpha$, así que tiene la misma cardinalidad que $\alpha$.

particular es ordinal límite ($\kappa_0 = \lambda^+$). De modo que si $\alpha+1 < \kappa_0$, entonces $|\alpha + 1| \leq \lambda$. Por lo tanto, $\forall\alpha < \kappa_0$, $|\kappa_0 - C_\alpha| \leq \lambda$. Podemos decir más: como $\kappa_0 - C_\alpha \subseteq \alpha+1 < \kappa_0$, por el inciso (1) tenemos que $\sigma(\alpha+1) = 0$ y, por lo tanto, $\sigma(\kappa_0 - C_\alpha) = 0$ . Por otro lado, toda columna tiene al menos una entrada de medida positiva. De lo contrario, como $C_\alpha = \bigcup_{\eta<\lambda} B^\eta_\alpha$, $\lambda < \kappa_0$ y $\sigma$ es $\kappa_o$-aditiva, entonces $\sigma(C_\alpha) = 0$. Pero si $\sigma(C_\alpha) = 0$, entonces $\sigma(\kappa_0 - C_\alpha) = \sigma(\kappa_0) = 1$! (pues $\sigma(\kappa_0 - C_\alpha) = 0$) . Así que $\forall\alpha < \kappa_0\ \exists\eta < \lambda(\sigma(B^\eta_\alpha) > 0)$. Por otro lado, sea $B_{MP} = \{B^\eta_\alpha : \sigma(B^\eta_\alpha) > 0\}$[22]. Como se tienen $\kappa_0$ columnas y acabamos de ver que cada una tiene al menos una entrada con medida positiva, entonces $|B_{MP}| = \kappa_0$. Sin embargo, puesto que hay $\lambda$ renglones entonces al menos un renglón tiene $\kappa_0$ entradas de medida positiva. Para toda $\eta < \lambda$ sea $R^\eta_{MP} = \{(B^\eta_\alpha) : \alpha < \kappa_o$ y $\sigma(B^\eta_\alpha) > 0\}$, de modo que $B_{MP} = \bigcup_{\eta<\lambda} R^\eta_{MP}$ y $|B_{MP}| = \lambda \bullet \sup_{\eta<\lambda} |R^\eta_{MP}|$. Como ya habíamos visto que $|B_{MP}| = \kappa_0 = \lambda^+$ entonces $\sup_{\eta<\lambda} |R^\eta_{MP}| = \lambda^+ = \kappa_0$. Así que existe $\eta_0 < \lambda$ tal que $|R^{\eta_0}_{MP}| = \lambda^+ = \kappa_0$. Pero los elementos de $R^{\eta_0}_{MP}$ son ajenos dos a dos, ya que pertenecen al renglón determinado por $\eta_0$ y como $\aleph_1 \leq \kappa_0$ todo esto querría decir que tenemos una familia $(R^{\eta_0}_{MP})$ de conjuntos de medida positiva y ajenos dos a dos que es no numerable!, pues como $\sigma$ es medida finita se cumple la versión generalizada del Lema 40 (ver Observación 10). Como esta última contradicción surge de suponer que todas las columnas tienen al menos una entrada de medida positiva y si llamamos a esta afirmación A, entonces tendríamos que concluir que $\neg A$. Es decir, que hay al menos una columna sólo con entradas de medida cero. Sin embargo, ya habíamos visto que eso tampoco podía ser. Así que de la hipótesis $\kappa_o$ es sucesor concluimos que $A$ y $\neg A$! Por lo tanto, $\kappa_0$ es cardinal límite.

Finalmente, con estas tres observaciones, junto con el hecho de que $\aleph_1 \leq \kappa_0 \leq 2^{\aleph_0}$, tenemos que $\kappa_0$ es un cardinal inaccesible menor igual que $2^{\aleph_0}$. ∎

Si se consideraran únicamente estos dos teoremas de Ulam, la hipótesis de la medida sería un elemento más para creer en la falsedad de la hipótesis del continuo y la hipótesis correspondiente al teorema generalizado de Ulam. Pero si fuéramos consistentes con esta postura, por la demostración de Vitali, tendríamos que desechar también el axioma de elección. Lebesgue y Borel, entre otros, creían que eso era lo que procedía (ver Kanamori [16] p. 22 y Wagon [24] p. 217). El problema es que desechar toda forma de elección infinita implicaría desechar una parte importante del análisis matemático y, por lo tanto, la razón de ser del problema de la medida[23]. Además se tendría que tener la certeza de que todos los conjuntos no medibles son conjuntos que se originan a partir del axioma de elección.

[22] Nótese que éste es el conjunto de las entradas de medida postiva de toda la matriz.

[23] Banach y Tarski se habían dado cuenta de esta problemática disyuntiva cuando descubrieron las descomposiciones paradójicas (ver Wagon [24] p. 217).

Los resultados de Solovay, que se presentarán en el siguiente capítulo, dan una solución feliz a este problema, pues no sólo confirman que el axioma de elección es condición necesaria para que existan conjuntos no Lebesgue medibles, sino que matizan que es el caso sólo con versiones relativamente fuertes de elección.

## 4.2. Resumen

Los resultados que se presentaron a lo largo de la primera parte se resumen en las siguientes tablas. Los rubros **Medida**, **Grupos** y **Conjuntos** se refieren a las teorías correspondientes. El rubro **Espacios** contiene las propiedades de los espacios euclideanos (incluidas aquellas relacionadas con sus respectivos grupos de isometrías) utilizadas en la demostración en cuestión.

| NEGACIÓN DE LA HIPÓTESIS DE LA MEDIDA (Existencia de conjuntos no medibles) | | | | | |
|---|---|---|---|---|---|
| | Hipótesis utilizadas | | | | |
| Pruebas | Medida | Espacios | Grupos | AE | Conjuntos |
| Vitali | $\sigma$-aditividad Invarianza | Traslaciones | | $BO_{\mathbb{R}}$ | |
| Banach Tarski | Finito-aditividad Invarianza | Rotaciones | Descom-posiciones Paradójicas | $BO_{\mathbb{R}}$ | |
| Bernstein | $\sigma$-aditividad Regularidad | Medida de Lebesgue | | $BO_{\mathbb{R}}$ | Cantor Bendixson |
| Ulam | $\sigma$-aditividad | | | $BO_{\mathbb{R}}$ | Hipótesis del Continuo |
| Ulam | $\sigma$-aditividad | | | $BO_{\mathbb{R}}$ | Cardinal Inacce-sible |

Es interesante notar la evolución de las condiciones: de aquellas fuertemente ligadas a los espacios euclideanos a las de buen orden y cardinalidad. Esto indica el carácter general del problema de la medida y su vínculo con la teoría de conjuntos.

En la siguiente tabla se incluyen resultados relacionados con la medida de Lebesgue que permiten entender las preocupaciones que genera la existencia de conjuntos no medibles y la importancia de caracterizarlos.

| MEDIDA DE LEBESGUE | |
|---|---|
| Prueba | Elementos de la prueba |
| $2^c$ conjuntos medibles | Completud de la medida<br>Conjunto de Cantor |
| $2^c$ conjuntos no medibles | Conjunto de Bernstein |

A continuación se presenta una tabla que contiene los resultados en los que la dimensión juega un papel importante. En ambos casos las medidas consideradas no sólo son finito-aditivas, sino también invariantes bajo transformaciones rígidas.

| HIPÓTESIS DE LA MEDIDA FINITO-ADITIVA | | | |
|---|---|---|---|
| Dimensión | Medida | Conjuntos no medibles | Elementos de la prueba |
| $\mathbb{R}^1$ | Banach | Ninguno | Hahn-Banach (Lema de Zorn) |
| $\mathbb{R}^3$ | Cualquiera | Análogo Vitali | Banach-Tarski (Descomposiciones paradójicas y $BO_{\mathbb{R}}$) |

La última tabla contiene las dos respuestas positivas respecto a la hipótesis de la medida que se presentan en esta tesis. El modelo de Solovay se expone con detalle hasta la segunda parte, pero se incluye en este resumen justamente porque indica qué tienen en común los conjuntos no medibles.

| HIPOTESIS DE LA MEDIDA | | |
|---|---|---|
| Prueba | Elementos de la prueba | Hipótesis Conjuntistas |
| Medida de Banach (Finito Aditiva) | $\mathbb{R}^1$ Hahn-Banach | Lema de Zorn |
| Modelo de Solovay | Medida de Lebesgue Método de Forcing | Cardinales Inaccesibles Fuertes |

# Parte II

# Un modelo en el que todos los conjuntos de reales son Lebesgue medibles

# Modelos de una teoría

Cuando nos enfrentamos a la demostración. vía el conjunto de Vitali, de la negación de la hipótesis de la medida se plantearon dos opciones para rescatarla: (1) debilitar las condiciones de la medida, (2) rechazar o debilitar el axioma de elección. La demostración con los conjuntos de Bernstein nos hizo ver que desechar la invarianza bajo traslación no era solución. Por otro lado, el teorema de Banach-Tarski muestra que abandonar la $\sigma-aditividad$ tampoco solucionaría el problema. Sin embargo, ambas pruebas, junto con la de Vitali, tienen una hipótesis en común: la existencia de un buen orden para los reales. Así que nos preguntamos si ésta es condición necesaria para la negación de la hipótesis de la medida.

Para confirmar que el axioma de elección es responsable de la existencia de conjuntos no medibles bastaría ver que $ZF \nvdash \neg HM$. Más adelante se discutirá con detenimiento qué es un modelo de la teoría de conjuntos, por lo pronto sólo cabe señalar que si existe un modelo de $ZF + HM$, entonces la negación de la hipótesis de la medida no es teorema de $ZF$. Sin embargo, al menos para el caso de la medida de Lebesgue, la $\sigma - aditividad$ de la medida presupone la posibilidad de hacer elecciones numerables. De modo que parece más sensato buscar un modelo de $ZF + AE_{débil}$, con $(AE_{débil})$ una versión del axioma de elección al menos tan fuerte como elección numerable, pero obviamente más débil que $BO_{\mathbb{R}}$. Un modelo con esas características fue justamente lo que Solovay construyó.

Un modelo de una teoría axiomática es una interpretación en la que todos los axiomas son verdaderos y, por lo tanto, también todos los teoremas que de ellos se derivan. Para ver que un enunciado no es teorema de una teoría basta encontrar una interpretación en la que se cumplan todos los axiomas, pero el enunciado sea falso. Por ejemplo, sea **TO (Teoría del Orden)** el conjunto de todos los teoremas que se desprenden de los siguientes axiomas:

**Axioma de Asimetría:** No existen $a, b$ tales que $a < b$ y $b < a$.

**Axioma de Transitividad:** Para todos $a, b$ y $c$, si $a < b$ y $b < c$ entonces $a < c$.

En este caso todo conjunto ordenado es modelo de **TO**. Por ejemplo el conjunto $\wp(\{\emptyset\}, \{\emptyset, \{\emptyset\}\})$ ordenado con la contención propia.

Consideremos el siguiente enunciado:

**Hipótesis de Linealidad (HL):** Para todos $a$ y $b$ ó bien $a < b$ ó $b < a$ ó $a = b$.

Como la contención propia no es un orden lineal sobre $\wp(\{\emptyset, \{\emptyset\}\})$, ya que $\{\{\emptyset\}\} \not\subseteq \{\emptyset\}$ y $\{\emptyset\} \not\subseteq \{\{\emptyset\}\}$, tenemos un modelo de **TO** en el que la hipótesis de linealidad es falsa. Así que **HL** no es teorema de **TO**. Desde luego que esto no quiere decir que la negación de la hipótesis de linealidad sea teorema de **TO**. De hecho, si tomamos el conjunto $\{\emptyset, \{\emptyset\}\}$ ordenado mediante la relación de pertenencia, tenemos un modelo de **TO** en el que se cumple la hipótesis de linealidad, así que su negación tampoco puede ser teorema de la teoría del orden.

Un enunciado, como la hipótesis de linealidad, que no se puede demostrar ni refutar a partir de una teoría se conoce como un enunciado independiente de la teoría. Esto quiere decir, entre otras cosas, que tanto él como su negación se pueden agregar como axiomas. La hipótesis del continuo es el ejemplo clásico de un enunciado independiente de la teoría de conjuntos, aunque no es el único. El trabajo de Solovay consistió justamente en demostrar que la hipótesis de la medida es un enunciado independiente de ZF (pues demostró que hay un modelo de la teoría de conjuntos sin el axioma de elección, en el que todos los conjuntos de reales son Lebesgue medibles).

## Modelos de la teoría de conjuntos.

En el ejemplo que dimos de la teoría del orden hablamos intuitivamente de los modelos como conjuntos ordenados cuyos elementos darían significado a las variables $a, b$ y $c$ y cuyo orden interpretaría el símbolo $<$. El caso de los modelos de la teoría de conjuntos es análogo, ya que también indican cuál es el universo (que puede ser un conjunto o una clase) sobre el que se toman los conjuntos y la relación entre ellos que interpreta el símbolo $\in$ y se comporta como lo establecen los axiomas. En ambos casos es posible formalizar la noción de interpretación.

El lenguaje de la teoría de conjuntos es un lenguaje de primer orden con igualdad que tiene cómo únicos parámetros el símbolo de cuantificación $\forall$ y el símbolo de predicado binario $\in$, pues es posible expresar todos los axiomas utilizando únicamente esos dos símbolos y los símbolos lógicos $(\longrightarrow, \neg, v_1, v_2 \ldots, =)$[24].

**Definición 18** *Una interpretación $\mathcal{I}$ del lenguaje de la teoría de conjuntos es una función que al símbolo de cuantificación $\forall$ le asigna un conjunto (o clase) no vacío $M$, conocido como el universo de la interpretación, y al símbolo $\in$ le asigna un relación binaria $R$ definida en los elementos de $M$ $(R \subset M \times M)$.*

Nótese que las interpretaciones del lenguaje de la teoría de conjuntos quedan completamente determinadas por el par $\langle M, R \rangle$, donde $M$ es el universo de la interpretación y $R$ la relación que interpreta al símbolo $\in$.

[24] Nótese que en el caso del lenguaje de la teoría del orden también se tiene sólo dos parámetros $(\forall, <)$.

Por otro lado, el teorema de completud de Gödel nos garantiza que si una teoría es consistente entonces tiene un modelo. Para el caso particular de ZFE parece sensato creer que se trata de una teoría consistente. Sin embargo, en sentido estricto no podemos saber si es así, pues el teorema de incompletud de Gödel señala que ZFE no puede probar su propia consistencia (ver Enderton [8] p. 275). Pero incluso si otra teoría pudiera demostrar la consistencia de ZFE nos veríamos después enfrentados al problema de demostrar, a su vez, la consistencia de ésta. Esto explica por qué las pruebas de independencia de un enunciado respecto a la teoría de conjuntos son pruebas de consistencia relativa en las que la consistencia de ZFE es un supuesto de vital importancia.

# Capítulo 5

# Teorema de Solovay

El teorema de Solovay es un ejemplo de los resultados que se obtienen mediante pruebas de consistencia relativa, aunque en este caso se presupone la consistencia de algo más que ZFE. La discusión sobre las hipótesis que Solovay agrega se dejará para más adelante. Por lo pronto, se enunciará el teorema y se discutirá su significado para el análisis matemático y el problema general del tamaño de un conjunto.

**Notación 4** *El enunciado "Existe un cardinal fuertemente inaccesible" se denotará con $I$.*

**Teorema 46** *Supongamos que existe un modelo estándar y transitivo de $ZFE+I$. Entonces existe un modelo estándar transitivo de $ZF$ en el que los siguientes enunciados son verdaderos:*

*(1) El principio de elección dependiente.*

*(2) Todo conjunto de reales es Lebesgue medible.*

*(3) Todo conjunto de reales tiene la propiedad de Baire.*

*(4) Todo conjunto no numerable de reales contiene un conjunto perfecto.*

## 5.1. Elección Dependiente y Análisis Matemático.

El principio de elección dependiente es pieza clave para entender la importancia del resultado de Solovay, ya que garantiza la posibilidad de tener elecciones numerables.

**Definición 19 (Principio de Elección Dependiente)** *Si $R$ es una relación en un conjunto no vacío $X$ tal que $\forall x \in X\ \exists y \in X\ (xR\ y)$, entonces existe una sucesión $\{x_n\}_{n=0}^{\infty}$ de elementos de $X$ tal que $\forall n \in \mathbb{N}$, $x_n R\ x_{n+1}$.*

**Definición 20 (Principio de Elección Numerable)** *Si $\{A_n\}_{n=0}^{\infty}$ es una familia numerable de conjuntos no vacíos entonces existe una sucesión $\{a_n\}_{n=0}^{\infty}$ tal que $\forall n \in \mathbb{N}.\ a_n \in A_n$.*

**Notación 5** *Denotaremos con ED al Principio de Elección Dependiente y con EN al Principio de Elección Numerable.*

**Teorema 47** ($ED \Longrightarrow EN$) *El principio de elección dependiente implica el principio de elección numerable.*

**Demostración.** Sea $A = \{A_n\}_{n=0}^{\infty}$ una familia de conjuntos no vacíos y sea $S_A$ el conjunto de todas las sucesiones finitas $s = \{a_i\}_{i=0}^{n}$ tales que $\forall i,\ 0 \leq i \leq n$, $a_i \in A_i$. Sea $R$ una relación sobre $S_A$ tal que $\forall s, t \in S_A$:

$$sR\,t \text{ ssi } s = \{a_i\}_{i=0}^{n} \text{ y } t = s \cup \{a_{n+1}\}, \text{con } a_{n+1} \in A_{n+1}.$$

Por el principio de elección dependiente tenemos una sucesión numerable $\{s_n\}_{n=0}^{\infty}$ de sucesiones finitas tal que $s_n R s_{n+1}$. Como ($sR\,t \Rightarrow s \subset t$), entonces $\{s_n\}_{n=0}^{\infty}$ es tal que $s_n \subset s_{n+1}$. Así que $\bigcup_{n \in \mathbb{N}} s_n = \{a_i\}_{i=0}^{\infty}$, con $a_i \in A_i$. ■

Muchos teoremas del análisis matemático se demuestran mediante el principio de elección numerable o el principio de elección dependiente. Desde resultados muy básicos como *unión numerable de conjuntos numerables es numerable* hasta el teorema de Hahn-Banach restringido a espacios de Banach separables[1]. Incluso, la $\sigma -$ *aditividad* de la medida de Lebesgue depende del principio de elección numerable, pues ZF es consistente con que los reales sean unión numerable de conjuntos numerables y, por lo tanto, también es consistente con que la medida de Lebesgue no sea $\sigma - aditiva$[2]. Esto explica porque el principio de elección dependiente es el primer enunciado en aparecer dentro del teorema de Solovay, pues sin él la medida de Lebesgue no ocuparía un papel central en el problema de la medida.

**Teorema 48** *Si el principio de elección numerable es válido, entonces la unión de una familia numerable de conjuntos numerables es numerable.*

**Demostración.** Supongamos que se cumple el principio de elección numerable. Sea $A = \{A_i\}_{i \in \mathbb{N}}$ tal que $\forall i \in \mathbb{N}$, $|A_i| = \aleph_0$. Sin pérdida de generalidad podemos además suponer que $A_i \cap A_j = \emptyset$ si $i \neq j$. Para demostrar que $\left|\bigcup_{i \in \mathbb{N}} A_i\right| = \aleph_0$ basta dar una función biyectiva F de $\mathbb{N} \times \mathbb{N}$ en $\bigcup_{i \in \mathbb{N}} A_i$, ya que $|\mathbb{N} \times \mathbb{N}| = \aleph_0$. Como $\forall i \in \mathbb{N}$, $|A_i| = \aleph_0$ entonces $\forall i \in \mathbb{N}$, $G_i = \{g : g \text{ es biyección entre } \mathbb{N} \text{ y } A_i\} \neq \emptyset$.

[1] Es decir, restringido a los espacios vectoriales normados completos (de Banach) y separables (ver Solovay [23] p.3).

[2] Hay un modelo de ZF en el que los reales son unión numerable de subconjuntos numerables de reales (ver Wagon [24] p. 208). Así que en ese modelo la medida de Lebesgue no es $\sigma-$aditiva, pues de otro modo $\mu_L(\mathbb{R}) = 0$!

Así que $\{G_i\}_{i\in\mathbb{N}}$ es una familia numerable de conjuntos no vacíos. De modo que, por el principio de elección numerable, hay una sucesión de funciones $\{g_i\}_{i\in\mathbb{N}}$ tal que $\forall i \in \mathbb{N}$, $g_i \in G_i$. Sea $F : \mathbb{N} \times \mathbb{N} \longrightarrow \bigcup_{i\in\mathbb{N}} A_i$ tal que $\forall (n, m) \in \mathbb{N} \times \mathbb{N}$:

$$F(n, m) = g_n(m).$$

Nótese que $F$ es inyectiva ya que las $g_i$ son inyectivas y dada $m \in \mathbb{N}$, $g_i(m) \neq g_j(m)$ si $i \neq j$, pues en ese caso $A_i \cap A_j = \emptyset$. Además $F$ es suprayectiva, pues dada $a \in \bigcup_{i\in\mathbb{N}} A_i$, existe $j \in \mathbb{N}$ tal que $a \in A_j$. De modo que $a \in rang(g_j)$. Así que existe $m \in \mathbb{N}$ tal $g_j(m) = a$. ■

También demostraremos, usando el principio de elección numerable, parte del teorema de Heine-Borel:

**Teorema 49** *Sea $A \subset \mathbb{R}^n$. Si $A$ es compacto entonces $A$ es cerrado y acotado*[3].

**Demostración.** En realidad demostraremos que si un conjunto A no es cerrado o no es acotado, entonces no es compacto:

(1) Supongamos primero que $A$ no es acotado. Eso quiere decir que $A_n = \{a \in A : |a| > n\} \neq \emptyset$. Nótese que $\{A_n\}_{n\in\mathbb{N}}$ es una familia numerable de conjuntos no vacíos. Así que, por el principio de elección numerable, hay una sucesión $S = \{a_n\}_{n\in\mathbb{N}}$ tal que $|a_n| > n$. Por otro lado, como dada $a \in A$ existe $n \in \mathbb{N}$ tal que $|a| < n$, entonces dada $a \in A$ existe $V_a$ tal que $a \in V_a$ y $V_a \cap S$ es finito. Claramente $\{V_a\}_{a\in A}$ es una cubierta abierta de $A$ que no tiene subcubierta finita, ya que $S$ es una sucesión infinita de elementos de $A$ y $V_a \cap S$ es finito. Así que $A$ no es compacto.

(2) Supongamos ahora que $A$ no es cerrado. Es decir, supongamos que existe $x$ tal que $x \notin A$ y $x \in A'$. Recordemos que: $x \in A' \Longrightarrow \forall n \in \mathbb{N}$, $A_n = \{a \in A : |x - a| < \frac{1}{n}\} \neq \emptyset$. Así que $\{A_n\}_{n\in\mathbb{N}}$ es una familia de conjuntos no vacíos y entonces, por el principio de elección numerable, hay una sucesión $s = \{a_n\}_{n\in\mathbb{N}} \subset A$ tal que $\lim_{n\to\infty} a_n = x$. Sea $V_n = \{a \in \mathbb{R}^n : |x - a| > \frac{1}{n}\}$. Nótese que $\{V_n\}_{n\in\mathbb{N}}$ es una cubierta abierta de $A$ que no tiene una subcubierta finita, pues de otro modo, existiría $n_o$ tal que para toda $a \in A$, $|a - x| > \frac{1}{n_0}$. Es decir, si $m \geq n_0$, $A_m = \{a \in A : |x - a| < \frac{1}{m}\} = \emptyset$ y, por lo tanto, $x \notin A'$! Así que $A$ no es compacto. ■

## 5.2. Propiedad de Baire y Conjuntos Perfectos.

Sobre el inciso (2) del teorema de Solovay, que afirma que en el modelo todos los conjuntos de reales son Lebesgue medibles, no hay mucho más que decir que lo expuesto a lo largo de la tesis; salvo por el siguiente teorema que nos demuestra que un conjunto medible es un conjunto "bien portado"[4].

[3] El teorema de Heine-Borel en realidad es el si y sólo si, pero el regreso se puede consultar en Rudin [21] (pp. 33-35).

[4] En el sentido de que es igual a un conjunto de Borel, salvo por un conjunto de medida cero.

**Teorema 50** *Si un conjunto $A$ es Lebesgue medible entonces es igual a un conjunto $G_\delta$ menos un conjunto de medida cero.*

**Demostración.** Si $A$ es medible dada $n \in \mathbb{N}$, por definición, existen $F_n$ cerrado y $G_n$ abierto tales que $F_n \subset A \subset G_n$ y $\mu^*(G_n - F_n) < \frac{1}{n}$. Si tomamos $G = \bigcap_{n \in \mathbb{N}} G_n$, como $(G - F_n) \subset (G_n - F_n)$, entonces $\mu^*(G - F_n) < \frac{1}{n}$. Sea $N = G - A$. Nótese que $A = G - N$ con $G$ un conjunto $G_\delta$. Así que sólo falta ver que $\mu^*(N) = 0$, pues en ese caso $N$ es medible y tiene medida cero[5]. Además dada $n \in \mathbb{N}$, como $F_n \subset A$, entonces $G - A = N \subset (G - F_n)$. Así que por la monotonía de la medida exterior (Lema 16) se tiene que $\mu^*(N) < \frac{1}{n}$ y, por lo tanto, $\mu^*(N) = 0$. ■

Por otro lado, es interesante ver que los incisos (3) y (4) nos hablan de un buen comportamiento de los conjuntos de reales en términos topológicos y conjuntistas, como el inciso (2) lo hace respecto a la medida. Es decir, dentro del modelo de Solovay, es posible decir cuál es el tamaño de un conjunto en los tres sentidos expuestos al principio de la tesis.

Para comprender el significado del inciso (3) conviene recordar qué se entiende por la diferencia simétrica de dos conjuntos y qué significa que un conjunto tenga la propiedad de Baire.

**Observación 13** *Recuérdese que un conjunto $F$ es de primera categoría si y solo si es unión numerable de conjuntos densos en ninguna parte. Es decir, si $F = \bigcup_{n \in \mathbb{N}} D_n$, con $int\bar{D}_n = \emptyset$.*

**Definición 21** $A \triangle B = (A - B) \cup (B - A) = A \cup B - A \cap B$.

$A \triangle B$ se conoce como la diferencia simétrica de A y B.

**Definición 22** *Un conjunto $A$ tiene la propiedad de Baire si y sólo si existe $G$ un conjunto abierto tal que $A \triangle G = F$, con $F$ un conjunto de primera categoría.*

El teorema de Solovay nos indica que los conjuntos que no tienen la propiedad de Baire son, en todo caso, conjuntos que se obtienen a partir de versiones más fuertes del axioma de elección que el principio de elección dependiente. Sin embargo, aún suponiendo el axioma de elección, ¿Realmente hay conjuntos que no tienen la propiedad de Baire? Es interesante ver que ni el conjunto de Vitali ni los de Bernstein tienen la propiedad de Baire. A continuación daremos la demostración para el caso de los conjuntos de Vitali, pues la demostración para los conjuntos de Bernstein se puede consultar en Oxtoby [19] (p. 24).

[5] Ya se había señalado antes que la primera parte de la demostración del Teorema 30 se puede generalizar para demostrar que la medida de Lebesgue es completa.

Para demostrar que el conjunto de Vitali no tiene la propiedad de Baire se necesitan los siguientes resultados:

**Lema 51** *Todo subconjunto de un conjunto de primera categoría es de primera categoría.*

**Demostración.** Sea $A = \bigcup_{n\in\mathbb{N}} A_n$ tal que $int\ \bar{A}_n = \emptyset$ y sea $B \subset A$. Sea $B_n = B \cap A_n$. Está claro que $B = \bigcup_{n\in\mathbb{N}} B_n$, así que sólo habría que demostrar que $int\ \bar{B}_n = \emptyset$, pero eso es inmediato, pues $\bar{B}_n \subset \bar{A}_n \Longrightarrow int\bar{B}_n \subset\ int\ \bar{A}_n = \emptyset$. ■

**Teorema 52 (Baire)** *Todo intervalo abierto no vacío es de segunda categoría.*

La demostración de este teorema no se dará aquí, pero puede consultarse en Oxtoby [19] (p.2)

**Teorema 53** *El conjunto de Vitali no tiene la propiedad de Baire.*

**Demostración.** Sea $V$ el conjunto de Vitali. Supongamos que $V$ tiene la propiedad de Baire, entonces existen $U$ abierto y $F$ de primera categoría tales que $U \triangle V = F$. Como $F$ es de primera categoría entonces $F = \bigcup_{n\in\mathbb{N}} A_n$, con $int\bar{A}_n = \emptyset$.

Tenemos dos casos:

(1) $U \neq \emptyset$. Como $U$ es abierto no vacío, existe $I = (a, b) \subset U$. De modo que $I \cap V \neq \emptyset$, pues de lo contrario, $I \subset U \triangle V = F$ y, por el Lema 51, $I$ sería de primera categoría! (teorema de Baire). Sea $r \in (a, b) \cap V$. Nótese que $(a, r) \subset (a, b)$. Sea $T = V \cap (a, r)$ y sea $q \in \mathbb{Q}$ tal que $r + q < b$. Afirmamos que $T_{tras(q)} \subset (a, b) - V$. Sea $s \in T_{tras(q)}$, eso quiere decir que existe $t \in T$ tal que $s = t + q$. Como $q \in \mathbb{Q}$ , entonces $s \sim_v t$. Pero $T = V \cap (a, r)$, así que $t \in V$ y, por lo tanto, $s \notin V$. Además como $t \in T$ y $T \subset (a, r)$ entonces $a < t < r$ y como $q$ es tal que $r + q < b$, entonces $a < t + q < r + q < b$. Por lo tanto, $s = t + q \in (a, b)$. De modo que para toda $s \in T_{trasq}$, $s \notin V$ y $s \in (a, b)$. Es decir, $T_{trasq} \subset (a, b) - V \subset U - V \subset U \triangle V = F$. Así que, por el Lema 51,

$T_{tras(q)} = \bigcup_{n\in\mathbb{N}} B_n$ con $int\bar{B}_n = \emptyset$. Sea $C_n = (B_n)_{tras(-q)} = \{s - q : s \in B_n\}$. Afirmamos que $T \subseteq \bigcup_{n\in\mathbb{N}} C_n$ y que $int\bar{C}_n = \emptyset$. Sea $t \in T = V \cap (a, r)$, entonces $s = t + q \in T_{tras(q)}$. Así que existe $m \in \mathbb{N}$ tal que $s \in B_m$. Como $s - q = t$, entonces $t \in (B_m)_{tras-q} = C_m$. Así que $T \subseteq \bigcup_{n\in\mathbb{N}} C_n$. Falta ver que cada miembro de la unión es un conjunto denso en ninguna parte, pero esto es trivial ya que dicha propiedad se preserva bajo traslaciones y $C_n = (B_n)_{tras-q}$, con $int\bar{B}_n = \emptyset$. Tenemos entonces que T está contenido en un conjunto de primera categoría.

Por el Lema 51, T es de primera categoría. Es decir, $T = \bigcup_{n\in N} D_n$, con $int\bar{D}_n = \emptyset$. Por otro lado, $(a-r)-T = (a-r)-V \subset U \triangle V = \bigcup_{n\in N} A_n$, así que $(a-r)-T = \bigcup_{n\in N} F_n$, con $int\bar{F}_n = \emptyset$. Juntando estos resultados tenemos que $(a,r) = ((a,r)-T)\cup T = \bigcup_{n\in N} F_n \cup \bigcup_{n\in N} D_n$ con $int\bar{F}_n = \emptyset$ e $int\bar{D}_n = \emptyset$. Es decir, $(a,r)$ es de primera categoría! (teorema de Baire).

(2) $U = \emptyset$. En este caso $U \triangle V = V = F$, con $F$ de primera categoría. Es decir, $V = \bigcup_{n\in N} \bar{A}_n$, con int $\bar{A}_n = \emptyset$. Recordemos que para la demostración del Teorema 4, definimos para toda $q \in \mathbb{Q} \cap [0,1)$ el conjunto $V_q = \{s+q \mid s \in V \wedge s+q < 1\} \cup \{(t+q)-1 \mid t \in V \wedge t+q > 1\} = (V^q_{<1})_{trasq} \cup (V^q_{>1})_{tras(q-1)}$. Como $V^q_{<1} \subset V$ y $V^q_{>1} \subset V$ y estamos suponiendo que V es de primera categoría, entonces $V_q$ también es de primera categoría. Eso quiere decir que si $\{q_m\}_{m\in N}$ es una numeración de $\mathbb{Q} \cap [0,1)$, entonces $V_{q_m} = \bigcup_{n\in N} A_{nm}$, con $int\ \bar{A}_{nm} = \emptyset$. Como además $\{V_{q_m}\}_{m\in N}$ es partición numerable del intervalo [0,1) tenemos que $[0,1) = \bigcup_{m\in N}\bigcup_{n\in N} A_{nm}$, con $int\bar{A}_m = \emptyset$. Esto querría decir que el intervalo $[0,1)$ es de primera categoría![6]. En resumen ninguno de los dos casos se puede dar y, por lo tanto, el conjunto de Vitali no tiene la propiedad de Baire. ■

El hecho de que todos los conjuntos de reales tengan la propiedad de Baire habla de un buen comportamiento topológico, en el sentido de que todos los conjuntos de reales son como los conjuntos abiertos salvo por un conjunto topológicamente despreciable.

El inciso (4) nos indica también la existencia de un buen comportamiento del tamaño de los conjuntos de reales en términos de cardinalidad. De hecho, este inciso se puede ver como una respuesta al problema del continuo.

Lo primero, como en el caso de la propiedad de Baire, es recordar qué significa que un conjunto sea perfecto.

**Definición 23** *Un conjunto es perfecto si y sólo si es no vacío, cerrado y no tiene puntos aislados.*

Para demostrar, desde ZFE, que hay subconjuntos de reales no numerables que no contienen un conjunto perfecto utilizaremos los conjuntos de Bernstein. Sin embargo, antes veremos algunos resultados muy importantes respecto a la cardinalidad de los conjuntos perfectos

**Teorema 54** *Todo conjunto perfecto es no numerable.*

**Demostración.** Sea $P$ un conjunto perfecto. Supongamos que $P$ es numerable. Sea $\{x_n\}_{n=0}^{\infty}$ una numeración de los elementos de $P$ y sea $V_0$ una vecindad

[6] Nótese que este segundo caso es totalmente análogo a la demostración de la negación de la hipótesis de la medida a partir del conjunto Vitali.

tal que $x_0 \in V_0$. Como $P$ no tiene puntos aislados, entonces $x_0$ es punto de acumulación de $P$. De modo que existe $V_1$ tal que $x_0 \notin \bar{V}_1$, $V_1 \cap P \neq \emptyset$ y $\bar{V}_1 \subset V_0$. Procedemos entonces recursivamente. Supongamos que ya tenemos definido a $V_n$. Como $x_n$ es punto de acumulación, entonces existe $V_{n+1}$ tal que $x_n \notin \bar{V}_{n+1}$, $V_{n+1} \cap P \neq \emptyset$ y $\bar{V}_{n+1} \subset V_n$. Nótese que $\bar{V}_{n+1} \cap P$ es compacto, ya que es intersección de un compacto con un cerrado. Si $K_n = \bar{V}_{n+1} \cap P$, por construcción, tenemos que para toda $x_n \in P$. $x_n \notin K_n$. Y como $\{x_n\}_{n=0}^{\infty}$ es una numeración de los elementos de $P$ y $K_n \subset P$, entonces $\bigcap_{n \in \mathbb{N}} K_n = \emptyset$. Por otro lado, sabemos que $\{K_n\}_{n \in \mathbb{N}}$ es una familia anidada de conjuntos compactos, de modo que su intersección es no vacía[7]. Es decir, $\bigcap_{n \in \mathbb{N}} K_n \neq \emptyset$! Por lo tanto, todo conjunto perfecto es no numerable. ■

**Corolario 55** *Todo conjunto perfecto tiene la cardinalidad de los reales .*

**Demostración.** Este resultado se desprende de manera inmediata del teorema anterior y del Corolario 34 que justamente nos garantiza que todo conjunto cerrado no numerable tiene la cardinalidad de los reales. ■

Fue Cantor quien, en su intento por verificar la hipótesis del continuo, demostró que todo conjunto perfecto tiene la cardinalidad de los reales, con la esperanza de poder demostrar después que todos los subconjuntos de reales no numerables contenían un conjunto perfecto. Sin embargo, el axioma de elección sirvió, también en este caso, para dar un contraejemplo.

**Teorema 56** *Hay conjuntos de reales no numerables que no contienen un conjunto perfecto.*

**Demostración.** Sea $B$ un conjunto de Bernstein. Tenemos dos casos: (1) Supongamos que $B$ es numerable, entonces $B^c$ es no numerable. Afirmamos que $B^c$ no contiene un conjunto perfecto. Supongamos que existe $P$ un conjunto perfecto tal que $P \subset B^c$, entonces $P \cap B = \emptyset$. Sin embargo, por el teorema anterior sabemos que $P$ es un conjunto cerrado no numerable, así que por la definición de un conjunto de Bernstein $B \cap P \neq \emptyset$! Por lo tanto, $B^c$ no contiene un conjunto perfecto.(2) Supongamos que $B$ es no numerable, por el mismo argumento que usamos para $B^c$, se puede ver que en este caso $B$ es un conjunto no numerable que no contiene un conjunto perfecto. ■

El teorema de Solovay es interesante respecto al problema del continuo, pues el inciso (4) nos garantiza que en el modelo en cuestión los conjuntos de reales son finitos, numerables o equipotentes con el conjunto de todos los reales. Desde luego que en este caso no se puede pensar la cardinalidad de los reales como un cardinal ordinal, pues eso equivaldría a tener un buen orden para los reales que, como ya vimos, es condición suficiente para tener conjuntos que no sean

[7] Este es un conocido resultado de topología que se puede consultar en Rudin [20] (p. 33).

Lebesgue medibles. Pero entonces ¿Qué sentido tiene decir que la hipótesis del continuo para los reales es válida en el modelo de Solovay?

Hasta ahora hemos planteado el problema del continuo suponiendo que hay un buen orden para los reales, pues lo que nos hemos preguntado es si $\aleph_1$ es o no el cardinal de los reales. Sin embargo, existe una versión de la hipótesis del continuo que no necesita del axioma de elección, ya que se puede expresar en términos de funciones inyectivas.

**Notación 6** *Sean $A,B$ conjuntos. Escribiremos $A \lnapprox B$ si existen funciones inyectivas, pero no existen funciones suprayectivas, de $A$ en $B$.*

**Proposición 57** *Hipótesis del continuo sin axioma de elección (HC-AE):*

$$\neg \exists A \subset \mathbb{R} \text{ tal que } \mathbb{N} \lnapprox A \text{ y } A \lnapprox \mathbb{R}.$$

Nótese que si nosotros tuviéramos que $ZFE \vdash \neg HC$, el teorema de Solovay nos garantizaría que los conjuntos que niegan la hipótesis del continuo son, en todo caso, como los conjuntos no medibles; es decir, son conjuntos que se generan a partir de formas relativamente fuertes de elección. Sin embargo, Gödel demostró que ZFE es consistente con la Hipótesis del Continuo (es decir, que $ZFE \nvdash \neg HC$), con lo que garantizó que para negar dicha hipótesis no bastan versiones fuertes de elección, se necesitan nuevos axiomas. En ese sentido, el resultado de Solovay es en realidad una versión más débil de lo que probó Gödel, pues demuestra que $ZF + ED \nvdash \neg(HC - AE)$.

Por otro lado, el hecho de que la hipótesis del continuo se cumpla en el modelo de Solovay parece contradecir el primer teorema de Ulam[a]. Sin embargo, recordemos que el teorema de Ulam usa la versión de la hipótesis del continuo con axioma de elección. De hecho, ya desde ese momento se había hecho mucho énfasis en la importancia del orden que los reales heredan de $\aleph_1$ para el desarrollo de la demostración de Ulam.

Finalmente cabe señalar que aunque la prueba de Solovay es, como la de Gödel, una prueba de consistencia relativa, hay profundas diferencias entre ellas. La primera, que en parte explica porqué el resultado de Solovay es más débil respecto al problema del continuo, tiene que ver con la hipótesis de la que parte, ya que Solovay supone algo más fuerte que la simple consistencia de ZFE. La segunda tiene que ver con la forma en que se construye el modelo que se desea, ya que la prueba de Solovay se basa en el método de forcing, que introdujo Cohen precisamente para demostrar la consistencia de ZFE con la negación de la hipótesis del continuo (ver Kunen [18] pp. 211-216).

[a] *Si la hipótesis del continuo se cumple entonces hay conjuntos que no son Lebesgue medibles* (ver Teorema 42).

## 5.3. Cardinales fuertemente inaccesibles .

Para analizar las hipótesis del teorema de Solovay hace falta recordar qué es un cardinal inaccesible fuerte:

**Definición 24** *$\kappa$ es un cardinal fuertemente inaccesible si y sólo si $\kappa$ es un cardinal límite, regular, mayor que $\aleph_0$ y tal que $\forall\lambda < \kappa$, $2^\lambda < \kappa$*[9].

La existencia de cardinales fuertemente inaccesibles, que denotaremos con $I$, no es teorema de ZFE, ya que se puede ver que $ZFE + \neg I$ es consistente. La idea es tomar el universo constructible de Gödel denotado por $L$[10]. Si $L$ no tiene cardinales inaccesibles, entonces $L$ es modelo $ZFE + \neg I$ . De lo contrario, sea $\kappa_0 = \text{mín}\{\kappa \in L : \kappa \text{ es fuertemente inaccesible}\}$. Se puede ver que $L_{\kappa_0}$ es modelo de $ZFE + \neg I$ (Hrbacek y Jech [14] p.279). Así que $ZFE + \neg I$ es consistente. Sin embargo, no se puede probar la consistencia relativa de $ZFE + I$ (Capítulo 2, Amor [3]). En otras palabras, suponer que hay cardinales fuertemente inaccesibles puede contradecir a ZFE. Sin embargo, existen argumentos intuitivos a favor de dicha hipótesis.

Si observamos que $\aleph_0$ es un cardinal límite y regular tal que $\forall\lambda < \aleph_0$, $2^\lambda < \aleph_0$, es natural pensar que el enunciado "existen cardinales fuertemente inaccesibles" no es sino una generalización del Axioma de Infinito[11]. Sin embargo, en el mismo tenor, podríamos decir que un argumento inmediato a favor del axioma de elección es que se trata de una generalización del principio de elección numerable, lo que no quiere decir que esté exento de dificultades.

Al margen de si la hipótesis de Solovay es aceptable o no, lo que es notable es que recupera la conexión entre propiedades de cardinalidad y de medida que se comenzaban a perfilar con los resultados de Ulam. De hecho, en el siguiente capítulo, dedicado al método de forcing, se podrá ver que las hipótesis sobre la existencia de cardinales inaccesibles suelen traducirse en resultados sobre la cardinalidad del continuo (i.e. de $\mathbb{R}$) y propiedades respecto a la potencia de los reales. Esto no sólo vuelve a acercar los resultados de Solovay y Ulam, sino que demuestra que, pese a que la cardinalidad no es una forma adecuada de considerar el tamaño de un conjunto de reales para la teoría de la integración, sí es factor determinante en el problema de la medida[12].

---

[9] Esta última propiedad establece la diferencia entre los cardinales fuertemente inaccesibles y los cardinales inaccesibles sin más.

[10] $L$ es el modelo que Gödel creó para demostrar la consistencia de $ZFE + HC$.

[11] Hrbacek y Jech presentan un argumento más sofisticado que involucra una separación de los conjuntos análoga a la distinción de la lógica de primer orden y la de segundo orden (ver Hrbacek y Jech [14] p.279).

[12] Shelah probó que la hipótesis sobre la existencia de un cardinal fuertemente inaccesible es condición necesaria para tener un modelo en el que todos los conjuntos de reales son Lebesgue medibles (ver Kanamori [16] p.136).

# Capítulo 6

# Forcing

En 1963 Cohen demostró que la negación de la hipótesis del continuo es consistente con ZFE mediante una idea aparentemente sencilla y natural: partir de un modelo de la teoría de conjuntos y extenderlo agregando nuevos conjuntos. Sin embargo, puesto que el objetivo central era hacer más grande la potencia de los reales era claro que aquellos conjuntos que se introducían debían tener propiedades particulares.

La posibilidad de extender de manera controlada modelos de la teoría de conjuntos fue la clave para que el método de forcing, introducido por Cohen, se convirtiera en un fecundo método de pruebas de consistencia relativa aplicable a problemas no sólo de la teoría de conjuntos (como el caso del continuo), sino de otras áreas como teoría de la medida. La revisión de los resultados que dan forma al método de forcing junto con un esbozo de su aplicación con álgebras de Boole permiten entender, en términos generales, cómo construyó Solovay un modelo en el que todos los conjuntos de reales son Lebesgue medibles.

## 6.1. Modelos transitivos y estándar

En la introducción a la segunda parte se señaló que la consistencia de una teoría implica la existencia de un modelo para ella. Sin embargo, el teorema de Solovay, como todos los resultados que se prueban mediante forcing, presupone la existencia de un modelo con propiedades muy especiales: que sea transitivo y estándar[1]. La justificación de que las pruebas mediante el método de forcing siguen siendo pruebas de consistencia relativa se basa en algunos teoremas importantes de la lógica matemática, como el teorema de Reflexión, el Colapso de Mostowski y el teorema de Löwenheim-Skolem(ver Amor [2]). Nosotros nos concentraremos en explicar en qué consisten esas propiedades y por qué se piden.

**Definición 25** *$M$ es transitiva ssi* $\forall x(x \in M \Longrightarrow x \subset M)$.

[1] También se suele pedir que sea numerable, pero eso lo veremos más adelante.

**Definición 26** *Un modelo de ZFE es estándar si el símbolo $\in$ se interpreta como la pertenencia entre conjuntos.*

Recordemos que los modelos de ZFE son interpretaciones del lenguaje de la teoría de conjuntos que hacen verdaderos a los axiomas y que dichas interpretaciones se pueden caracterizar como pares ordenados conformados por un conjunto o clase que funciona como el "universo" y una relación que interpreta el símbolo $\in$. De modo que si $\langle M, R\rangle$ es modelo de ZFE, entonces interpreta a toda fórmula $\varphi$ del lenguaje de la teoría de conjuntos. Es posible definir formalmente (ver Kunen [18] p.112) la interpretación de una fórmula $\varphi$ en $\langle M, R\rangle$, que se escribe como $\varphi^{\langle M,R\rangle}$ y se conoce como la relativización de $\varphi$ a $\langle M, R\rangle$. Por ejemplo, $(x \in y)^{\langle M,R\rangle} = xRy$ y $(\exists x\ \alpha)^{\langle M,R\rangle} = \exists x \in M(\alpha^{\langle M,R\rangle})$.

La ventaja de trabajar con modelos estándar es que la interpretación de una fórmula del lenguaje de la teoría de conjuntos sólo requiere acotar los cuantificadores. Por ejemplo, $(\exists x\ (x \in y))^{\langle M,\in\rangle} = \exists x \in M(x \in y))$. Dado que en forcing se trabaja con modelos estándar, de ahora en adelante en lugar de referirnos a los modelos de ZFE mediante el par $\langle M, \in\rangle$, lo haremos simplemente mediante $M$.

### 6.1.1. Absolutez y Relativización

Una de las ventajas de los modelos transitivos es que nos garantizan que el significado de cierto tipo de fórmulas (llamadas fórmulas absolutas para modelos transitivos) no varía si se relativizan a un modelo transitivo o se piensan en el universo de todos los conjuntos. Por ejemplo, si los elementos de un conjunto no vacío que pertenece a una clase $M$ no están en M, entonces dicho conjunto es vacío según $M$. Esto no pasa si $M$es una clase transitiva, ya que por definición contiene a todos los elementos de sus elementos.

**Definición 27** *Sea $\varphi$ una fórmula del lenguaje de la teoría de conjuntos. $\varphi$ es una fórmula absoluta para clases transitivas si y sólo si para toda clase transitiva $M$*

$$\forall x_1,...,x_n \in M[\varphi^M(x_1,...,x_n) \Leftrightarrow \varphi(x_1,...,x_n)].$$

Existen muchas fórmulas que son absolutas para modelos transitivos de ZFE. Por ejemplo: $x = \emptyset, x = \omega$, R es relación y R es función (ver Kunen [18] pp.120-121). Sin embargo más interesante aún son las fórmulas que no son absolutas para modelos transitivos. Por ejemplo, $\alpha$ es cardinal ó $z = \wp(x)$. Más adelante se verá que la no absolutez de la noción de potencia de un conjunto juega un papel central en forcing.

## 6.2. Definición

Hemos señalado ya que la idea general de forcing es agregar a un modelo de ZFE nuevos conjuntos, de manera que se obtenga un modelo más grande[2] que dé cuenta de la consistencia de ZFE junto con el enunciado con el que se quiere probar la consistencia. Lo interesante es que los conjuntos que se agregan, aunque no están en el modelo base, pueden conocerse a partir de los elementos que están en éste, ¿Qué quiere decir esto?

El conjunto de reales puede servirnos como analogía para responder de manera intuitiva esta pregunta. Aunque más adelante, cuando se vea la definición de forcing, se dará un respuesta más formal. Sea $T$ la teoría determinada por los siguientes axiomas:

**Axioma de Asimetría:** No existen $a$ y $b$ tales que $a < b$ y $b < a$.

**Axioma de Transitividad:** Para todos $a, b$ y $c$, si $a < b$ y $b < c$ entonces $a < c$.

**Axioma de Linealidad:** Para todos $a$ y $b$ ó bien $a < b$ ó $b < a$ ó $a = b$.

**Axioma de Densidad:** Para todos $a$ y $b$ si $a < b$ entonces existe $c$ tal que $a < c < b$.

Nótese que $T$ es una extensión de la Teoría del Orden que presentamos en el Capítulo II y que $\langle \mathbb{Q}, < \rangle$ es un modelo de $T$ si el símbolo $<$ se interpreta como el orden canónico de los racionales, ya que ese es un orden lineal denso.

Consideremos el siguiente enunciado:

**Hipótesis de Completud (HCM):** Toda colección acotada superiormente tiene supremo.

Como no existe $q \in \mathbb{Q} \cap [0,1]$ tal que $q = \sup\{x \in \mathbb{Q} \cap [0,1] : x^2 < 2\}$, está claro que en $\langle \mathbb{Q}, < \rangle$ no se cumple la Hipótesis de Completud. Sin embargo, es posible agregar nuevos objetos a $\mathbb{Q}$ de manera que se tenga un modelo de $T + HCM$. La construcción de los reales mediante cortaduras de Dedekin no sólo permite ver que $\langle \mathbb{R}, < \rangle$ es una extensión de $\langle \mathbb{Q}, < \rangle$ en la que todos los axiomas de $T$ son válidos y la Hipótesis de Completud también, sino que los conjuntos que se agregan se pueden describir mediante los elementos del modelo inicial (es decir, mediante los racionales) y el orden entre ellos. El caso del modelo base y el modelo que se obtiene con forcing, conocido como extensión genérica, es análogo no sólo porque ahí también se tiene que el primero está contenido en el segundo, sino porque los elementos que están en el segundo se pueden describir mediante los elementos del primero[3].

Por otro lado, nótese que los ejemplos que demuestran que la Hipótesis de Completud no es válida en $\langle \mathbb{Q}, < \rangle$ son los que señalan qué tipo de objetos hay que agregar para extender el modelo. Cuando pensamos en modelos de ZFE no queda tan claro que haya modelos para los cuales existan conjuntos que no están en él, pero que se pueden describir a partir de sus elementos.

[2] Si consideramos la jerarquía acumulativa habría que decir que mediante forcing se obtiene en realidad un modelo más "gordo", ya que el modelo base y el modelo extendido tienen los mismos ordinales (ver Kunen [18] p. 191).

[3] Hrbacek y Jech [14] (pp. 275-277) exponen con más detalle esta analogía.

La clave está en el uso de órdenes parciales particulares que también permiten ver qué tipo de conjuntos no están en el modelo base y, por lo tanto, sugieren cómo hay que extenderlo.

**Definición 28** *Un orden parcial es un par* $\langle P, \leq \rangle$, *donde* $\leq$ *es una relación reflexiva, antisimétrica y transitiva sobre un conjunto* $P$.

**Observación 14** *En forcing se suele suponer además que existe* $1_P \in P$ *tal* $\forall p \in P$, $p \leq 1_P$. *Así que, en lo que sigue, cuando nos refiramos a* $P$ *un orden parcial se estará suponiendo que tiene un elemento máximo.*

**Definición 29** *Sea* $P$ *un orden parcial.* $G \subset P$ *es un filtro sobre* $P$ *si y sólo si:*

1. $G \neq \emptyset$.
2. $\forall p, q \in G, \exists r \in G(r \leq p \wedge r \leq q)$.
3. $\forall p \in G, \forall q \in P(p \leq q \Longrightarrow q \in G)$[4].

**Observación 15** *Si* $G$ *es un filtro sobre* $P$ *entonces* $1_P \in G$.

A continuación daremos una definición de densidad con respecto a órdenes parciales, que no se debe confundir con la noción topológica de densidad.

**Definición 30** *Un conjunto* $D$ *es denso en* $P$ *si y sólo si* $D \subseteq P$ *y* $\forall p \in P$ $\exists q \in D(q \leq p)$.

**Definición 31** *Sea* $M$ *un modelo estándar y transitivo de ZFE y* $P$ *un orden parcial tal que* $P \in M$. $G$ *es* $P$*-genérico sobre* $M$ *si y sólo si* $G$ *es filtro y* $\forall D$ $(D \in M \wedge D$ *es denso en* $P \Longrightarrow G \cap D \neq \emptyset)$.

**Definición 32** $\forall p, q \in P$, $p$ *y* $q$ *son compatibles ssi* $\exists r \in P(r \leq p \wedge r \leq q)$[5].

**Notación 7** $p \perp q$ *denotará el caso en que* $p$ *y* $q$ *son incompatibles* $(p \perp q \Leftrightarrow \neg \exists r \in P(r \leq p \wedge r \leq q))$.

**Definición 33** *Un orden parcial es frondoso si y sólo si* $\forall p \in P$ $\exists$ $s, t \in P$ $(s \leq p \wedge t \leq p \wedge s \perp t)$.

Si pensamos el orden parcial como un árbol invertido, donde $1_P$ es el punto en la que comienzan las primeras ramificaciones, se tiene una idea bastante más gráfica e intuitiva de lo que significa que un orden parcial sea frondoso. Se trata de algo así como un árbol con muchísimas ramas, ya que en cada una hay al menos una bifurcación.

Los órdenes parciales frondosos junto con los filtros $P$-genéricos son los que nos garantizan la existencia de conjuntos que, aunque relacionados con los elementos del modelo base, no pertenecen a éste.

[4] Esta propiedad permite ver por qué el nombre de filtro, pues nos dice que G "absorbe a todo lo que se encuentra por encima".

[5] Nótese que la condición (2) de filtro, nos dice que cualesquiera dos elementos de un filtro son compatibles.

**Lema 58** *Si $M$ es un modelo estándar y transitivo de ZFE, $P$ un orden parcial frondoso que pertenece a $M$ y $G$ un filtro $P$-genérico sobre $M$, entonces $G \notin M$.*

**Demostración.** Supongamos que $G \in M$. Sea $D = \{x \in P : x \notin G\}$. Afirmamos que $D$ es denso en $P$ y que $D \in M$. Como estamos suponiendo que $G \in M$ y $M$ es modelo de ZFE, por el axioma de separación, tenemos que $D \in M$. Por otro lado, como $P$ es frondoso, dada $p \in P$ $\exists s, t \in P(s \leq p \wedge t \leq p \wedge s \perp t)$, así que $s \notin G$ ó $t \notin G$, pues de otro modo $s$ y $t$ serían compatibles, por la condición (2) de filtro. Supongamos, sin pérdida de generalidad, que $s \notin G$, entonces $s \in D$. Y como $s \leq p$, $s$ es testigo de que existe $x \in D$ tal que $x \leq p$. Por lo tanto, $D$ es denso en $P$ y como $G$ es un filtro P-genérico entonces $G \cap D \neq \emptyset$! (pues $D = P - G$). ■

Nótese que $G \in \wp(P)$ pero $G \notin M$. Esto coincide con el hecho, fundamental para las pruebas de consistencia relativa, de que la potencia de un conjunto no es una noción absoluta. En particular, el hecho de que todo modelo estándar, transitivo y numerable con un orden parcial frondoso tenga un filtro $P$-genérico que no pertenezca al modelo es fundamental para el método de forcing.

La demostración de la existencia de filtros $P$-genéricos se basa en la tercera propiedad que se le pide a un modelo base, a saber: que sea numerable. Es interesante notar que este requisito involucra un resultado paradójico, pues si $M$ es un modelo de ZFE numerable entonces el conjunto de reales en $M$ es numerable. Como toda paradoja, este resultado (conocido como la paradoja de Skolem)[6], no encierra una contradicción. Para entenderlo basta notar que ser no numerable no es una noción absoluta, ya que un conjunto es no numerable si no hay una función suprayectiva de los naturales en él. Así que un conjunto numerable puede no ser numerable según un modelo simplemente porque la función que lo biyecta con los naturales no está en el modelo.

**Lema 59** *Si $M$ es un modelo numerable de ZFE, $P$ un orden parcial que pertenece a $M$ y $p \in P$, entonces hay un filtro $P$-genérico $G$ sobre $M$ tal que $p \in G$.*

**Demostración.** Sea $\{D_n\}_{n \in \mathbb{N}}$ una numeración de los densos en P que pertenecen a $M$ (dicha numeración existe pues por hipótesis $M$ es numerable). Sea $A_0 = \{p\}$ y $A_{n+1} = D_n$. Tomamos $X = \{\{p_n\}_{n=0}^{k} : p_n \in A_n, k \in \mathbb{N} \text{ y } \forall n < k, p_{n+1} \leq p_n\}$ y definimos la siguiente relación sobre X: $\forall s, t \in X$, $sRt$ ssi $s = \{p_n\}_{n=0}^{k}$ y $t = s \cup \{p_{k+1}\}$. Sea $s = \{p_n\}_{n=0}^{k} \in X$. Como $p_k \in A_k \subset P$ y $A_{k+1} = D_k$ es denso en P, entonces existe $q \in D_{k+1}$ tal que $q \leq p_k$. Así que basta tomar $t = s \cup \{q\}$ para ver que existe $t \in X$ tal que $sRt$. Esto quiere decir que $X$ y $R$ cumplen las hipótesis del principio de elección dependiente (ver Definición 19). Por lo tanto, existe una sucesión $\{s_k\}_{k=0}^{\infty}$ tal que $s_k R\, s_{k+1}$(i.e. $s_k = \{p_n\}_{n=0}^{k} \subset \{p_n\}_{n=0}^{k+1} = s_{k+1}$ y $p_{n+1} \leq p_n$). Sea $S = \bigcup_{k \in \mathbb{N}} s_k$. Nótese que $S$ es un sucesión infinita $\{p_n\}_{n=0}^{\infty}$ tal que $p_o = p$ y $\forall n \in \mathbb{N}$, $p_{n+1} \in D_n$ y $p_{n+1} \leq p_n$. Sea $G = \{q \in P : \exists n \in \mathbb{N} \text{ tal que } p_n \leq q\}$. Afirmamos que $G$ es filtro $P$-genérico.

[6] Véase Kunen [18] (p. 141)

La condición (1) de filtro se cumple porque $p = p_0 \in G$, la condición (2) porque $S$ tiene un orden lineal y la condición (3) por la forma en que se definió G. Finalmente G es $P$-genérico, ya que $\forall n \in \mathbb{N}$, $p_{n+1} \in G \cap D_n$. Así que G intersecta a todo denso en $P$ que pertenece a $M$. ■

Antes de continuar cabe señalar un par de cosas importantes. La primera es que la numeración de los densos en $P$ que pertenecen a $M$ existe, aunque eso no quiere decir que esté en $M$. De hecho, si $P$ es frondoso no hay numeración en $M$ de los densos en cuestión. Si hubiera una numeración de los densos en $M$, como $M$ es modelo del axioma de separación y $G = \{q \in P : \exists n \in \mathbb{N} \text{ tal que } p_n \leq q\}$ tendríamos, contrario a lo que demostramos en el Lema 58, que $G$ está en $M$. Esto nos conduce de manera natural a la segunda observación: ZFE se utiliza en dos niveles, (1) con $M$ un modelo de ZFE, (2) como metateoría desde donde se prueban cosas como el lema anterior, pues el principio de elección dependiente se aplicó "fuera" de $M$ (i.e. ZFE es el marco teórico desde donde se trabajan muchos de los resultados de forcing).

Lo siguiente es mostrar cómo se puede extender un modelo de manera que siga siendo modelo de ZFE. La idea es nombrar los elementos de la extensión genérica mediante elementos del modelo base, de forma que los nombres nos permitan responder preguntas sobre cómo son los elementos de la extensión genérica.

**Definición 34** *Sea $P$ un orden parcial. Decimos que $\tau$ es un $P$-nombre si y sólo si $\tau$ es una relación tal que $\forall(\sigma, p) \in \tau$, $\sigma$ es $P$-nombre y $p \in P$* [7].

**Notación 8** *Sean $V$ el universo de todos los conjuntos, $M$ un modelo de ZFE y $P$ un orden parcial en $M$. $V^P$ denotará a la clase de los $P$-nombres y $M^P$ a la clase de los $P$-nombres en $M$ (i.e. $M^P = V^P \cap M$).*

**Definición 35** *Sea $G$ un filtro $P$-genérico y $V$ el universo de todos los conjuntos. Sea $i_G : V^P \to V$ tal que $\forall \tau \in V^P$:*

$$i_G(\tau) = \{i_G(\sigma) : \exists p \in G \text{ tal que } (\sigma, p) \in \tau\}^8.$$

Con este funcional podemos definir la extensión genérica de $M$, dada por un filtro P-genérico $G$ y denotada por $M[G]$ como:

$$M[G] = \{i_G(\tau) : \tau \in M^P\} = i_G[M^P].$$

Nótese que $M^P$, aunque definido originalmente como un conjunto de elementos particulares de $M$, puede pensarse simplemente como un conjunto de símbolos de constante que se agregan al lenguaje de la teoría de conjuntos para

[7] En otras palabras, es una relación tal que $\text{dom}(\tau)$ está contenido en el conjunto de los $P-$nombres y $\text{rang}(\tau)$ en el orden parcial. Se puede dar una definición formal de manera recursiva (ver Kunen [18] pp. 188-189).

[8] Según Kunen [18] (p. 188) ésta también es una definición que se puede formalizar mediante recursión. Sin embargo, para los fines de este capítulo basta con la versión más intuitiva.

extenderlo a un lenguaje $L'$ (conocido como lenguaje de forcing), lo que explica que nos refiramos a ellos como "nombres". En ese sentido, podríamos decir que la función $i_G$ es una interpretación (es decir, una función que a cada nombre le asocia un objeto) entre muchas otras. Sin embargo, es la que nos interesa, pues, como veremos más adelante, existe una relación fundamental entre lo que es verdadero en la extensión genérica y lo que es verdadero en el modelo base.

No es exagerado decir que los $P$-nombres son el ADN de los conjuntos de la extensión genérica, pues así como el núcleo de la célula contiene información sobre un cuerpo que no estará contenido en ella, los $P$-nombres guardan también información sobre un mundo más grande que aquel al que pertenecen. De hecho, los $P-$nombres se introducen con la idea de que toda fórmula de $L'$ esté codificada por una fórmula en el modelo base.

Para ver que $G$ un filtro $P-$genérico tiene un $P$-nombre y, por lo tanto, es elemento de la extensión genérica es necesario introducir un tipo especial de $P-$nombres:

**Definición 36** *Dada $x$, $\check{x} = \{(\check{y}, 1_P) : y \in x\}$ es el nombre canónico de $x$*[9].

Con el principio de $\in -$inducción (ver Kunen [18] p.102) se puede demostrar que $i_G(\check{x}) = x$, ya que dicho principio, que no es sino una generalización del principio de inducción canónico, nos garantiza que todos los conjuntos cumplen una propiedad, si para cualquier conjunto $x$ el hecho de que sus elementos cumplan la propiedad implica que él también la cumple[10].

**Lema 60** $i_G(\check{x}) = x$.

**Demostración.** Sea x tal que para todo $y \in x$, $i_G(\check{y}) = y$. Como $\check{x} = \{(\check{y}, 1_P) : y \in x\}$ y para cualquier filtro $G$, $1_P \in G$, entonces $i_G(\check{x}) = \{i_G(\check{y}) : y \in x\}$. Así que usando la hipótesis inductiva tenemos que $i_G(\check{x}) = x$. De modo que por el principio de $\in -$inducción, dado cualquier conjunto $x$, $i_G(\check{x}) = x$. ∎

Este lema no sólo nos permitirá construir un $P$-nombre para $G$ en $M$, sino también ver que $M[G]$ es una extensión de $M$.

**Lema 61** $M \subset M[G]$.

**Demostración.** Sea $m \in M$. Sea $\check{m}$ el nombre canónico de $m$. Como $i_G(\check{m}) = m$ y $\check{m} \in M^P$, entonces $m = i_G(\check{m}) \in i_G[M^P] = M[G]$. ∎

**Lema 62** *Sea $P$ un orden parcial y $G$ un filtro $P$-genérico. Si $\Gamma = \{(\check{p}, p) : p \in P\}$, entonces $i_G(\Gamma) = G$.*

---

[9] Nótese que ésta también es una definición recursiva y que $\check{\emptyset} = \emptyset$.

[10] En términos más formales, si A es una propiedad acerca de conjuntos, el principio de $\in -$inducción nos dice que: $\forall x[\forall y(y \in x \Longrightarrow A(y)) \Longrightarrow A(x)] \Longrightarrow \forall x A(x)$.

**Demostración.** Lo primero que hay que notar es que $\Gamma$ es un $P$-nombre, pues es un conjunto de pares ordenados con dominio en los nombres canónicos y rango en el orden parcial. Por definición tenemos que $i_G(\Gamma) = \{i_G(\check{p}) : p \in G\}$. Por la observación anterior $i_G(\check{p}) = p$. Así que $i_G(\Gamma) = \{p : p \in G\} = G$ y $G \in M[G]$. ■

Estos dos últimos lemas, junto con el Lema 58, garantizan que si $P$ es un orden parcial frondoso, entonces $M \subsetneq M[G]$. Lo siguiente sería demostrar que $M[G]$ es modelo de ZFE. A partir de la definición de $i_G$ es posible demostrar que $M[G]$ es modelo de los axiomas de extensionalidad, regularidad, par, infinito y elección (ver Kunen [18] p.191). Sin embargo, para la demostración del resto de los axiomas, como separación y potencia, se necesita la noción de forzar.

Si retomamos la analogía con el conjunto de reales, aunque en términos de su expresión binaria, podríamos pensar que las sucesiones finitas de ceros y unos son condiciones que nos describen parcialmente a un número irracional $x$ del intervalo $[0,1]$. En el sentido de que si la sucesión $(1,1,0,1)$ es una condición que describe a $x$, entonces $\frac{1}{2}+\frac{1}{4}+\frac{1}{16} < x < \frac{1}{2}+\frac{1}{4}+\frac{1}{8}$. Es decir, la condición $(1,1,0,1)$ fuerza a que $x$ esté entre $\frac{13}{16}$ y $\frac{7}{8}$. La idea de tener condiciones que fuerzan ciertos tipo de propiedades es lo que caracteriza al método de forcing.

**Definición 37** *Sea $\varphi(x_1, ..., x_n)$ una fórmula del lenguaje de la teoría de conjuntos , donde $x_1, ..., x_n$ son variables libres. Sean $M$ un modelo estándar transitivo y numerable (i.e. un modelo ETN) de ZFE, $P$ un orden parcial sobre $M$, $p \in P$ y $\tau_1, ..., \tau_n \in M^P$. Decimos que "$p$ fuerza a $\varphi$" (escrito como $p \Vdash \varphi(\tau_1, ..., \tau_n)$) si dado cualquier filtro $P$-genérico $G$ tal que $p \in G$ entonces $\varphi(i_G(\tau_1), ..., i_G(\tau_n))$ es válida en $M[G]$ (es decir, $\varphi(i_G(\tau_1), ..., i_G(\tau_n))^{M[G]}$).*

Sabemos que los $P$-nombres son formas de referirnos a los elementos de $M[G]$ a partir de $M$ y que las condiciones están en $M$, ya que son elementos de $P$ que pertenece a $M$ un modelo transitivo. Sin embargo, la definición que hemos dado de forzar involucra a $M[G]$, lo que haría imposible demostrar que $M[G]$ es modelo de ZFE, a menos que la noción de forzar también se pudiera expresar en términos de $M$. Afortunadamente se puede definir una relación auxiliar de forcing, que llamaremos *forcing estrella* y que denotaremos con $\Vdash^*$ .

**Lema 63 (Definibilidad)** $\forall p \in P$, $p \Vdash^* \varphi(\tau_1, ..., \tau_n) \iff p \Vdash \varphi(\tau_1, ..., \tau_n)$.

No daremos aquí la definición formal de *forcing estrella* ni la demostración del lema de definibilidad (véase Kunen [18] pp 195-197). Sin embargo, daremos un ejemplo, expuesto por el propio Solovay [23] (p. 7), que ilustra el funcionamiento de $\Vdash^*$. Sean $p, q \in P$ y $G$ un filtro $P$- genérico con $\Gamma$ su $P$-nombre en $M$ y sea $\varphi$ la fórmula $x_1 \in x_2$, entonces $p \Vdash^* \varphi(\check{q}, \Gamma)$ si y sólo si $M \models p \leq q$. Como. por el lema anterior, $p \Vdash^* \varphi(\check{q}, \Gamma) \Longrightarrow p \Vdash \varphi(\check{q}, \Gamma)$, entonces $p \Vdash^* \varphi(\check{q}, \Gamma) \Longrightarrow (p \in G \Longrightarrow q \in G)^{M[G]}$, lo que quiere decir que podemos garantizar cosas en la extensión genérica a partir de lo que sucede en el modelo base.

Nótese además que tanto en el caso de los reales como en la definición de forzar se trata de condicionales. En forcing si $p \in G$ y $p \leq q$ entonces $q \in G$, en el caso de los reales si (1,1,0,1) es una condición de describe a $x$ entonces (1) también lo es. De hecho, el orden que heredan las sucesiones de ceros y unos de los subconjuntos de naturales finitos ordenados mediante la contención invertida, nos permite decir que si la sucesión $s_n$ describe a $x$ y $s_n \leq s_m$, entonces $s_m$ también describe a $x$.

Cabe señalar que muchas de las demostraciones de forcing se basan en la idea intuitiva de extender el conjunto de reales. Por ejemplo, el conjunto que da cuenta de que hay una extensión del modelo base en el que no se cumple el axioma de elección se conoce como los *reales de Cohen*[11]. Esto explica porqué en forcing se suele trabajar con órdenes parciales definidos sobre conjuntos de sucesiones finitas. A continuación daremos un ejemplo de ello presentado en el artículo de Solovay [23] (p.8).

### 6.2.1. Ejemplo de Forcing: el colapso de un cardinal.

**Definición 38** *Sea $\lambda \in OR - \{0\}$, entonces definimos el siguiente conjunto:*

$$S_\lambda = \{p : p \text{ es función} \wedge |p| < \aleph_0 \wedge dom(p) \subset \omega \wedge rang(p) \subset \lambda\}.$$

Nótese que la contención invertida es una relación de orden sobre $S_\lambda$ (i.e. $p < q \iff q \subset p$) tal que $P_\lambda = \langle S_\lambda, \supseteq\rangle$ es un orden parcial con elemento máximo $1_{P_\lambda} = \emptyset$. Además si $M$ es un modelo estándar, transitivo y numerable de ZFE y $\lambda \in OR \cap M$, entonces $P_\lambda \in M$.

**Lema 64** *Sean $M$ un modelo estándar, transitivo y numerable de ZFE, $\lambda \in OR \cap M$ tal que $\aleph_0 < |\lambda|$ y $P_\lambda$ el orden parcial antes descrito. Sea $G$ un filtro $P_\lambda-$genérico, entonces $G \notin M$.*

**Demostración.** Por el Lema 58 basta demostrar que $P_\lambda$ es frondoso. Sea $p \in P_\lambda$, entonces existe $n \in \omega$ tal que $n \notin dom(p)$. Como $\aleph_0 < |\lambda|$ y $\lambda \in OR$, entonces $\{0, 1\} \subset \lambda$. Así que $r = p \cup \{\langle n,0\rangle\} \in P_\lambda$ y $q = p \cup \{\langle n, 1\rangle\} \in P_\lambda$. Claramente $r < p$ (*i.e.* $p \subset r$) y $q < p$ (*i.e.*$p \subset q$). Además $p \perp q$, ya que $\forall s[s \leq r \implies r \subseteq s \implies \langle n, 0\rangle \in s \implies \langle n, 1\rangle \notin s$ (pues $s$ es función) $\implies q \subsetneq s \implies s \nleq q]$. En resumen, dada $p \in P_\lambda$ existen $r, q \in P_\lambda$ tales que $r \leq p$, $q \leq p$ y $r \perp q$. Así que $P_\lambda$ es frondoso y, por el Lema 58, $G \notin M$. ■

**Teorema 65** *Sean $M$ un modelo estándar, transitivo y numerable de ZFE, $\lambda \in OR \cap M$ $(\aleph_0 < |\lambda|)^M$ y $P_\lambda$ el orden parcial antes descrito. Sea $G$ un filtro $P_\lambda-$genérico, entonces $\bigcup G$ es una función $F : \omega \longrightarrow \lambda$ suprayectiva.*

---

[11] Solovay también agrega mediante forcing un tipo de conjuntos llamados *reales aleatorios* (*random reals*) para demostrar su teorema.

**Demostración.** (1) $F = \bigcup G$ es función. Nótese que $\bigcup G \subset \omega \times \lambda$. Sea $n \in dom(\bigcup G)$ y supongamos que existen $\alpha, \beta \in \lambda$ tales que $\langle n, \alpha \rangle \in \bigcup G$ y $\langle n, \beta \rangle \in \bigcup G$. Afirmamos que $\alpha = \beta$. Como $\langle n, \alpha \rangle \in \bigcup G$ y $\langle n, \beta \rangle \in \bigcup G$ entonces existen $p, q \in G$ tales que $\langle n, \alpha \rangle \in p$ y $\langle n, \beta \rangle \in q$. Por otro lado, por la segunda condición de filtro, si $p, q \in G$ entonces existe $r \in G$ tal que $r \leq p$ y $r \leq q$ ($i.e. \exists r \in G(p \subseteq r \wedge q \subseteq r)$). Así que existe $r \in G$ tal que $\langle n, \alpha \rangle \in r$ y $\langle n, \beta \rangle \in r$. Finalmente, como $r$ es función (pues $r \in G \subset P_\lambda$) entonces $\alpha = \beta$.

(2) $dom(F) = \omega$. Como $dom(\bigcup G) = \bigcup_{p \in G} dom(p)$, basta demostrar que dada $n \in \omega$ existe $p \in G$ tal que $n \in dom(p)$. Sea $D_n = \{q \in P_\lambda : n \in dom(q)\}$. Nótese que $D_n$ está en $M$ por el axioma de separación. Afirmamos que $D_n$ es denso en $P_\lambda$. Sean $n \in \omega$ y $p \in P_\lambda$. Hay que demostrar que existe $q \in D_n$ tal que $q \leq p$. Tenemos dos casos: (i) $n \in dom(p)$. Así que $p \in D_n$. (ii) $n \notin dom(p)$. Sea $q = p \cup \{\langle n, 0 \rangle\}$. Claramente $q$ está en $D_n$ y $q \leq p$. Así que $D_n$ es denso. Por otro lado, hay que recordar que si $G$ es filtro $P_\lambda$−genérico entonces para todo $D \in M$ tal que es denso en $P_\lambda$, $G \cap D \neq \emptyset$. Así que dada $n \in \omega$, $G \cap D_n \neq \emptyset$ y, por lo tanto, dada $n \in \omega$ existe $p \in G$ tal que $n \in dom(p)$.

(3) $F$ es suprayectiva. Sea $\alpha \in \lambda$ y sea $D_\alpha = \{q \in P_\lambda : \exists n \in \omega\ (\langle n, \alpha \rangle \in q)\}$. Por un argumento completamente análogo al del inciso anterior, para ver que $F$ es suprayectiva basta demostrar que si $\alpha \in \lambda$ entonces $D_\alpha$ es denso en $P_\lambda$. Sea $p \in P_\lambda$. Otra vez tenemos dos casos: (i) Existe $n \in \omega$ tal que $\langle n, \alpha \rangle \in p$. Este caso es trivial. (ii) No existe $n \in \omega$ tal que $\langle n, \alpha \rangle \in p$. Como $p \in P_\lambda$, $p$ es una función finita. Así que existe $m \in \omega$ tal que $m \notin dom(p)$. Si tomamos $q = p \cup \{\langle m, \alpha \rangle\}$ tenemos que $q \in D_\alpha$ y $q \leq p$. ■

Nótese que $F \notin M$, ya que si $F \in M$, como $G = \{p \in P_\lambda : p \subset \bigcup G = F\}$ y $M$ es modelo del axioma de separación, entonces $G \in M$! Sin embargo, $F \in M[G]$ ya que $G \in M[G]$ y $M[G]$ es modelo del axioma de unión.

**Corolario 66** $(|\lambda| = \aleph_0)^{M[G]}$.

**Demostración.** Por hipótesis $\lambda$ es un ordinal en $M$ tal que $(\aleph_0 < |\lambda|)^M$. Así que existe $g \in M \subset M[G]$ tal que $g$ es un función inyectiva de $\omega$ en $\lambda$. Además, por el teorema anterior, tenemos que existe $F \in M[G]$ tal que $F$ es una función suprayectiva de $\omega$ en $\lambda$. Como $M[G]$ es modelo de ZFE, entonces en $M[G]$ existe una función biyectiva de $\omega$ en $\lambda$[12]. ■

**Observación 16** *Muchas veces estas funciones, conocidas como funciones de colapso, son las que permiten agregar nuevos reales. Esto quedará más claro cuando se exponga el colapso de Levy, sobre el que se basa la demostración del teorema de Solovay.*

[12] Es por este corolario que el ejemplo que hemos dado se suele conocer como el colapso de un cardinal.

Por otro lado, supongamos que $\sigma$ es un $P-$nombre de $F$ y $p = \langle 0, \beta \rangle$, con $\beta \in \lambda$. Entonces $\sigma(\check{0}) = \check{\beta}$ es una fórmula del lenguaje de forcing $L'$ que interpretada en $M[G]$ es el enunciado $F(0) = \beta$. Este enunciado puede ser verdadero o falso en $M[G]$, pero si $p$ es una condición que describe a $F$ (i.e. $p \in G$) entonces dicho enunciado es verdadero en $M[G]$. Esto lo hemos formalizado como $p \Vdash \sigma(\check{0}) = \check{\beta}$ .

Hemos dicho ya que la noción de forzar es en el fondo un condicional, así que parecería que todo lo que podemos decir respecto a $F$ o un irracional depende del filtro genérico o las condiciones que lo describen. Sin embargo, nótese que para el caso de los irracionales hay propiedades generales que son independientes de las condiciones que los determinan de manera más particular, como por ejemplo que su expansión binaria es infinita y no periódica. Algo análogo sucede con forcing, como se vió con el corolario del teorema anterior.

Como se verá más adelante, el hecho de que cierto tipo de fórmulas se cumplan en $M[G]$ independientemente del filtro genérico de que se trate es fundamental para las pruebas de consistencia relativa, sobre todo cuando involucran afirmaciones respecto a cardinalidad.

Por otro lado, cabe señalar que la noción de forzar sirve esencialmente para demostrar que la extensión genérica es modelo de ZFE, pues, como bien lo señala Solovay y se confirma con el siguiente lema, "un hecho fundamental acerca de forcing es la relación que existe entre la verdad y la relación forzar "(Solovay [23] p.7).

**Lema 67 (Verdad)** *Sea $G$ un filtro P-genérico, entonces*

$$\varphi(i_G(\tau_1), ..., i_G(\tau_n))^{M[G]} \iff \exists p \in G(p \Vdash \varphi(\tau_1, ..., \tau_n)).$$

Este lema tampoco se demostrará (ver Kunen [18] pp.195-197), pero cabe señalar desde ahora que la parte más útil será la ida de este bicondicional.

El siguiente teorema es una muestra de cómo con la definición de forzar y los lemas de definibilidad y verdad se puede ver que $M[G]$ es modelo de ZFE.

**Teorema 68** *Sea $M$ un modelo estándar, transitivo y numerable de ZFE, $P$ un orden parcial en $M$, $G$ un filtro P-genérico sobre $M$. Entonces $M[G]$ es modelo del axioma de separación.*

**Demostración.** Sea $\varphi(x, z, y_1, ..., y_n)$ una fórmula del lenguaje de la teoría de conjuntos y sean $a, b_1, ..., b_n \in M[G]$. Hay que demostrar que existe $c \in M[G]$ tal que $c = \{x \in a : (\varphi(x, a, b_1, ..., b_n))^{M[G]}\}$. Como $M[G] = i_G[M^P]$, hay que encontrar un $P$-nombre en $M$ para la colección $\{x \in a : (\varphi(x, a, b_1, ..., b_n))^{M[G]}\}$. Nótese que $a, b_1, ..., b_n \in M[G] \Longrightarrow \exists \sigma, \tau_1, ..., \tau_n \in M^P$ tales que $i_G(\sigma) = a, i_G(\tau_1) = b_1, ... i_G(\tau_n) = b_n$. Además, como $\Vdash^*$ está en términos del modelo base y $dom(\sigma) \times P \in M$, usando el hecho de que $M$ es modelo del axioma de separación tenemos que existe $\rho \in M$ tal que $\rho = \{\langle \delta, p \rangle \in dom(\sigma) \times P : p \Vdash^* (\delta \in \sigma) \wedge \varphi(\delta, \sigma, \tau_1, ..., \tau_n)\}$. Afirmamos que $\rho$ es el $P-$nombre de $c$, así que hay que demostrar que $i_G(\rho) = c$. Esto se hará por doble contención.

(1) $i_G(\rho) \subset c = \{i_G(\lambda) \in i_G(\sigma) : \varphi(i_G(\lambda), i_G(\sigma), i_G(\tau_1), ..., i_G(\tau_n))^{M[G]}\}$.

Sea $b \in i_G(\rho)$, entonces, por la definición de $i_G$, existe $\delta \in dom(\rho)$ tal que $i_G(\delta) = b$ y existe $p \in G$ tal que $\langle\delta, p\rangle \in \rho$. Por la forma en que se tomó $\rho$, $\langle\delta, p\rangle \in \rho \Longrightarrow p \Vdash^* (\delta \in \sigma) \wedge \varphi(\delta, \sigma, \tau_1, ..., \tau_n) \Longrightarrow p \Vdash (\delta \in \sigma) \wedge \varphi(\delta, \sigma, \tau_1, ..., \tau_n)$ (por el lema de definibilidad). Y como $p \in G$, por la definición de *forzar* ($\Vdash$), $p \Vdash (\delta \in \sigma) \wedge \varphi(\delta, \sigma, \tau_1, ..., \tau_n) \Longrightarrow \varphi(i_G(\delta), i_G(\sigma), i_G(\tau_1), ..., i_G(\tau_n))^{M[G]} \Longrightarrow b = i_G(\delta) \in c$.

(2) $c \subset i_G(\rho)$. Sea $b = i_G(\delta) \in c$, entonces $\varphi(i_G(\delta), i_G(\sigma), i_G(\tau_1), ..., i_G(\tau_n))^{M[G]}$. Así que por el lema de verdad existe $p \in G$ tal que $p \Vdash (\delta \in \sigma) \wedge \varphi(\delta, \sigma, \tau_1, ..., \tau_n)$ y, por el lema de definibilidad y la definición de $\rho$, $\langle\delta, p\rangle \in \rho$ con $p \in G$. De modo que $b = i_G(\delta) \in i_G(\rho)$. ■

Para demostrar que la extensión genérica es modelo del axioma de elección no se necesita la noción de forzar, aunque si es necesario introducir los $P$-nombres de los pares ordenados y una versión particular del axioma de elección.

**Definición 39** *Sean $\sigma$ y $\tau$ $P$-nombres. Entonces:*

$$\begin{aligned} pno(\sigma, \tau) &= \{\langle\sigma, 1_P\rangle, \langle\tau, 1_P\rangle\}. \\ po(\sigma, \tau) &= pno(pno(\sigma, \sigma), pno(\sigma, \tau)). \end{aligned}$$

Nótese que tanto $pno(\sigma, \tau)$ como $po(\sigma, \tau)$ son $P-$nombres. La denominación *pno* (par no ordenado) y *po* (par ordenado) se debe a que con ellos se puede nombrar a los pares y pares ordenados de la extensión genérica, como se ve con el siguiente lema:

**Lema 69** *Sean $i_G(\sigma)$, $i_G(\tau)$ elementos de $M[G]$. Entonces:*

$$i_G(pno(\sigma, \tau)) = \{i_G(\sigma), i_G(\tau)\} \text{ e } i_G(po(\sigma, \tau)) = \langle i_G(\sigma), i_G(\tau)\rangle.$$

**Demostración.** La primera igualdad se desprende inmediatamente de la definición de $i_G$ y del hecho de que $1_P \in G$, para todo $G$ filtro $P-$genérico. La demostración de la segunda se basa en la primera:

$$i_G(po(\sigma, \tau)) = i_G(pno(pno(\sigma, \sigma), pno(\sigma, \tau))) = \{i_G(pno(\sigma, \sigma)), i_G(pno(\sigma, \tau))\}$$

$$= \{\{i_G(\sigma), i_G(\sigma)\}, \{i_G(\sigma), i_G(\tau)\}\} = \{\{i_G(\sigma)\}, \{i_G(\sigma), i_G(\tau)\}\} = \langle i_G(\sigma), i_G(\tau)\rangle$$

■

Así que $pno(\sigma, \tau)$ es el $P$-nombre del par $\{i_G(\sigma)\ , i_G(\tau)\}$ y $po(\sigma, \tau))$ el $P-$nombre del par ordenado $\langle i_G(\sigma), i_G(\tau)\rangle$. Lo siguiente es demostrar que el axioma de elección es equivalente a que *para todo conjunto $X$ exista una función que tenga como dominio a un ordinal y que su imagen contenga a $X$.*

**Lema 70** $AE \Longleftrightarrow \forall x \exists \alpha \in OR \exists f(f : \alpha \to V \wedge x \subset f[\alpha])$.

**Demostración.** ($\Longrightarrow$) Como ($AE \Longrightarrow PBO$) (ver Apéndice A) entonces dado un conjunto $x$ existe $\langle x, <\rangle$ un buen orden. Por el teorema de enumeración, existe a su vez un único ordinal $\alpha$ tal que $\langle \alpha, \in\rangle \cong_f \langle x, <\rangle$. Como $f$ es un isomorfismo entre $\langle \alpha, \in\rangle$ y $\langle x, <\rangle$, claramente $\alpha$ y $f$ cumplen con las condiciones buscadas.

($\Longleftarrow$) Como también tenemos que ($PBO \Longrightarrow AE$), basta demostrar que dado $x$ un conjunto tal que $\exists\alpha \in OR\ \exists f(f : \alpha \to V \wedge x \subset f[\alpha])$, existe $\langle x, <\rangle$ un buen orden. La idea es calcar el orden de $\alpha$. Sin embargo, no sabemos si $f^{-1}$ es función, así que hay que construir una función inyectiva de $x$ en $\alpha$. Nótese que $x \subset f[\alpha] \Longrightarrow \forall y \in x(A_y = \{\beta \in \alpha : f(\beta) = y\} \neq \emptyset)$. Como $\langle \alpha, \in\rangle$ es un buen orden y $\forall y \in x(A_y \neq \emptyset \wedge A_y \subset \alpha)$, sea $g : x \to \alpha$ tal que $g(y) = \min_{\in} A_y$. Nótese que $g$ está bien definida. Además como $f$ es función, entonces $g$ es inyectiva. Por otro lado, sabemos que todo subconjunto no vacío de un conjunto bien ordenado está bien ordenado, así que $\langle g[x], \in\rangle$ es un buen orden. Por lo tanto, $x$ se puede bien ordenar de la siguiente manera:

$$\forall y, z \in x,\ y < z \Longleftrightarrow g(y) \in g(z).$$

■

**Teorema 71** *Sea M un modelo estándar, transitivo y numerable de ZFE, P un orden parcial en M y G un filtro P-genérico sobre M. Entonces M[G] es modelo del axioma de elección.*

**Demostración.** Sea $a \in M[G]$, entonces existe $\sigma \in M^P$ tal que $i_G(\sigma) = a$. Como $dom(\sigma) \in M$ y en $M$ es válido el axioma de elección, entonces existe $\langle dom(\sigma), <\rangle$ un buen orden. Por el teorema de enumeración, existe un único ordinal $\alpha$ tal que $\langle dom(\sigma), <\rangle \cong_f \langle \alpha, \in\rangle$, con $f \in M$. Es decir, podemos enumerar en $M$ a los elementos del dominio de $\sigma$, indexándolos con su imagen bajo $f$. De modo que $dom(\sigma) = \{\delta_\beta : f(\delta) = \beta < \alpha\}$. Sea $\tau = \{po(\check{\beta}, \delta_\beta) : \beta < \alpha\} \times \{1_P\}$. Nótese que $\tau$ es un $P-$nombre en $M$ y que $i_G(\tau) = \{i_G(po(\check{\beta}, \delta_\beta)) : \beta < \alpha\}$. Por el Lema 69 $i_G(po(\check{\beta}, \delta_\beta) = \langle i_G(\check{\beta}), i_G(\delta_\beta)\rangle$, así que $i_G(\tau) = \{\langle \beta, i_G(\delta_\beta)\rangle : \beta < \alpha\}$. Nótese que $i_G(\tau)$ es una función cuyo dominio es $\alpha$ y tal que $a = i_G(\sigma) \subset rang(i_G(\tau))$. ■

## 6.3. Modelo de Solovay

Ya hemos señalado que en la extensión genérica de un modelo estándar, transitivo y numerable de ZFE también se cumplen todos los axiomas de la teoría de conjuntos (incluido el axioma de elección), así que $M[G]$ no puede ser el modelo al que se refiere el teorema de Solovay. Sin embargo, sí es el punto de partida para su construcción.

En términos muy generales, la demostración de Solovay, que no es exagerado calificar de auténtica ingeniería genética conjuntista, se divide en dos partes: (1) La construcción de una extensión del modelo base $M$[13] utilizando el método

[13] Aunque en este caso el modelo base $M$ es modelo de ZFE+$\exists\kappa$($\kappa$ es fuertemente inaccesible).

de forcing con un orden parcial semejante al del colapso de un cardinal (Colapso de Levy). (2) El "adelgazamiento" de la extensión genérica a partir de sus sucesiones de ordinales mediante un principio similar al de la construcción del universo $L$ de los conjuntos definibles. Sin embargo, lo más interesante de la prueba de Solovay es el uso que hace de filtros sobre álgebras de Boole para confirmar que este doble proceso genera un modelo en el que todos los conjuntos de reales son Lebesgue medibles.

### 6.3.1. El colapso de Levy

En esta sección veremos que es posible colapsar todos los cardinales que hay entre $\aleph_o$ y $\kappa$ un cardinal fuertemente inaccesible del modelo base. La idea, como en el ejemplo que dimos del colapso de un solo cardinal, es que a partir del filtro $P$- genérico se construyan las funciones suprayectivas de $\omega$ en cada una de los ordinales (incluidos cardinales según $M$) menores que $\kappa$. Para ello es necesario modificar un poco el orden parcial sobre el que se trabaja.

**Definición 40** *Sea $\lambda \in OR - \{0\}$. Entonces definimos el siguiente conjunto:*

$$Col(\lambda) = \{p : p \text{ es función} \wedge \ |p| < \aleph_0 \wedge dom(p) \subset \lambda \times \omega \wedge rang(p) \subset \lambda\}.$$

Nótese que también en este caso la contención invertida es una relación de orden sobre $Col(\lambda)$(i.e. $p < q \iff q \subset p$) tal que $P_\lambda = \langle Col(\lambda). \supseteq\rangle$ es un orden parcial con elemento máximo $1_{P_\lambda} = \emptyset$. Además si $M$ es un modelo estándar, transitivo y numerable de ZFE y $\lambda \in M$, entonces $P_\lambda \in M$.

**Lema 72** *Sea $G$ un filtro $P_\lambda$−genérico y para toda $\alpha < \lambda$ sea $f_\alpha \subseteq \omega \times \lambda$ tal que $f_\alpha = \{\langle n, \beta\rangle : \langle\langle\alpha, n\rangle, \beta\rangle \in \cup G\}$. Entonces $f_\alpha$ es una función suprayectiva de $\omega$ sobre $\alpha$.*

No daremos la demostración de este lema ya que es completamente análoga a la demostración del colapso de un cardinal (Teorema 65). Sin embargo, es importante señalar que gracias a las modificaciones que se hicieron en el orden parcial, en este caso, el filtro $P$− genérico permite construir una familia de $\lambda$ funciones $\{f_\alpha\}_{\alpha<\lambda}$, donde cada $f_\alpha$ es la función que da cuenta del colapso de $\alpha$. De modo que forcing con $Col(\lambda)$ es en realidad el colapso simultaneo de todos los cardinales menores que $\lambda$. Esto nos conduce de manera natural al siguiente resultado:

**Lema 73** *Sea $G$ un filtro $P_\lambda$−genérico. Entonces $(\lambda \leq \aleph_1)^{M[G]}$.*

**Demostración.** Como $M[G]$ es modelo del axioma de unión y del axioma de separación, entonces para toda $\alpha < \lambda$, $f_\alpha \in M[G]$. Así que $(|\alpha| \leq \aleph_0)^{M[G]}$ y, por definición de $\aleph_1$, $(\lambda \subseteq \aleph_1)^{M[G]}$. Por lo tanto, $(\lambda \leq \aleph_1)^{M[G]}$. ∎

Si tomamos $\lambda = \omega$ tenemos el orden parcial con el que se generan los *reales de Cohen.*

**Definición 41** *Sea $M$ un modelo estándar, transitivo y numerable de ZFE. Sea $P_\omega = \langle Col(\omega), \supseteq \rangle$ y $G$ un filtro $P_\omega$–genérico. Para cada $n \in \omega$ sea $x_n = \{m \in \omega : \langle \langle n, m \rangle, 1 \rangle \in \cup G\} = \{m \in \omega : \exists p \in G(p(\langle n, m \rangle) = 1\}$. $x_n$ se conoce como un real de Cohen.*

$\mathcal{R} = \{x_n : n \in \omega\}$ es el conjunto de reales de Cohen que sirve para demostrar que ZF es consistente con la negación del axioma de elección (ver Jech [15]).

Solovay también usa $Col(\lambda)$ para obtener la extensión genérica a partir de la cual construye su modelo, aunque con $\lambda = \kappa$ el cardinal fuertemente inaccesible de las hipótesis de su teorema. Arziel Levy (ver Kanamori [16] p.126) fue el primero en trabajar con el colapso de cardinales fuertemente inaccesibles, así que $Col(\kappa)$, con $\kappa$ cardinal fuertemente inaccesible, suele conocerse como el colapso de Levy. Uno de los principales resultados respecto al colapso de Levy es que $(\kappa = \aleph_1)^{M[G]}$.

Por otro lado, si con $P_\omega = \langle Col(\omega), \supseteq \rangle$ la extensión genérica contiene al menos una cantidad numerable de reales que no estaban en la modelo base, parece natural esperar que con $P_\kappa = \langle Col(\kappa), \supseteq \rangle$ se tendrán muchos reales nuevos. La definición de *real aleatorio (random real)* que Solovay introduce en su artículo sirve, entre otras cosas, para confirmarlo.

Existen diversas razones por las cuales el colapso de Levy es un buen punto de partida para la construcción del modelo de Solovay. Una de ellas se encuentra ligada a la posibilidad de restringir el universo de los conjuntos a aquellos que son definibles a partir de una sucesión de ordinales.

**Definición 42** *Sea $X \in M[G]$ y sea $S$ la colección de sucesiones infinitas de ordinales en $M[G]$. Decimos que $X$ es un conjunto definible a partir de sucesiones infinitas (numerables) de ordinales en $M[G]$ si y sólo si existe $s \in S$ y $\varphi$ una fórmula del lenguaje de la teoría de conjuntos tales que $X = \{x \in M[G] : M[G] \models \varphi(x, s)\}$. Denotaremos con $OD(S)$ a la clase de los conjuntos definibles a partir de sucesiones infinitas de ordinales en $M[G]$.*

**Definición 43** *La cerradura transitiva de un conjunto $A$, denotada por $ct\ (A)$, es la intersección de todos los conjuntos transitivos que contienen a $A$*[14]*.*

**Definición 44** $HOD(S) = \{X \in M[G] : ct(X) \in OD(S)\}$.

La idea de tener un modelo que sólo contenga los conjuntos que se pueden definir mediante fórmulas del lenguaje de la teoría de conjuntos tuvo sus orígenes en la construcción que hizo Gödel de un modelo de ZFE en el que la hipótesis del continuo es verdadera. Uno de los aspectos más sobresalientes de este modelo (conocido como el universo constructible $L$) es que la definición de potencia de un conjunto es más precisa, lo que en el fondo significa que existe control sobre la forma en que crecen los conjuntos.

[14] Por recursión sobre $n$ se puede definir $\cup_0 = A$ y $\cup_{n+1} = \bigcup(\cup_n)$. De modo que $ct(A) = \bigcup\{\cup_n : n \in \omega\}$. Es decir, para obtener la cerradura transitiva de A se agregan los elementos de sus elementos y luego los elementos de estos y así sucesivamente hasta tener un conjunto transitivo.

La noción de definibilidad a partir de ordinales ciertamente es más restrictiva que la definibilidad sin más, pues en este caso se pide que el conjunto esté definido a partir de una fórmula que involucre ordinales. Sin embargo, es justamente por esto que la restricción de la extensión genérica (producida por $Col(\kappa)$) a la clase de los conjuntos definibles a partir de una sucesión de ordinales en ella es el modelo de Solovay, pues se necesitaba un modelo en el que no sólo hubiera control respecto a cardinalidad, sino también respecto a propiedades de medida.

Lo primero que hay que notar es que las funciones $f_\alpha$ que definimos en el Lema 72 pueden verse como sucesiones numerables de ordinales. De modo que, aunque no sepamos todavía cuán grande es el conjunto de reales en la extensión genérica, sí podemos decir que hay $\kappa$ sucesiones infinitas de ordinales que probablemente permiten definir más conjuntos de reales a partir de ordinales. Más aún, aunque $Col(\kappa)$ no permita saber exactamente cómo se extienden los conjuntos de reales en general, sí podremos responder preguntas respecto a los conjuntos definibles a partir de ordinales, dado que se tiene un cierto control sobre las sucesiones de ordinales.

**Proposición 74** *Si $M$ es un modelo estándar y transitivo de $ZFE+I$ y $M[G]$ la extensión genérica producida por $P_\kappa = \langle Col(\kappa), \supseteq \rangle$, entonces $HOD(S)$ es el modelo de Solovay.*

**Notación 9** *A partir de ahora nos referiremos a $HOD(S)$ como $M_S$.*

**Observación 17** $M_S \subset M[G]$.

La demostración de que $M_S$ es modelo de ZF es análoga a la demostración de que $HOD$ es un modelo interno de ZF (ver Kunen [18] pp. 160-163), así que sólo nos concentraremos en ver cómo se demuestra que en $M_S$ todos los conjuntos de reales son Lebesgue medibles. El colapso de Levy incluye resultados, como el siguiente lema, que son de gran utilidad y que explican porqué la extensión genérica producida por él es un buen punto de partida para la construcción del modelo de Solovay.

Lo primero que cabe señalar es que M[G], aunque definido originalmente a partir del funcional $i_G$, puede verse que el mínimo modelo estándar transitivo y numerable de ZFE que contiene a $M$ y que tiene como elemento a $G$. Esto permite generalizar la noción de extensión de un modelo con respecto a un elemento que no está en él, sin necesidad de hablar de filtros.

**Definición 45** *Dado $x$ un conjunto, $M[x]$ es un modelo de ZFE*[15] *tal que:*

**1** $M \subset M[x]$.

**2** $x \in M[x]$.

**3** *Si $N$ es un modelo estándar y transitivo de ZFE tal que $M \subset N$ y $x \in N$, entonces $M[x] \subset N$.*

[15] Suponiendo que existe, claro está.

**Lema 75** *Sea $X \in M_S = HOD(S)$. Entonces existe una fórmula $\psi$ del lenguaje de la teoría de conjuntos tal que:*

$$x \in X \iff M[x] \models \psi(x).$$

No se dará la prueba de este lema, ya que requiere de muchos resultados previos (ver Jech p.547). Sin embargo, puesto que $X \in M_S$, sabemos que existen $\varphi$ una fórmula del lenguaje de la teoría de conjuntos y $s$ una sucesión de ordinales en $M[G]$ tales que $x \in X \iff M[G] \models \varphi(x, s)$. Así que este lema se puede resumir como: dada $X \in M_S$ existe $\psi$ tal que para toda $x \in X$, $M[G] \models \varphi(x, s) \iff M[x] \models \psi(x)$.

Nótese que esto es similar a la relación que se tiene entre la verdad en el modelo base y la verdad en la extensión genérica[16]. Por ejemplo, para demostrar que la extensión genérica es modelo del axioma de elección se utilizó el hecho de que en $M$ es válido dicho axioma. En general, aunque aquí no lo hemos incluido, cada uno de los axiomas de la teoría de conjuntos es válido en la extensión genérica gracias a que lo es en el modelo base y gracias a la relación de verdad que existe entre ambos mediante la noción de forzar . La contribución más importante del artículo de Solovay consistió en extender esa relación a propiedades de medida mediante el concepto de *real aleatorio sobre el modelo base.*

### 6.3.2. Reales aleatorios

La clasificación de los borelianos del modelo base en aquellos que tienen medida cero y aquellos que no, induce una clasificación de los reales de la extensión genérica que permite demostrar que en $M_S$ todos los conjuntos de reales son Lebesgue medibles.

#### Extensión de los conjuntos borelianos

Ya hemos señalado que en la extensión genérica que se obtiene mediante el colapso de Levy se tendrán nuevos reales. Esto significa que los conjuntos borelianos del modelo base también se extienden.

**Notación 10** $\mathfrak{B}_{\mathbb{R}}^{M}$ *denota la $\sigma-$álgebra de Borel según $M$ y $\mathfrak{B}_{\mathbb{R}}^{M[G]}$ la $\sigma-$álgebra de Borel según $M[G]$.*

**Lema 76** *Si $B \in \mathfrak{B}_{\mathbb{R}}^{M}$ entonces existe $B^* \in \mathfrak{B}_{\mathbb{R}}^{M[G]}$ tal que $B = B^* \cap M$.*

[16]Aunque en este caso se trata de la relación de verdad entre dos modelos que extienden al modelo base.

La demostración de este lema se basa en el hecho, garantizado por ZFE, de que los borelianos se pueden codificar mediante las sucesiones infinitas de naturales, de manera que la sucesión que se le asigne indique cuál ha sido su proceso de construcción. Por ejemplo, sea $(p_n, q_n)_{n\in\mathbb{N}}$ una numeración de los intervalos abiertos con extremos racionales. Si $A$ es abierto entonces existe $I \subset \mathbb{N}$ tal que $A = \bigcup_{i\in I}(p_i, q_i)$ . En la codificación propuesta por Kanamori [16] (p.137) $c_A : \omega \longrightarrow \omega$ es tal que $c_A(0) = 0$ (lo que indica que $A$ es abierto) y $c_A(i+1) = \chi_I$ (lo que indica qué intervalos son parte de su unión). Nótese además que $A^*$ será el boreliano en la extensión genérica construido de la misma manera.

Por otro lado, el inverso de este lema no necesariamente se cumple (i.e. $D \in \mathfrak{B}_{\mathbb{R}}^{M[G]} \not\Rightarrow \exists B(B \in \mathfrak{B}_{\mathbb{R}}^{M} \wedge D = B^*)$. Basta con que exista $c : \omega \longrightarrow \omega$ tal que $c \in M[G]$ y $c \notin M$[17], pues el boreliano que se construye con $c$ estaría en la extensión genérica pero no en el modelo base[18]. En resumen, existe una función $* : \mathfrak{B}_{\mathbb{R}}^{M} \longrightarrow \mathfrak{B}_{\mathbb{R}}^{M[G]}$ que no necesariamente es suprayectiva. Sin embargo, para nuestros fines basta con que cada boreliano del modelo base tenga su extensión en $M[G]$.

**Definición 46** *Sea $x$ un real de $M[G]$. $x$ es un real aleatorio (random real) sobre $M$ si y sólo si para todo $B \in \mathfrak{B}_{\mathbb{R}}^{M}$ tal que $\mu_L(B) = 0$, $x \notin B^*$.*

**Definición 47** $Ra(M) = \{x \in M[G] : x$ *es un real aleatorio sobre* $M\}$.

**Observación 18** *Si denotamos con $\mathbb{R}^*$ a los reales de $M[G]$, entonces $Ra(M) = \mathbb{R}^* - \bigcup\{B^* : B \in \mathfrak{B}_{\mathbb{R}}^{M} \wedge \mu_L(B) = 0\}$.*

Con el siguiente lema veremos que los reales aleatorios sobre $M$ son un ejemplo de los reales que extienden al modelo base. Es decir, son conjuntos que están en $M[G]$ pero no en $M$.

**Lema 77** *Si $x \in Ra(M)$, entonces $x \notin M$.*

**Demostración.** Supongamos que $x \in M$. Como $M$ es modelo de ZFE, $\{x\} \in \mathfrak{B}_{\mathbb{R}}^{M}$ y $\mu_L(\{x\}) = 0$. Pero por el Lema 76 tenemos que $x \in \{x\}^*$! (pues $x \in Ra(M) \Longrightarrow \neg B \in \mathfrak{B}_{\mathbb{R}}^{M}[\mu_L(B) = 0 \wedge x \in B^*]$. ■

Por el siguiente lema, para ver que cualquier conjunto de reales $X$ que está $M_S$ es Lebesgue medible, basta ver que existe $A \in \mathfrak{B}_{\mathbb{R}}^{N}$ tal que $M_S \models \mu_L(X \triangle A) = 0$.

**Lema 78** *Sea $X$ un conjunto de reales. Si existe $A \in \mathfrak{B}_{\mathbb{R}}$ tal que $\mu^*(X \triangle A) = 0$, entonces $X$ es Lebesgue medible.*

[17] Esto es posible dado que la potencia no es una noción absoluta.

[18] Solovay prueba el siguiente resultado desde ZFE: Sea B un conjunto. B es un conjunto boreliano si y sólo existe $c : \omega \to \omega$ que lo codifica (ver Solovay [23] p. 25).

**Demostración.** Sea X un conjunto de reales y $A \in \mathfrak{B}_{\mathbb{R}}$ tal que $\mu^*(X \triangle A) = 0$. Lo primero es observar que $X = (X - A) \cup (A \cap X)$. Así que, como la clase de los conjuntos Lebesgue medibles es una $\sigma-$álgebra, para ver que $X$ es Lebesgue medible basta demostrar que $(X - A)$ y $(A \cap X)$ lo son. Por un lado, tenemos que $X \triangle A = A - X \cup X - A$. Así que, por la monotonía de la medida exterior, como $\mu^*(X \triangle A) = 0$ entonces $\mu^*(X - A) = 0$. Por otro lado, como la medida de Lebesgue es completa entonces cualquier conjunto con medida exterior igual a cero es Lebesgue medible, de modo que $X \triangle A$ y $(X - A)$ son Lebesgue medibles. Así que sólo falta demostrar que $A \cap X$ es Lebesgue medible. Nótese que $A - (X \triangle A) = A \cap X$ . Como $X \triangle A$ es Lebesgue medible, entonces $(X \triangle A)^c$ también lo es y por lo tanto $A - (X \triangle A) = A \cap (X \triangle A)^c$ es Lebesgue medible (por ser intersección de medibles). De modo que $A \cap X$ es Lebesgue medible. ■

Para encontrar el conjunto de Borel correspondiente a cada uno de los conjuntos de reales de $M_S$ se utilizará una sorprendente relación entre los filtros del álgebra de Boole generada por los borelianos del modelo base y los reales aleatorios.

### 6.3.3. Álgebras de Boole

**Definición 48** *Un álgebra de Boole es un conjunto $\mathcal{B}$ con dos funciones binarias $\wedge, \vee$, una función unaria $\neg$ y dos elementos distinguidos 1 y 0 que cumplen lo siguiente:*

1. $\wedge$ ,$\vee$ son operaciones cerradas sobre $\mathcal{B}$, conmutativas, asociativas y distributivas.

2. $x \wedge (x \vee y) = x$.

3. $x \vee (x \wedge y) = x$.

4. $x \vee \neg x = 1$.

5. $x \wedge \neg x = 0$.

En las álgebras de Boole se cumplen otras propiedades que no se incluyeron en la lista, ya que se pueden deducir a partir de aquellas.

**Lema 79** *Sea $\mathcal{B}$ un álgebra de Boole y sean $x, y \in \mathcal{B}$. Entonces (a) $x \vee x = x$, (b) $x \wedge x = x$, (c) $x \vee 1 = 1$, (d) $\neg(x \wedge y) = \neg x \vee \neg y$, (e) $\neg(x \vee y) = \neg x \wedge \neg y$, (f) $\neg\neg x = x$, (g) $x \vee 0 = x$.*

**Demostración.** Sólo demostraremos los incisos $(a)$ y $(c)$.

$(a)$ $x \vee x = x$. Por la propiedad (3) si tomamos $y = \neg x$ tenemos que $x = x \vee (x \wedge \neg x)$. Así que usando la distributividad y la propiedad (4) $x = (x \vee x) \wedge (x \vee \neg x) = (x \vee x) \wedge 1$. Por otro lado, por las propiedades (4) y (2) tenemos que $x' \wedge 1 = x' \wedge (x' \vee \neg x') = x'$. Así que si tomamos $x' = (x \vee x)$ entonces tenemos que $x = x' \wedge 1 = x' = x \vee x$.

(c) $x \vee 1 = 1$. Por la propiedad (4) $x \vee 1 = x \vee (x \vee \neg x)$. Utilizando la asociatividad de $\vee$ y el inciso (a) tenemos que $x \vee (x \vee \neg x) = (x \vee x) \vee \neg x = x \vee \neg x$. De modo que $x \vee 1 = x \vee \neg x$ y por la propiedad (4) podemos conlcuir que $x \vee 1 = 1$. ■

Aunque la notación parece más cercana a la lógica, las funciones $\wedge$, $\vee$ y $\neg$ pueden interpretarse como las operaciones conjuntistas intersección, unión y complemento respectivamente aplicadas al conjunto potencia de un conjunto $X$. De modo que en ese caso el conjunto vacío es el 0, X el 1 y los incisos (d) y (e) del Lema 79 las leyes de De Morgan. De hecho, parte de la contribución de las álgebras de Boole es que se trata de una estructura que permite simplificar las relaciones de verdad entre los modelos de la teoría de conjuntos generados mediante forcing.

En general el método de forcing, incluso para obtener la extensión genérica de un modelo, se puede trabajar con álgebras de Boole, ya que toda álgebra de Boole tiene como orden natural la relación $x \leq_{\mathcal{B}} y \iff x \vee y = y \iff x \wedge \neg y = 0$. De modo que $P_{\mathcal{B}} = \langle \mathcal{B} - \{0\}, \leq_{\mathcal{B}} \rangle$ no es sino el álgebra de Boole vista como orden parcial, con 1 su elemento máximo. En la interpretación conjuntista de un álgebra de Boole la relación $\leq_{\mathcal{B}}$ es la contención (es decir, $x \leq_{\mathcal{B}} y \iff x \subseteq y$).

Por otro lado, para cualesquiera $x, y \in \mathcal{B}$ , como $\mathcal{B}$ es cerrado bajo la operación $\vee$, entonces $x \vee y \in \mathcal{B}$ . Esto quiere decir que el supremo de $A = \{x, y\}$ con respecto a la relación $\leq_{\mathcal{B}}$ está en $\mathcal{B}$, pues es fácil ver que $x \vee y$ es la mínima cota superior. De manera análoga se tiene que $x \wedge y \in \mathcal{B}$ y es la mínima cota inferior de A. Sin embargo, no podemos garantizar que $A$ un subconjunto arbitrario de $\mathcal{B}$ tenga supremo e ínfimo en $\mathcal{B}$.

**Definición 49** *$\mathcal{B}$ un álgebra de Boole es completa si y sólo si para todo subconjunto $A \subseteq \mathcal{B}$, $A$ tiene supremo e ínfimo en $\mathcal{B}$.*

**Notación 11** *$\bigvee A$ denotará el supremo de $A$ con respecto a la relación $\leq_{\mathcal{B}}$ y $\bigwedge A$ denotará al ínfimo de $A$.*

**Observación 19** *Si la relación $\leq_{\mathcal{B}}$ es la contención entre conjuntos, entonces $\bigvee A = \bigcup A$ y $\bigwedge A = \bigcap A$.*

A continuación daremos la definición de filtro sobre un álgebra de Boole, que más adelante se verá que es equivalente a la de filtro $P_{\mathcal{B}}$−genérico.

**Definición 50** *Sea $\mathcal{B}$ un álgebra de Boole. $F \subseteq \mathcal{B} - \{0\}$ es filtro sobre $\mathcal{B}$ si y sólo si:*

(1) $F \neq \emptyset$.
(2) $\forall x, y \in F, x \wedge y \in F$.
(3) $\forall x \in F, \forall y \in \mathcal{B}(x \vee y = y \Longrightarrow y \in F)$[19].

[19]Si pensamos en las operaciones $\wedge$ y $\vee$ como la intersección y la unión respectivamente, el inciso (2) nos diría que si $p, q \in F$ entonces $p \cap q \in F$ y el incisio (3) que si $p \in F$ y $p \subset q$ entonces $q \in F$.

**Observación 20** *$1 \in F$, si $F$ es filtro sobre $\mathcal{B}$. Esto se desprende inmediatamente de las propiedades (1) y (3) de $F$ y del hecho de que para toda $x \in \mathcal{B}$, $x \vee 1 = 1$ (Lema 79).*

**Definición 51** *$F$ es ultrafiltro sobre $\mathcal{B}$ si y sólo si para toda $x \in \mathcal{B}$, $x \in F$ ó $\neg x \in F$.*

**Definición 52** *Sean $M$ un modelo estándar, transitivo y numerable de ZFE, $\mathcal{B} \in M$ un álgebra de Boole, $F$ un ultrafiltro sobre $\mathcal{B}$. $F$ es $M-$completo si y sólo si para todo $S \subseteq \mathcal{B}$ ($S \in M$ y $\bigvee S \in F \Longrightarrow S \cap F \neq \emptyset$).*

Con el siguiente lema queda claro porqué se puede trabajar forcing con álgebras de Boole, pues nos muestra que la noción de ultrafiltro $M-$completo sobre un álgebra de Boole es equivalente a la noción de filtro $P_{\mathcal{B}}-$genérico sobre M.

**Lema 80** *Sean $M$ un modelo estándar, transitivo y numerable de ZFE y $\mathcal{B} \in M$ un álgebra de Boole completa. Sea $F \subseteq \mathcal{B} - \{0\}$. Entonces $F$ es ultrafiltro M-completo en $\mathcal{B}$ si y sólo si $F$ es un filtro $P_{\mathcal{B}}-$genérico sobre $M$.*

**Demostración.** ($\Longrightarrow$) Sea $F$ un ultrafiltro $M$-completo. Lo primero que haremos es demostrar que $F$ es filtro con respecto al orden parcial $\leq_{\mathcal{B}}$. Así que verificaremos cada una de las condiciones que se le piden: (1) $F \neq \emptyset$ por la condición (1) de ultrafiltro. (2) Sean $p, q \in F$ entonces tomamos $r = p \wedge q$, ya que $r \leq_{\mathcal{B}} p$ , $r \leq_{\mathcal{B}} q$ y además $r \in F$ por la condición (2) de ultrafiltro. De modo que $r$ es testigo de la compatibilidad de $p$ y $q$ según el orden $\leq_{\mathcal{B}}$. (3) Sea $p \in F$ y $q \in P_{\mathcal{B}}$ tal que $p \leq_{\mathcal{B}} q$. Por definición de $\leq_{\mathcal{B}}$, $p \leq_{\mathcal{B}} q \Longrightarrow p \vee q = q$. Así que por la condición (3) de ultrafiltro tenemos que $p \leq_{\mathcal{B}} q \Longrightarrow q \in F$. Finalmente para ver que $F$ es $P_{\mathcal{B}}$-genérico hay que demostrar que $F \cap D \neq \emptyset$, para todo $D \in M$ tal que $D$ es denso en $P_{\mathcal{B}}$. Para ello demostraremos primero que $\bigvee D = 1$. Como 1 es el elemento máximo de $P_{\mathcal{B}}$ es obvio que $\bigvee D \leq 1$. Por otro lado, sea $t \in \mathcal{B}$ una cota superior de $D$ según el orden $\leq_{\mathcal{B}}$. Supongamos que $1 \nleq t$, entonces $1 \wedge \neg t \neq 0$. Así que $1 \wedge \neg t = p \in \mathcal{B} - \{0\}$. Además, como por hipótesis $D$ es denso en $P_{\mathcal{B}}$ entonces existe $q \in D$ tal que $q \leq_{\mathcal{B}} p$. En otras palabras, existe $q \in D$ tal que $q \wedge \neg p = 0$. Por otro lado, como $\neg p = \neg(1 \wedge \neg t)$, por los incisos ($d$) (Leyes de De Morgan), ($f$) y ($g$) del Lema 79 tenemos que $\neg p = 0 \vee (\neg\neg t) = 0 \vee t = t$. Así que $q \wedge t = 0$. Pero como $t$ es cota superior de $D$ y $q \in D$ entonces $q \leq_{\mathcal{B}} t$ , así que también tenemos que $q \wedge \neg t = 0$. Juntando los dos últimos resultados tenemos que $q = 0$! (pues $q \in D \subset \mathcal{B} - \{0\}$). De modo que para toda $t$ cota superior de $D$, $1 \leq t$. Por lo tanto, $1 \leq \bigvee D$. Además $1 \in F$ (Observación 20). Así que $\bigvee D \in F$ y por la definición de ultrafiltro $M$-completo tenemos que $F \cap D \neq \emptyset$.

($\Longleftarrow$) Sea $F$ un filtro $P_{\mathcal{B}}-$genérico sobre M. (1) $F \neq \emptyset$ por la condición (1) de filtro $P_{\mathcal{B}}$. (2) Dadas $p, q \in F$, por la condición (2) de filtro $P_{\mathcal{B}}$, existe $r \in F$ tal que $r \leq_{\mathcal{B}} p$ y $r \leq_{\mathcal{B}} q$. Así que $r \vee q = q$ y $r \vee p = p$ y, por lo tanto, $r \vee (p \wedge q) = p \wedge q$. De modo que $r \leq_{\mathcal{B}} p \wedge q$. Como $r \in F$ y $F$ es $P_{\mathcal{B}}-$genérico sobre M entonces $p \wedge q \in F$. La condición (3) se sigue inmediatamente de la definición del orden $\leq_{\mathcal{B}}$ ($x \leq_{\mathcal{B}} y \iff x \vee y = y$) y de la condición (3) de filtro $P_{\mathcal{B}}$–genérico sobre M. (4) Para ver que $F$ es ultrafiltro sea $r \notin F$. Hay que demostrar que $\neg r \in F$. Sea $D_r = \{q \in \mathcal{B} - \{0\} : q \leq_{\mathcal{B}} r$ ó $q \leq_{\mathcal{B}} \neg r\}$. Afirmamos que $D_r$ es denso en $P_{\mathcal{B}}$. Sea $p \in P_{\mathcal{B}} - D_r$ y sea $q = p \wedge r$. Nótese que $q \leq_{\mathcal{B}} r$ y $q \leq_{\mathcal{B}} p$, así que $q \in D_r$ y $q \leq_{\mathcal{B}} p$. De modo que $D_r$ es denso en $P_{\mathcal{B}}$. Además $D_r$ está en $M$, ya que $M$ es modelo del axioma de separación. Como $F$ es $P_{\mathcal{B}}$–genérico entonces $F \cap D_r \neq \emptyset$. Sea $t \in F \cap D_r$. Por la definición de $D_r$, $t \leq_{\mathcal{B}} r$ ó $t \leq_{\mathcal{B}} \neg r$. Como $t \in F$ y $r \notin F$ entonces $t \nleq_{\mathcal{B}} r$ (pues de lo contrario, por la condición (3) de filtro $r \in F$). De modo que $t \leq_{\mathcal{B}} \neg r$ y, por lo tanto, $\neg r \in F$. Sólo falta demostrar que $F$ es $M$–completo. Sea $S \subseteq \mathcal{B}$ tal que $S \in M$ y $\bigvee S \in F$, entonces sea $D_S = \{q \in \mathcal{B} - \{0\} : q \leq_{\mathcal{B}} \neg(\bigvee S)$ ó existe $s \in S$ tal que $q \leq_{\mathcal{B}} s\}$. $D_S$ es denso en $P_{\mathcal{B}}$. Sea $p \in P_{\mathcal{B}}$ tal que $p \wedge \bigvee S \neq 0$. Esto quiere decir que existe $s \in S$ tal que $p \wedge s \neq 0$. Sea $r = s \wedge p$. Nótese que $r \leq_{\mathcal{B}} s$ y $r \leq_{\mathcal{B}} p$. Así que $r \in D_s$ y $r \leq_{\mathcal{B}} p$. Por lo tanto $D_S$ es denso en $P_{\mathcal{B}}$. En este caso $D_S$ también está en $M$ gracias al axioma de separación. De modo que, como $F$ es $P_{\mathcal{B}}$–genérico, $F \cap D_S \neq \emptyset$. Sea $t \in F \cap D_S$. Por definición de $D_S$, $t \leq_{\mathcal{B}} \neg(\bigvee S)$ ó existe $s \in S$ tal que $t \leq_{\mathcal{B}} s$. Pero como $t \in F$ y $\bigvee S \in F$, entonces $t \wedge \bigvee S \in F$ y, por lo tanto, $t \wedge \bigvee S \neq 0$. Así que $t \nleq_{\mathcal{B}} \neg(\bigvee S)$. De modo que existe $s \in S$ tal que $t \leq_{\mathcal{B}} s$. Finalmente como $t \in F$, entonces $s \in F$. Así que $F \cap S \neq \emptyset$. ■

Cabe señalar que es posible definir el conjunto de los $\mathcal{B}$ -nombres en $M$, que denotaremos con $M^{\mathcal{B}}$, de la misma manera en que se definen los $P$–nombres en $M$.

La ventaja de trabajar con álgebras de Boole, en lugar de órdenes parciales, es que podemos simplificar la relación entre verdad y forzar. Esto se hace introduciendo la noción de valor booleano de una fórmula.

**Definición 53** *Sea $\mathcal{B} \in M$ un álgebra de Boole tal que $(\mathcal{B}$ es completa$)^M$. Sean $\varphi$ una fórmula del lenguaje de teoría de conjuntos con $n$ variables libres y sean $\tau_1, ..., \tau_n \in M^{\mathcal{B}}$. Se define el valor de verdad (valor booleano) de $\varphi$, denotado por $\|\varphi(\tau_1, ..., \tau_n)\|$ como:*

$$\|\varphi(\tau_1, ..., \tau_n)\| = \bigvee\{p \in \mathcal{B} : p \Vdash \varphi(\tau_1, ..., \tau_n)\}$$

Nótese que $\{p \in \mathcal{B} : p \Vdash \varphi(\tau_1, ..., \tau_n)\} = \{p \in \mathcal{B} : p \Vdash^* \varphi(\tau_1, ..., \tau_n)\}$ gracias al lema de definibilidad. Además, como $M$ es modelo del axioma de separación, $\{p \in \mathcal{B} : p \Vdash \varphi(\tau_1, ..., \tau_n)\} \in M$ y $\|\varphi(\tau_1, ..., \tau_n)\|$ está bien definido, pues por hipótesis $\mathcal{B}$ es un álgebra de Boole completa según $M$. Para simplificar la notación, a partir de ahora escribiremos el valor booleano de una función $\varphi$ como $\|\varphi\|$, omitiendo la referencia de $\tau_1, ..., \tau_n \in M^{\mathcal{B}}$.

A continuación se presentan un par de resultados que en parte explican porqué es útil trabajar con álgebras de Boole, aunque cuando se analice la prueba de Solovay esto quedará más claro.

**Teorema 81** *Sea $\varphi$ una fórmula y $F$ un ultrafiltro $M$-completo sobre $\mathcal{B}$, entonces*

$$M[F] \models \varphi \iff \|\varphi\| \in F$$

**Demostración.** ($\Longrightarrow$) Como $F$ es un ultrafiltro $M$-completo sobre $\mathcal{B}$, por el Lema 80 sabemos que $F$ es un filtro $P_{\mathcal{B}}-$genérico sobre M, de modo que podemos utilizar todos los resultados de forcing presentados en la Sección 6.2. En particular, por lema de verdad tenemos que $M[F] \models \varphi \Longrightarrow \exists p \in F(p \Vdash \varphi)$. Como $\|\varphi\| = \bigvee\{p \in \mathcal{B} : p \Vdash \varphi\}$, entonces $(M[F] \models \varphi \Longrightarrow \exists p \in F(p \leq_{\mathcal{B}} \|\varphi\|))$. Así que por la condición (3) de filtro tenemos que $M[F] \models \varphi \Longrightarrow \|\varphi\| \in F$.

($\Longleftarrow$) Supongamos que $\|\varphi\| \in F$. Como $\|\varphi\|$ es en realidad el supremo de las condiciones que fuerzan a $\varphi$ y estamos suponiendo que $F$ es un filtro $M$-completo, entonces basta demostrar que $S = \{p \in \mathcal{B} : p \Vdash \varphi\} \in M$, ya que en ese caso $F \cap S \neq \emptyset$. Sin embargo ya habíamos señalado que, gracias a la noción de *forcing estrella*, se puede usar el axioma de separación en $M$ para ver que $S \in M$. ■

### Filtros y reales aleatorios

Solovay introduce el álgebra de Boole generada por la $\sigma-$álgebra de Borel para demostrar que todos los conjuntos de reales de $M_S$ son Lebesgue medibles.

**Definición 54** *Sea $\sim$ la siguiente relación de equivalencia sobre $\mathfrak{B}_{\mathbb{R}}^{M}$ :*

$$\forall B_1, B_2 \in \mathfrak{B}_{\mathbb{R}}^{M}(B_1 \sim B_2 \Longleftrightarrow \mu_L(B_1 \triangle B_2) = 0).$$

**Notación 12** *Sea $B \in \mathfrak{B}_{\mathbb{R}}^{M}$. Denotaremos con $[B]$ la clase de equivalencia de $B$ según $\sim$ y con $\mathcal{B}_{\mathbb{R}}$ al conjunto de todas las clases de equivalencia (i.e. $\mathcal{B}_{\mathbb{R}} = \{[B] : B \in \mathfrak{B}_{\mathbb{R}}^{M}\}$).*

$\mathcal{B}_{\mathbb{R}}$ es un álgebra de Boole completa (ver Jech [15] p. 542). Así que tiene sentido hablar de los ultrafiltros $M-$completos de $\mathcal{B}_{\mathbb{R}}$. Lo interesante es que existe una correspondencia uno a uno entre los reales aleatorios sobre M y los ultrafiltros M-completos de $\mathcal{B}_{\mathbb{R}}$, aunque para el teorema de Solovay es suficiente el siguiente resultado:

**Teorema 82** *Para todo real aleatorio $x$ existe un ultrafiltro $M$-completo $F$ tal que $M[F] = M[x]$ y para todo $B \in \mathcal{B}_{\mathbb{R}}^{M}$ :*

$$x \in B^* \Longleftrightarrow [B] \in F.$$

No daremos la demostración de este teorema (ver Jech [15] p. 542), no sólo porque incluye muchos detalles técnicos, sino porque es más interesante ver que con él se tiene una hermosa aplicación del concepto de valor booleano para la demostración de que todos los conjuntos de reales de $M_S$ son Lebesgue medibles. Simplemente cabe señalar que del mismo modo que existe $\Gamma$ un $P-$nombre estándar para los filtros genéricos, existe un $\mathcal{B}_{\mathbb{R}}$- nombre estándar para los reales aleatorios.

**Teorema 83** *Sea $X$ un conjunto de reales en $M_S$. Entonces existe $A \in \mathfrak{B}_{\mathbb{R}}^{M_S}$ tal que $X \cap Ra(M) = A \cap Ra(M)$.*

Nótese que si demostramos este teorema, para ver que $X$ es Lebesgue medible sólo tendríamos que demostrar que $\mu_L(\mathbb{R}^* - Ra(M)) = 0$[20], pues como $X \cap Ra(M) = A \cap Ra(M)$ entonces $X \triangle A = A \cup X - A \cap X \subset \mathbb{R}^* - Ra(M)$.

**Demostración.** Si $X \in M_S$ entonces por el Lema 75 existe $\psi$ una fórmula del lenguaje de la teoría de conjuntos tal que $x \in X \iff M[x] \models \psi(x)$. Sea $x \in X \cap Ra(M)$. Por el teorema anterior sabemos que existe $F$ un ultrafiltro M-completo sobre $\mathcal{B}_{\mathbb{R}}$ tal que $M[F] = M[x]$ y $x \in B^* \iff [B] \in F$. Así que, como para $x \in X \cap Ra(M)$ existe $F$ un ultrafiltro M-completo sobre $\mathcal{B}_{\mathbb{R}}$ tal que $M[F] = M[x]$, entonces $M[F] \models \psi(x)$. Sea $\tau$ el $\mathcal{B}_{\mathbb{R}}$-nombre estándar de un real aleatorio. Como el valor booleano de una fórmula no depende del filtro, tiene sentido hablar del valor booleano de $\psi(\tau)$. Nótese además que como $\|\psi(\tau)\| \in \mathcal{B}_{\mathbb{R}}$, entonces existe $B \in \mathfrak{B}_{\mathbb{R}}^M$ tal que $\|\psi(\tau)\| = [B]$. Por otro lado, el Teorema 81 nos garantiza que $M[F] \models \psi(x) \iff \|\psi(\tau)\| \in F$ entonces $M[F] \models \psi(x) \iff [B] \in F$. Finalmente, por el teorema anterior, sabemos que $[B] \in F \implies x \in B^*$, de modo que si $x \in X \cap Ra(M)$ entonces $x \in B^*$. Así que $A = B^* \cap M_S$, con $B$ tal que $[B] = \|\psi(\tau)\|$, es el boreliano que buscamos. Hasta aquí sólo hemos demostrado que $X \cap Ra(M) \subset A \cap Ra(M)$. Para la otra contención, sea $x \in A \cap Ra(M)$. Entonces $x$ es un real aleatorio tal que $x \in B^*$. De modo que, por el teorema anterior, existe $F$ un ultrafiltro M-completo sobre $\mathcal{B}_{\mathbb{R}}$ tal que $[B] \in F$ y $M[F] = M[x]$. Como $[B] = \|\psi(\tau)\|$ y $[B] \in F$ entonces $M[F] \models \psi(x)$ y como $M[F] = M[x]$ entonces $M[x] \models \psi(x)$. Así que $x \in X \cap Ra(M)$. Por lo tanto, $A \cap Ra(M) \subset X \cap Ra(M)$. ∎

**Corolario 84** *Todo conjunto de reales en $M_S$ es Lebesgue medible.*

**Demostración.** Sea $A$ el conjunto de Borel dado por el teorema anterior. Hay que ver que $\mu_L(\mathbb{R}^* - Ra(M)) = 0$. Como ser un conjunto de medida cero es una noción absoluta para modelos estándar y transitivos (ver Solovay [23] p. 31) trabajaremos desde $V$, el universo de todos los conjuntos, del mismo modo en que se trabajó con los densos en $M$ un modelo numerable de ZFE para ver que sí hay filtros $P-$genérico (Lema 59). Como $M$ es un modelo numerable, entonces $\mathfrak{B}_{\mathbb{R}}^M$ es numerable. Sea $\{B_n : B \in \mathfrak{B}_{\mathbb{R}}^M \wedge \mu_L(B) = 0\}_{n\in\mathbb{N}}$ una numeración de los borelianos de $M$ que tienen medida cero. Por la $\sigma - aditividad$ de la medida de Lebesgue, $\mu_L(\bigcup_{n\in N} B_n) = 0$. Como la medida de Lebesgue es monótona y $\mathbb{R}^* - Ra(M) \subset \bigcup_{n\in\mathbb{N}} B_n^*$ entonces $\mu_L(\mathbb{R}^* - Ra(M)) = 0$. Por lo tanto, $\mathbb{R}^* - Ra(M)$ tiene medida cero en cualquier modelo transitivo y estándar de ZF. Finalmente, como $(X \triangle A) \subset (\mathbb{R}^* - Ra(M))$ y $M_S \models \mu_L(\mathbb{R}^* - Ra(M)) = 0$ entonces $M_S \models \mu_L(X \triangle A) = 0$. ∎

[20] Ser un conjunto de medida cero es una noción absoluta para modelos transitivos (ver Solovay [23] p.31)

# Conclusiones

Dado que el teorema de Solovay confirma las sospechas de Lebesgue respecto a la responsabilidad del axioma de elección en la negación de la hipótesis de la medida, resulta muy sorprendente que Solovay. después de enunciar su teorema, afirme que "por supuesto que el axioma de elección es verdadero y, por lo tanto, hay conjuntos que no son Lebesgue medibles"[21]. Más aún, si la mayoría de los resultados del análisis matemático sólo requieren del principio de elección numerable y otros, como el teorema de Hahn-Banach para espacios de Banach separables[22], se pueden demostrar mediante el principio de elección dependiente, realmente resulta tentador abandonar las versiones más fuertes del axioma de elección para evitar las paradojas. Sin embargo, los múltiples resultados que hemos tratado a lo largo de esta tesis, como lo son la medida de Banach, el teorema de Banach-Tarski y los teoremas de Ulam. no sólo dan cuenta de la riqueza teórica que se perdería si se abandonara el axioma de elección, sino que se convierten en un argumento contra dicha postura. En ese sentido, podríamos decir que el problema de la medida parece generar. él mismo, una paradoja, pues la búsqueda por el responsable de la negación de la hipótesis de la medida terminó por demostrar la grandeza del culpable.

Por otro lado, cabe señalar que Solovay demostró que hay un modelo en el que se cumple tanto el principio de elección dependiente como el que todos los conjuntos de reales sean Lebesgue medibles, lo que no quiere decir que la hipótesis de la medida se pruebe a partir de dicha versión de elección. De hecho, es imposible que $ZF+ED \vdash HM$ , ya que cualquier modelo de ZFE (incluso el de la hipótesis misma del teorema de Solovay) es modelo de $ZF+ED+\neg HM$. De modo que $ZF+ED \nvdash HM$ y como, gracias a Solovay, sabemos que $ZF+ED \nvdash \neg HM$, tenemos entonces que la hipótesis de la medida es un enunciado independiente de la teoría $ZF+ED$, de modo que tanto él como su negación se pueden agregar como axiomas.

[21] Ver Solovay [23] Pág.3.
[22] Ver Beckenstein y Narici [5].

Hay quienes sugieren, como Garnir [10] la existencia de dos tipos de análisis matemático: el solovayano y el zorniano, identificados con las teorías ZF+HM y ZFE respectivamente. Esta perspectiva parece menos radical que la postura de Solovay, pues apela a la diversidad teórica. En otras palabras, no parece necesario determinar si el axioma de elección es verdadero o no, sino que se trata de explorar qué sucede si se acepta o no, de la misma manera en la que la geometría se exploró dependiendo de si se aceptaba o no el quinto postulado de Euclides sobre las paralelas. Sin embargo, del mismo modo en que la geometría euclidiana se planteaba inicialmente como la más intuitiva y las geometrías no euclidianas como las más bizarras, la hipótesis de la medida negaría un mundo teórico que, aunque inicialmente antintuitivo, es fascinante, como se ve con la paradoja de Banach-Tarski.

Es cierto que habría que explorar más los teoremas de ZF+HM que no lo son de ZFE. Pero por lo pronto no podemos decir que la contribución de los resultados de Solovay se centre en una taxonomía de los distintos tipos de análisis. De hecho, si tomamos en cuenta que hay conjuntos borelianos tales que sus imágenes bajo funciones continuas no son borelianos (ver Hrbacek-Jech [14] p. 278), queda claro que el teorema de Solovay es valioso porque permite aceptar con cierta tranquilidad la existencia de conjuntos no medibles, pues nos señala que todos ellos son producto del axioma de elección.

En resumen, lo que Solovay dice es que hay que celebrar que es el axioma de elección el responsable de la negación de la hipótesis de la medida (claro está siempre y cuando creamos que la teoría de conjuntos es consistente con la existencia de cardinales fuertemente inaccesibles).

Finalmente de esta tesis sólo queda decir que aunque de la esfera de las matemáticas se obtengan dos esferas, la del análisis y la lógica, igual de importantes que la original, los resultados que se han presentado conforman un modelo de la unidad fundamental de la matemática.

$\mathcal{FIN}$

Todos los teoremas de esta tesis son teóricos. Cualquier descomposición paradójica en el mundo físico es milagro.

Ningún animal irracional ha sido maltratado o torturado en la realización de esta tesis.

**DIRECCION**

*Ana Meda Guardiola    José Alfredo Amor Montaño*
*(Palma de Oro)*

**ACTORES INVITADOS**

*Magali Folch    Guillermo Grabinsky    Alejandro Bravo*

**GUION**

*Ana Alvarez Velasco*

**ASISTENTES DE DIRECCION**

*Alejandro Odgers    Miguel Angel Mota*

**EFECTOS ESPECIALES**

*Federico Alvarez    Alicia Silva*

**PRODUCCION**

*Lourdes Velasco    Alejandro Alvarez    Begoña Arregui*

**DIRECCION DE FOTOGRAFIA**

*Myriam Berubé    Juan Andrés Arango*

**CASTING**

*Raúl Saucedo*

**LITERATURA GOTICA**

*Olwen Rowe*

**ESCENOGRAFIA**

*Valentina Rojas Loa    Christian Wiessel*

**DIRECCION DE ARTE**

*Maestro Emilio*

**SONIDO**

*Hugo Solís García*

**VESTUARIO**

*Miriam Gerade*

**MAQUILLAJE**

*Magali Folch Gabayet*

**ASISTENTES TECNICOS**

*Lev Jardón    Lourdes Velasco    Ana Meda*

**GESTION CULTURAL**

*Alicia Gómez*

**GESTION ACADEMICA**

*Andrés Porta    Osvaldo Téllez*

**MUSICA**

| | | |
|---|---|---|
| "Paseo sin rumbo" | "Lucha Villa" | "Las cocineras" |
| Hugo Solís | Rapazound | Los Vega |
| Nueva York | El FARO de Oriente | Edificio Cogordán |

**A TODOS, MIS MAS PROFUNDOS AGRADECIMIENTOS**

La realización de esta tesis no habría sido posible sin la colaboración especial de:

*Compañía de danzas de medio oriente* ***AMRA*** *y el* ***Pan del Escocés.***



*Ciudad de México, 2003.*

110

# Apéndice A

# Axioma de Elección y Buen Orden

La idea de que todo conjunto puede bien ordenarse parece ser bastante natural, pues contar un conjunto no es sino bien ordenarlo. Sin embargo, basta con que consideremos el caso de los reales para que deje de parecernos tan obvio, pues ¿cómo podemos "contar" a los reales?

Lo primero que habría que señalar es que un buen orden para los reales sería algo muy distinto a su orden canónico, ya que en ese caso todo subconjunto no vacío debería tener un elemento mínimo. Pero además veremos que, aunque la demostración de que el axioma de elección implica el principio del buen orden se basa en la idea de que para bien ordenar un conjunto basta biyectarlo con un ordinal[1], se trata de un resultado que, como todos los que se desprenden del axioma de elección, no es constructivo sino puramente existencial.

**Teorema 85** $AE \Leftrightarrow PBO$.
*Demostración.* $\Leftarrow$) *Sea $A$ un conjunto de conjuntos no vacíos. Si aplicamos el principio del buen orden a $\bigcup A$, tenemos una relación $<_{\bigcup A}$ que bien ordena a $A$. Sea $f : A \to \bigcup A$ la siguiente función, $\forall x \in \bigcup A$ :*

$$f(x) = \min_{<_{\bigcup A}} x$$

*Como $\forall x(x \in A \Rightarrow x \neq \emptyset \wedge x \subseteq \bigcup A)$ entonces la función $f$ está bien definida. Claramente $f$ es una función de elección sobre $A$, ya que $\forall x \in A$, $\min_{<_{\bigcup A}} x \in x$.*
■

Para demostrar el regreso introduciremos el teorema de recursión transfinita, que no es sino una generalización del teorema de recursión usual.

[1] Son justamente los números ordinales los que sirven para "contar".

**Definición 55** *Si denotamos con $V$ a la clase de todos los conjuntos, entonces decimos que $G$ es un funcional sobre $V$ si y sólo si $G \subset V \times V$ y $\forall X \in V \exists! Y \in V$ tal que $(X, Y) \in G$. En otras palabras, $G$ es funcional de $V$ si se comporta como función sobre la clase $V$.*

**Teorema 86** *Sea $G$ un funcional de $V$ en $V$. Entonces se puede definir un único funcional $F$ tal que: (i) $domF = OR$, (ii) $\forall \alpha \in OR$, $F(\alpha) = G(F[\alpha])$.*

**Teorema 87** *Un conjunto $A$ es bien ordenable ssi $\wp(A) - \{\emptyset\}$ tiene una función de elección.*

**Demostración.** $\Rightarrow$) Igual que en el teorema anterior, la función de elección está dada por el elemento mínimo del subconjunto no vacío.

$\Leftarrow$) Sea $f$ una función de elección para $\wp(A) - \{\emptyset\}$. Como el universo de los conjuntos no es un conjunto, entonces existe $c \notin A$. Sea $f'$ una función con dominio igual a $\wp(A)$ tal que $\forall X \in \wp(A) - \{\emptyset\}$, $f'(X) = f(X)$ y para $X = \emptyset$, $f'(\emptyset) = c$. Sea G un funcional tal que $\forall X$, $G(X) = f'(A - X)$. Nótese que $\forall X (A - X = \emptyset \iff G(X) = c)$ y que $rang(f') = A \bigcup \{c\}$. Gracias al Teorema 86 se puede definir un único funcional $F$ tal que $dom(F) = OR$ y $\forall \alpha \in OR$ :

$$F(\alpha) = G(F[\alpha]) = f'(A - F[\alpha]).$$

Afirmamos que:

1. $F(\alpha) \neq c \implies \forall \beta \leq \alpha$, $F(\beta) \neq c$. Supongamos que $F(\alpha) \neq c \implies G(F[\alpha]) \neq c \implies A - F[\alpha] \neq \emptyset$. Y como $\forall \beta \leq \alpha$ $F[\beta] \subseteq F[\alpha] \implies A - F[\alpha] \subseteq A - F[\beta]$, entonces $A - F[\beta] \neq \emptyset \implies \forall \beta \leq \alpha$, $F(\beta) \neq c$.

2. $F(\alpha) \neq c \implies \forall \beta \neq \delta \leq \alpha$, $F(\beta) \neq F(\delta)$. Supongamos que $F(\alpha) \neq c$ y que $\beta \neq \delta \leq \alpha$. Sin pérdida de generalidad, sea $\beta < \delta$, entonces $F(\beta) \in F[\delta]$ y como $\delta \leq \alpha$, por el inciso (1), $F(\delta) \neq c \implies G(F[\delta]) \neq c \implies A - F[\delta] \neq \emptyset \implies A - F[\delta] \in \wp(A) - \emptyset \implies F(\delta)$

   $= f'(A - F[\delta]) = f(A - F[\delta]) \implies F(\delta) \in A - F[\delta] \implies F(\delta) \notin F[\delta] \implies F(\delta) \neq F(\beta)$, pues ya habíamos visto que $F(\beta) \in F[\delta]$.

3. Existe $\gamma \in OR$ tal que $F(\gamma) = c$. Supongamos que $\forall \alpha \in OR$, $F(\alpha) \neq c$. Entonces, como el inciso (2) nos asegura que $F$ es inyectiva antes de tomar el valor $c$, tendríamos que $OR$ es biyectable con $F[OR] \subset A$ . Por el axioma de separación $F[OR]$ sería conjunto y por el axioma de reemplazo aplicado a $F^{-1}$ tendríamos que $OR$ es conjunto! Ya que $OR$ es clase.

Finalmente como los ordinales están bien ordenados[2] entonces como, por el inciso (3), $\{\gamma \in OR : F(\gamma) = c\} \neq \emptyset$, sea $\gamma_0 = \min\{\gamma \in OR : F(\gamma) = c\}$. Como $\forall \beta < \gamma_0 \; F(\beta) \neq c$, entonces $F[\gamma_o] \subset A$, ya que $ran(F) = A \bigcup \{c\}$. Por otro lado, $F[\gamma_o] = c \Longrightarrow A - F[\gamma_o] = \emptyset \Longrightarrow A \subseteq F[\gamma_o]$. Así que $F[\gamma_o] = A$. Como, por el inciso (2), $F$ es inyectiva antes de tomar el valor c, entonces $F$ es una biyección entre $\gamma_o$ y $A$. Así que el buen orden para $A$ será el orden que hereda de $\gamma_o$. Es decir, $\forall x, y \in A$, $\; x <_A y$ si y sólo si $F^{-1}(x) \in F^{-1}(y)$, entonces $<_A$ bien ordena a A. ■

Puede ser que se crea que se ha dado un buen orden explícito para A, pero nótese que el ordinal y la biyección que lo generan están garantizados por la función de elección sobre la potencia de A, de la que sólo sabemos que existe.

[2] En el sentido de que toda clase no vacía de ordinales tiene un elemento mínimo con respecto a la pertenencia. Esto se conoce como el Principio del Mínimo Ordinal (P.M.O)

# Apéndice B

# Categoría y medida

Al principio de la tesis se señaló que algunas generalizaciones del conjunto de Cantor demuestran que un conjunto topológicamente despreciable no necesariamente lo es desde un punto de vista geométrico. También se mencionó que estos ejemplos obligaron a abandonar las propuestas sobre criterios de integración basadas en propiedades topológicas[1] e impulsaron el desarrollo de la Teoría de la Medida. A continuación veremos un resultado que muestra con mucha claridad la diferencia entre categoría y medida.

**Definición 56** *Un conjunto $A$ es de primera categoría si y sólo si es unión numerables de conjuntos densos en ninguna parte. En cuyo caso decimos que $A$ es un conjunto flaco*[2].

*Teorema 88* *$\mathbb{R}^1$ se puede partir en dos conjuntos ajenos $A$ y $B$ tales que $A$ tiene medida de Lebesgue cero y $B$ es de primera categoría.*

**Demostración.** Sea $\{q_i\}_{i\in\mathbb{N}}$ una numeración de $\mathbb{Q}$, entonces $\forall i, j \in \mathbb{N}$ definimos los siguientes intervalos:

$$I_{ij} = (q_i - \frac{1}{2^{i+j}}, q_i + \frac{1}{2^{i+j}}).$$

Sea $G_j = \bigcup_{i\in\mathbb{N}} I_{ij}$ y $B = \bigcap_{j\in\mathbb{N}} G_j$. Afirmamos que $B$ tiene medida cero. Dada $\epsilon > 0$, existe $j \in \mathbb{N}$ tal que $\frac{1}{2^j} < \epsilon$ y, por lo tanto, existe $j \in \mathbb{N}$ tal que $\sum_{i\in\mathbb{N}} |I_{ij}| = \sum_{i\in\mathbb{N}} \frac{1}{2^{i+j}} = \frac{1}{2^j} < \epsilon$ y $\mu_L(B) = 0$, ya que $\forall j \in \mathbb{N}$, $B \subset G_j = \bigcup_{i\in\mathbb{N}} I_{ij}$. Así que sólo falta demostrar que $B^c = \mathbb{R} - B = \bigcup_{j\in\mathbb{N}} (G_j)^c$ es de primera categoría.

[1] En particular la idea de Hankel (1870) de que una función era Riemann integrable si y sólo si el conjunto de sus discontinuidades era denso en ninguna parte (ver Hawkings [12] p.32).

[2] "Menger" en inglés.

Nótese que $\forall j \in \mathbb{N}$, por construcción, $\mathbb{Q} \subset G_j$. Además $G_j$ es abierto, por ser unión de abiertos. De modo que $\forall j \in \mathbb{N}$, $G_j$ es un conjunto denso abierto y $\forall j \in \mathbb{N}$, $(G_j)^c$ es denso en ninguna parte, pues supongamos que int $\overline{G_j^c} \neq \emptyset$, entonces existe $B$ abierto no vacío tal que $B \cap G_j = \emptyset$. Por lo tanto, $G_j$ no es denso. Finalmente como $B^c = \bigcup_{j\in\mathbb{N}} (G_j)^c$, entonces $B^c$ es de primera categoría.
■

Nótese que puesto que $\mathbb{R}^1$ es de medida positiva y de segunda categoría, entonces lo que demostramos es que $B$ es de medida cero y de segunda categoría y que $B^c$ es de primera categoría y de medida positiva. Así que efectivamente la noción de ser flaco topológicamente no es equivalente a la de ser nulo geométricamente.

# Apéndice C

# Descomposiciones paradójicas

En general, una descomposición paradójica de un conjunto es una partición que bajo la acción de un grupo arroja dos o más copias del conjunto original. En la Sección 2.2 se demostró que la superficie esférica menos un conjunto numerable tenía una descomposición paradójica bajo la acción del grupo $G_{\tau\sigma}$. Se dijo además que esto era suficiente para ver que la superficie esférica completa se podía partir en una cantidad finita de pedazos que rotados daban dos copias.

Para ver que si $S^2 \backslash D$ es $G_{\tau\sigma}-$paradójico entonces $S^2$ también es $G_{\tau\sigma}-$paradójico es necesario introducir algunas definiciones y resultados preliminares:

**Definición 57** *Sean $A, B \subseteq \mathbb{R}^3$ diremos que $A$ y $B$ son $G$-equidescomponibles ( $A \sim_G B$) si y sólo si existen particiones finitas $\{A_i\}_{i=0}^n$ y $\{B_i\}_{i=0}^n$ de $A$ y $B$ respectivamente tales que si $0 \leq i \leq n$ entonces $A_i \cong_G B_i$.*

En términos menos formales, pero quizás más intuitivos, $A$ y $B$ son $G$-equidescomponibles si $A$ se puede partir en una cantidad finita de pedazos que bajo la acción de $G$ permiten reconstruir a $B$.

Esta definición da un versión alternativa de la noción de ser $G-$ paradójico, en la que se ve con mayor claridad que esta propiedad hace eco de resultados paradójicos que tienen que ver con las distintas acepciones de tamaño de un conjunto:

**Lema 89** *$E$ es $G$-paradójico ssi existen $A$ y $B$ subconjuntos ajenos de $E$ tales que $A \sim_G E$ y $B \sim_G E$.*

**Demostración.** ($\Leftarrow$)Si en lugar de $A$ y $E$ tomamos $\{A_i\}_{i=1}^r$ la partición de $A$, $\{E_i\}_{i=1}^r$ la de $E$ y $G_1 = \{g_i : g_i[A_i] = E_i\}$ el conjunto de rotaciones que dan cuenta de que $A \sim_G E$ y hacemos lo mismo con $B \sim_G E$, donde $\{B_j\}_{j=1}^s$ es la partición de $B$, $\{E_j\}_{j=1}^s$ la de $E$ y $G_2 = \{h_j : h_j[B_j] = E_j\}$, entonces tenemos que $A_1, ..., A_r, B_1, ...B_s$ son subconjuntos ajenos de $E$[1], $G_1 \cup G_2 \subset G$ y $E = \bigcup_{i=1}^{r} g_i[A_i]$ y $E = \bigcup_{j=1}^{s} h_j[B_j]$.

($\Rightarrow$)Para este caso tomamos simplemente $A = \bigcup_{i=1}^{r} A_i$ y $B = \bigcup_{i=1}^{s} B_i$. ■

Como bien los señala Wagon [24] (p.3) detrás de las descomposiciones paradójicas está la idea de duplicar conjuntos. En ese sentido, nótese que si en lugar de la relación $\sim_G$ pensáramos en la relación de equipotencia entre conjuntos, tendríamos frente a nosotros otra de las clásicas paradojas de la matemática relacionada con la noción de tamaño de un conjunto: los números naturales se pueden partir en dos subconjuntos ajenos (los números pares e impares) que son del mismo tamaño (en términos de cardinalidad) que todo el conjunto de naturales[2].

Lo siguiente será demostrar que la equidescomponibilidad es una relación de equivalencia, del mismo modo que la equipotencia[3].

**Teorema 90** *La relación $\sim_G$es una relación de equivalencia sobre $X$.*

**Demostración.** Sólo se demostrará la transitividad, ya que es la que se utilizará más adelante. Sean $A, B, C \subseteq X$ si $A \sim_G B$ y $B \sim_G C$ entonces $A \sim_G C$. Sean $\{A_i\}_{i=0}^{n}$, $\{B_i\}_{i=0}^{n}$ y $\{g_i\}_{i=0}^{n}$ las particiones y elementos del grupo que dan cuenta de que $A \sim_G B$ y sean $\{B'_j\}_{j=0}^{m}$, $\{C_j\}_{j=0}^{m}$ y $\{h_j\}_{j=0}^{m}$ las que dan cuenta de que $B \sim_G C$. Como $G$ es un grupo tenemos que los $g_i$ tienen inversos en $G$. Daremos particiones dentro de las particiones de A y C. Sea $A_{ij} = g_i^{-1}[B_i \cap B'_j]$ y $C_{ij} = h_j[B_i \cap B'_j]$. Está claro que $A_{ij} \cap A_{lk} = \emptyset$ y $C_{ij} \cap C_{lk} = \emptyset$ si $i \neq l$ o $j \neq k$, ya que en esos casos $B_i \cap B_l = \emptyset$ o $B'_j \cap B'_k = \emptyset$. Sea $f_{ji} = h_j g_i$. Nótese que $f_{ji} \in G$, ya que G es grupo. Si tomamos $P = \{A_{ij} : B_i \cap B'_j \neq \emptyset\}$ y $Q = \{C_{ij} : B_i \cap B'_j \neq \emptyset\}$ es fácil ver que $P$ y $Q$ son particiones de $A$ y $C$ respectivamente. Pues sólo falta comprobar que $\bigcup P = A$ y $\bigcup Q = C$. Lo primero es observar que $\forall 0 \leq i \leq n$ $B_i = \bigcup_{j=0}^{m} B_i \cap B'_j$, así que $A_i = g_i^{-1}[B_i] = g_i^{-1}(\bigcup_{j=0}^{m} B_i \cap B'_j) = \bigcup_{j=0}^{m} g_i^{-1}[B_i \cap B'_j] = \bigcup_{j=0}^{m} A_{ij}$. Por otro lado, como $\{A_i\}_{i=0}^{n}$ es por hipótesis partición de A, entonces $\bigcup_{i=0}^{n} A_i = A$ y, como $\bigcup_{i=0}^{n} A_i = \bigcup_{i=0}^{n} \bigcup_{j=0}^{m} A_{ij} = \bigcup P$, tenemos que $\bigcup P = A$. El caso de $C$ es análogo. Finalmente las $f_{ji}$ correspondientes a los elementos de $P$ y $Q$ permiten ver que $A \sim_G C$. ■

---

[1] $\{A_i\}_{i=0}^{r}$ y $\{B_i\}_{i=r}^{s+r}$ son particiones de $A$ y $B$ subconjuntos ajenos de C.

[2] Este antintuivo resultado se podía resolver de dos maneras:(1) rechazando la posibilidad de pensar en la cardinalidad de esos conjuntos como un "número", (2) aceptando que las reglas de la aritmética transfinita son distintas a las de la aritmética finita. De alguna manera frente al problema de la medida se tiene una disyuntiva análoga: (1) Aceptar que hay conjuntos no medibles, (2)Modificar las propiedades de la medida. Sin embargo, hemos visto que en este último caso la segunda opción es demasiado costosa y con los resultados de Solovay queda claro, además, que no resuelve nada.

[3] Aunque en el caso de la equipotencia se trate de un relacional, es interesante ver las fuertes anologías entre una noción y otra. Para profundizar en ello véase el Capítulo 3 de Wagon [24]
.

**Corolario 91** *Si $A$ y $B$ son $G$-equidescomponibles y $A$ es $G$-paradójico entonces $B$ es $G$-paradójico.*

**Demostración.** Como $A$ es $G$-paradójico entonces por el Lema 89 existen $P_1$ y $P_2$ subconjuntos ajenos de $A$ tales que $P_1 \sim_G A$ y $P_2 \sim_G A$. Por otro lado como $A$ y $B$ son $G$-equidescomponibles, sean $\{A_i\}_{i=0}^n$ , $\{B_i\}_{i=0}^n$ y $\{g_i\}_{i=0}^n$ las particiones y elementos de $G$ que dan cuenta de ello. Entonces tomamos $B_1 = \bigcup g_i[A_i \cap P_1]$ y $B_2 = \bigcup g_i[A_i \cap P_2]$. Claramente $B_1$ y $B_2$ son subconjuntos de B y como $P_1$ y $P_2$ son ajenos también tenemos que $B_1$ y $B_2$ lo son, pero además $B_1 \sim_G P_1$ y $B_2 \sim_G P_2$, así que usando la transitividad de $\sim_G$ tenemos

que $B_1 \sim_G B$ y $B_2 \sim_G B$. Utilizando otra vez el Lema 89, tenemos que $B$ es $G$-paradójico. ■

De modo que, como ya se vio que $S^2 \setminus D$ es $S0_3$−paradójico, para ver que $S^2$ es $S0_3$−paradójico basta ver que:

**Teorema 92** *$S^2$ es $SO_3$ equidescomponible con $S^2 \setminus D$ .*

**Demostración.** Como $D$ es un conjunto numerable, sea $l$ una línea que pase por el origen tal que $l \cap D = \emptyset$. Sea $G_l$ el grupo de rotaciones en $\mathbb{R}^3$ con eje de rotación $l$ y ángulo de rotación entre -$\pi$ y $\pi$ radianes módulo $2\pi$. Sea $p \in D$ y $A_p = \{\rho : \rho \in G_l \wedge \rho(p) \in D\}$. En otras palabras sea $A_p \subset G_l$ tal que $O_{A_p} = O_{G_l} \cap D$. Nótese que $A_p$ es numerable, pues como los ángulos de rotación están entre -$\pi$ y $\pi$, todo elemento de $D$ proviene de una rotación de $p$. Sea $A = \bigcup\limits_{p \in D} A_p$. Como A también es numerable, por ser unión numerable de numerables, y $|G_l| = c$, entonces $G_l - A \neq \emptyset$. Sea $\tau \in G_l - A$. Entonces $\tau(D) \cap D = \emptyset$, pues de otro modo $\tau$ pertenecería a A. Por lo tanto $\tau(D) \subset S^2 \setminus D$ y $S^2$ es $SO_3$ equidescomponible con $S^2 \setminus D$, ya que $S^2 = S^2 \setminus D \bigcup D$ y $S^2 \setminus D = S^2 \setminus D \bigcup \tau(D)$, con $\tau \in S0_3$. ■

Con esto completamos el detalle que faltaba para convencernos de que la superficie esférica de la bola unitaria se puede partir en una cantidad finita de pedazos que bajo rotaciones nos permiten obtener dos superficies esféricas del mismo tamaño que la original.

Finalmente, como toda partición de la superficie esférica induce (mediante la proyección radial) una partición sobre la bola unitaria sin el origen , tenemos que $B^3 - \{0\}$ también es $S0_3$−paradójico. Así que, por el Corolario 91, para ver que la bola unitaria completa es un conjunto $SO_3$−paradójico , basta demostrar que $B^3 - \{0\}$ es equidescomponible con $B^3$ (ver Wagon [24]).

# Apéndice D

# Teorema de Hahn-Banach

Los comentarios que Solovay hace respecto al teorema de Hahn-Banach llaman especialmente la atención, a pesar de no ser muy exhaustivos, pues sugieren una relación interesante entre categoría y medida. Detallar el argumento de Solovay es un ejemplo de los problemas abiertos que deja esta tesis. Sin embargo, identificaremos los detalles que se tendrían que completar para seguir el hilo de su reflexión.

Solovay señala que del inciso (3) de su teorema (*Todo conjunto de reales tiene la propiedad de Baire*) se puede deducir que "no hay una medida de probabilidad finito-aditiva definida en la potencia de los naturales que sea cero en los unitarios [singletons]. De modo que el teorema de Hahn -Banach no es válido en el modelo"[1] (Solovay [23] p. 3).

La medida de Banach (Sección 2.1) permite entender parte de este enigmático comentario, pues $\mu_B$ es una medida de probabilidad finito-aditiva sobre el intervalo $[0,1)$ que se construyó utilizando el teorema de Hahn-Banach. Como el intervalo $[0,1)$ es biyectable con $\wp(\mathbb{N})$, el teorema de Hahn-Banach también garantiza la existencia de una medida de probabilidad finito-aditiva para $\wp(\mathbb{N})$ que se anula en los unitarios. Esto quiere decir que si en el modelo de Solovay no se pueden tener medidas de probabilidad finito aditivas para $\wp(\mathbb{N})$, entonces el teorema de Hanh-Banach no es verdadero ahí.

El siguiente teorema muestra el vínculo entre este tipo de medidas y la propiedad de Baire.

**Teorema 93** *Si $\mu$ es una medida sobre una $\sigma-$álgebra $\Im \subset \mathbb{R}^1$ que es finito-aditiva, pero no $\sigma-$aditiva, entonces hay un conjunto de reales que no tiene la propiedad de Baire.*

No daremos los detalles de la demostración (ver Wagon [24] p.211). Sin embargo, cabe señalar que en ella no se hace uso del axioma de elección. De modo que se cumple en el modelo de Solovay.

[1] Solovay [23] p. 3.

En otras palabras, puesto que en el modelo de Solovay todos los conjuntos de reales tienen la propiedad de Baire y el teorema que hemos enunciado se cumple, entonces ahí no hay medidas de probabilidad que se anulen en los unitarios y que sean finito aditivas, pero no $\sigma$-aditivas.

Nótese que para concluir que el teorema de Hahn Banach no se cumple en dicho modelo faltaría demostrar que la medida de Banach no es $\sigma - aditiva$ sin recurrir a las versiones del axioma de elección que niegan la hipótesis de la medida, pues de otro modo, como esas versiones no están garantizadas en el modelo de Solovay, podría ser que en dicho modelo la medida de Banach fuera $\sigma - aditiva$ y, por lo tanto, no se podría concluir lo que sugiere Solovay por la vía que hemos expuesto.

En resumen, si $MP$ denota al enunciado *Existe una medida de probabilidad finito aditiva tal que* $\mu(\{a\}) = 0$, $PB$ al enunciado *Todos los conjuntos tienen la propiedad de Baire y* $M_s$ al modelo de Solovay, lo que tenemos es:

1. $ZF \vdash$ (Hahn-Banach $\Longrightarrow$Medida de Banach ($MP$)).

2. $ZF \vdash$ ($MP$+$\neg\sigma$-aditiva) $\Longrightarrow \neg PB$ ó $ZF \vdash (PB \Longrightarrow \neg(MP$+$\neg\sigma$-aditiva).

3. Como $M_s \models PB$ y $M_s$ es modelo de ZF, entonces $M_s \models \neg(MP + \neg\sigma$-aditiva).

4. Si ($ZF \vdash$ Medida de Banach no es $\sigma$-aditiva)[2], entonces $M_s \not\models$ Hahn-Banach.

Finalmente sólo cabe mencionar que, aunque el teorema de Hahn-Banach no se cumpla en el modelo de Solovay (i.e. no es teorema del análisis solovayano), el hecho de que el teorema de Hahn-Banach para espacios de Banach separables quede garantizado por el principio de elecciones dependientes (ver Beckenstein y Narici [5]) ya es bastante poderoso.

[2] Lo que faltaría demostrar.

# Bibliografía


[1] Amor Montaño, José Alfredo. *La hipótesis generalizada del continuo y su relación con el axioma de elección.* Crítica, Revista Hispanoamericana de Filosofía. Vol XXI. No. 62. Agosto 1989. México.

[2] Amor Montaño, José Alfredo. *Forcing y pruebas de independencia.* Aportaciones Matemáticas. Comunicaciones 9. XXIII Congreso SMM. México.1990.

[3] Amor Montaño, José Alfredo. *Pequeños grandes cardinales.* Tesis de Maestría. UAM-Iztapalapa. 1984.

[4] Bachman, George y Narici, Lawrence. *Functional Analysis.* Academic Press. New York. 1972.

[5] Beckenstein, Edward y Narici, Lawrence. *The Hahn-Banach theorem: the life and times.* Proceeding of the First Ibero-American Conference on Topology and its Aplications (Bencassim, 1995).Topology and its Applications.77. No 2. pp. 193-211. North Holland. 1997.

[6] Campero, Gabriela. *¿Es V distinto de L? Independencia del Axioma de Constructibilidad y algunas reflexiones sobre la no constructibilidad del universo conjuntista.* Tesis Licenciatura. UNAM. 1998.

[7] Devlin.*The joy of Sets.* 2a. Edición. Springer-Verlag. New York. 1993.

[8] Enderton, Herbert. *A mathematical introduction to logic.* 2a. Edición. Academic Press. E.E.U.U. 2001.

[9] Folland, Gerald. *Real Analysis.* John Wiley and Sons. E.E.U.U.1984.

[10] Garnir, H.G. *Functional analysis and its aplications: Solovay's axiom and functional analysis.* Lecture notes in mathematics and its applications.Vol. 399. pp 189-204. Springer-Verlag. Berlin. 1974

[11] Halmos, Paul. *Measure Theory.* 4a. Edición. Van Nostrand Company. E.E.U.U. 1956.

[12] Hawkings, Thomas. *Lebesgue Theory of Integration.* Chelsea Publishing Company. New York. 1975.

[13] Hernández Hernández, Fernando. *Teoría de Conjuntos.* Aportaciones Matemáticas. Textos Nivel Medio 13. SMM. México 1998.

[14] Hrbacek, Karel y Jech, Thomas. *Introduction to set theory.* 3a. Edición. Marcel Dekker. New York. 1999.

[15] Jech, Thomas. *Set Theory.* Academic Press. New York. 1978.

[16] Kanamori. Akihiro. *The higher infinite.* Sringer-Verlag. Berlin. 1997.

[17] Kelly, John. *General Topology.* Springer-Verlag. New York. 1955.

[18] Kunen, Kenneth. *Set Theory. An introduction to independence proofs.* North Holland. Amsterdam. 1980.

[19] Oxtoby, John. *Measure and Category.* 2a. Edición. Springer-Verlag. New York. 1987.

[20] Rubin, H y Rubin, J.E. *Equivalents of the axiom of choice.* North Holland. Amsterdam. 1985.

[21] Rudin, Walter. *Principles of mathematical analysis.* 2a. edición. Mc Graw Hill. México. 1965.

[22] Stromberg, Karl. *Universally Nonmesurable Subgroups of* $\mathbb{R}$. Amer. Math. Monthly 99. 1992.

[23] Solovay, Robert. *A model of set theory in which every set of reals is Lebesgue measurable.* Annals of Mathematics. 92. 1970.

[24] Wagon, Stan. *The Banach-Tarski Paradox.* Cambridge University Press. E.E.U.U. 1985.

[25] Zamora Carrillo, Jerónimo. *Reflexiones sobre Forcing: metamatemática y aplicaciones.* Tesis de Maestría U.A.M. 1995.