Universidad de San Carlos de Guatemala
Escuela de Ciencias Físicas y Matemáticas
Departamento de Matemática

UNA DEMOSTRACIÓN DEL TEOREMA DE LOS NÚMEROS PRIMOS USANDO ÁLGEBRAS DE BANACH

Cristian José Álvarez Bran

Asesorado por M.Sc. Alejandro José Vargas de León

Guatemala, julio de 2021

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA

ESCUELA DE CIENCIAS FÍSICAS Y MATEMÁTICAS

UNA DEMOSTRACIÓN DEL TEOREMA DE LOS NÚMEROS PRIMOS USANDO ÁLGEBRAS DE BANACH

TRABAJO DE GRADUACIÓN
PRESENTADO A LA JEFATURA DEL
DEPARTAMENTO DE MATEMÁTICA
POR

CRISTIAN JOSÉ ÁLVAREZ BRAN
ASESORADO POR M.SC. ALEJANDRO JOSÉ VARGAS DE LEÓN

AL CONFERÍRSELE EL TÍTULO DE
LICENCIADO EN MATEMÁTICA APLICADA

GUATEMALA, JULIO DE 2021

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
ESCUELA DE CIENCIAS FÍSICAS Y MATEMÁTICAS

CONSEJO DIRECTIVO

DIRECTOR	M.Sc. Jorge Marcelo Ixquiac Cabrera
SECRETARIO ACADÉMICO	Ing. José Rodolfo Samayoa Dardón

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

EXAMINADOR	Lic. Ronald Oliverio Chubay Gallina
EXAMINADOR	Lic. Frank Jorge Fritzsche Barrios
EXAMINADOR	Lic. Rubén Darío Narciso Cruz

Ref. D.DTG. 002-2021
Guatemala 01 de julio de 2021

El Director de la Escuela de Ciencias Físicas y Matemáticas de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Coordinador de la Licenciatura en Matemática Aplicada, al trabajo de graduación titulado: **"UNA DEMOSTRACIÓN DEL TEOREMA DE LOS NÚMEROS PRIMOS USANDO ÁLGEBRAS DE BANACH"** presentado por el estudiante universitario **Cristian José Álvarez Bran**, autoriza la impresión del mismo.

IMPRÍMASE.

"ID Y ENSEÑAD A TODOS"

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA · ESCUELA DE CIENCIAS FÍSICAS Y MATEMÁTICAS · DIRECCIÓN -USAC-ECFM-

M.Sc. Jorge Marcelo Ixquiac Cabrera
Director

201604169
JMIC/Paola

AGRADECIMIENTOS

Agradezco a mis papás por apoyarme y velar por mi desarrollo académico y personal, instándome a perseguir la mejor versión de mí. A mis hermanos, por servir como mis primeros amigos y ejemplos a seguir. Agradezco a mis profesores por compartir su conocimiento académico y por ampliar mi perspectiva de la vida laboral de un matemático.

También, agradezco al grupo de olimpiadas universitarias de matemática de la Universidad de San Carlos, alias «Los Tropikongs», por compartir conmigo mañanas y tardes de estudio, muchas grandes ideas, mucha matemática, pero sobretodo mucha amistad. Entre ellos, en especial quiero agradecer a Damián Ochoa y a Luis Mack, por motivarme a no dudar de mis capacidades y trabajar duro durante toda la carrera. Finalmente, quiero agradecer a mi novia, por apoyarme incondicionalmente durante toda la carrera, contribuir enormemente en mi felicidad y desarrollo personal.

DEDICATORIA

A mis papás y hermanos,

por su apoyo casi literalmente en todas las formas posibles. Nunca me hizo falta nada y gracias a ustedes mi mente está llena de queridas memorias.

A Sofía Garcia,

por compartir conmigo su felicidad, su amistad y su amor. Estos hicieron que cada momento junto a ti me diera paz y que todas nuestras experiencias juntos sean recuerdos muy preciados.

A Damián Ochoa,

como tutor, por apoyarme a alcanzar mis metas como competidor de olimpiadas de matemática y despedir esa etapa con la frente en alto, con muchas ansias de recibir mi siguiente reto académico. Como amigo, por estar presente en momentos complicados y compartir tantas risas y hamburguesas.

A Alejandro Vargas,

por aserorar mi tesis, pero principalmente por ser el primer impulsor de mi gusto hacia la matemática y continuar compartiendo su entusiasmo por la búsqueda de belleza y sentido en la misma.

A la ECFM,

por darle un punto de convergencia a la gente apasionada por la ciencia en Guatemala.

ÍNDICE GENERAL

ÍNDICE DE FIGURAS III

LISTA DE SÍMBOLOS V

OBJETIVOS VII

INTRODUCCIÓN IX

1. Conceptos elementales de topología y teoría de la medida **1**

1.1. Topología 1

1.2. Funciones continuas 2

1.3. Conjuntos compactos 3

1.4. Conjuntos y funciones medibles 6

1.5. Medida 8

1.6. Integración 9

1.7. Derivada de Radon-Nikodym 12

2. Tópicos de teoría de números y análisis **13**

2.1. Propiedades básicas de números primos 13

2.2. Inversión de Möbius 15

2.3. Notación de Landau 17

2.4. Estimaciones importantes 18

2.5. Fórmula de simetría de Selberg 21

3. Tópicos de análisis funcional **23**

3.1. Completación de un espacio normado 24

3.2. Topología débil* 26

3.3. El espacio $C_c(\mathbb{R})$ 28

3.4. Álgebras de Banach 29

3.5. Caracteres 31

4. Análisis de Fourier **35**

4.1. Definición y propiedades de la transformada de Fourier 35

4.2. Caracteres de $L_1(\mathbb{R})$ 41

5. Convolución y convergencia de medidas **45**

5.1. Medidas de Radon 45

5.2. Teoremas de Convergencia 47

6. Demostración del Teorema **53**

6.1. Primera Reformulación del Teorema 53

6.2. Segunda Reformulación 55

CONCLUSIONES **65**

RECOMENDACIONES **67**

BIBLIOGRAFÍA **69**

ÍNDICE DE FIGURAS

5.1. f_l con $l = 3$. 50

6.1. Funciones g_+ y g_- 58

LISTA DE SÍMBOLOS

Símbolo	Significado
$:=$	es definido por
$\cong$	es isomorfo a
$E \setminus F$	diferencia entre E y F
$\mathcal{P}(X)$	conjunto potencia de X
$\mathcal{X}$	σ-álgebra de conjuntos
$\mathcal{B}$	σ-álgebra de Borel
μ_ℓ	medida de Lebesgue
Λ	función de Von Mangoldt
$L_1(X)$	espacio de funciones integrables
χ_E	función característica de E
$\underset{D}{*}$	Convolución de Dirichlet
O	O-grande
o	o-pequeña
$\|\| \cdot \|\|$	norma o seminorma
X^*	espacio dual de X
$\overset{*}{\rightharpoonup}$	convergencia débil*
$\sigma(x)$	espectro de x
$r_\sigma(x)$	radio espectral de x
$\sigma(\mathcal{A})$	espacio de caracteres $\psi\colon \mathcal{A} \to \mathbb{C}$
$\operatorname{supp}(f)$	soporte de f
$C(\mathbb{R})$	espacio de funciones continuas
$C_c(\mathbb{R})$	espacio de funciones continuas de soporte compacto
$C_c^\infty(\mathbb{R})$	espacio de funciones suaves de soporte compacto
$\mathscr{S}(\mathbb{R})$	espacio de Schwartz
$\widehat{f}$	transformada de Fourier de f
$\check{f}$	transformada inversa de Fourier de f

Símbolo	Significado
δ_p	medida de Dirac en p

OBJETIVOS

General

Desarrollar conceptos de espacios y álgebra de Banach para demostrar el teorema de los números primos.

Específicos

1. Desarrollar las nociones básicas de topología, teoría de la medida y teoría de integración.

2. Introducir los conceptos básicos de teoría de números y teoría de estimación, así como también la importante conexión entre ambas, la Fórmula de Simetría de Selberg.

3. Presentar los conceptos necesarios de análisis funcional para definir álgebras de Banach y elementos afines. Demostrar la fórmula de Gelfand.

4. Establecer resultados que muestren el comportamiendo de la transformada de Fourier y la convolución de funciones en los espacios $C_c(\mathbb{R})$, $\mathscr{S}(\mathbb{R})$ y $L_1(\mathbb{R})$.

5. Definir medidas de Radon, convolución de medidas, convergencia de medidas como funcionales de $C(\mathbb{R})$.

6. Demostrar el teorema de los números primos usando álgebras de Banach, análisis de Fourier, teoría de números y teoría de la medida.

INTRODUCCIÓN

El Teorema Fundamental de la Aritmética provee una motivación de por qué los números primos son tan importantes «Todo número natural mayor a 2 es expresable de manera única como producto de números primos». Por esto, se puede considerar a estos números como los bloques fundamentales de la estructura multiplicativa de los enteros $\mathbb{Z}$. Una gran cantidad de las preguntas relevantes en el área de Teoría de Números pueden reducirse a su estudio sobre los primos.

Desde tiempos de antaño los matemáticos estaban conscientes de la relevancia de los números primos. Por ejemplo, Euclides (aproximadamente 300 AC) conocía que existen infinitos primos. Con el paso del tiempo, resultados y aplicaciones igualmente interesantes acerca de estos números han surgido. Sin embargo, uno de los misterios que aún elude nuestro conocimiento es su distribución o ubicación dentro de los números naturales; si bien existen algoritmos para determinar si un número dado es primo, no existe un procedimiento rápido para «hallar el siguiente primo», pues su ocurrencia es a primera vista aleatoria.

Debido a esto, en tiempos más modernos (aproximadamente 1700) se optó por atacar un problema más débil: considerar la función $\pi(x)$ que cuenta la cantidad de números primos menores a $x \in \mathbb{R}$ y examinar el crecimiento de esta función. Gauss conjeturó que dicha función posee un comportamiento similar al de $\frac{x}{\log x}$ para x muy grande; y a finales del siglo XIX Jacques Hadamard y Vallé-Poussin demostraron este resultado de manera independiente y pasaría a llamarse el Teorema de los Números Primos (TNP) [12].

En una historia paralela, el Análisis Funcional surgió como necesidad de una nueva teoría para estudiar algunos objetos en Ecuaciones Diferenciales Parciales y Ecuaciones Integrales. Sin embargo, su aplicabilidad ha trascendido a otras ramas como la Teoría de Probabilidad, Álgebra Abstracta y Teoría de Números [8, Capítulo 1]. Aquí, presentamos una prueba del TNP utilizando elementos del análisis funcional.

Para esto, primero desarrollamos los conceptos necesarios de teoría de la medida, teoría de números, teoría de estimación y análisis de Fourier. Luego definimos una seminorma $||\cdot||_\Lambda$ en el álgebra de funciones $f\colon \mathbb{R} \to \mathbb{C}$ continuas de soporte compacto $C_c(\mathbb{R})$, con la convolución de funciones como producto. Utilizando esta norma podemos enunciar una proposición que implica el Teorema de los Números Primos; si $||\cdot||_\Lambda$ es trivial en $C_c(\mathbb{R})$ entonces TNP se sigue.

Suponer que $||\cdot||_\Lambda$ es no trivial, permite la construcción de un homomorfismo entre las funciones Lebesgue-integrables $L_1(\mathbb{R})$ y la completación del álgebra $C_c(\mathbb{R})$, lo cual eventualmente lleva a una contradicción.

1. Conceptos elementales de topología y teoría de la medida

1.1. Topología

En la recta $\mathbb{R}$ o el plano $\mathbb{R}^2$, la métrica euclideana usual induce de manera natural y formal una noción de cercanía. Gracias a esto, por ejemplo, asociamos el término «vecindad» de un punto $(x, y) \in \mathbb{R}^2$ con un círculo centrado en este punto. La topología nos permite hablar de vecindades y establecer una relación de cercanía entre puntos en un espacio sin apelar al concepto de métrica. Para hacerlo, utiliza únicamente la relación de contención entre conjuntos.

Sea X un conjunto. Una familia $\mathcal{T}$ de subconjuntos de X es una *topología* para X si $\mathcal{T}$ cumple las siguientes tres propiedades:

a) $\emptyset, X \in \mathcal{T}$,

b) Si $A \in \mathcal{T}$ y $B \in \mathcal{T}$ entonces $A \cap B \in \mathcal{T}$,

c) Si $A_i \in \mathcal{T}$ para todo $i \in I$, I un conjunto de índices, entonces $\bigcup_{i \in I} A_i \in \mathcal{T}$.

A la pareja $(X, \mathcal{T})$ se le conoce como *espacio topológico*. A los conjuntos que pertenecen a $\mathcal{T}$ se les llama *conjuntos abiertos* y a sus complementos *conjuntos cerrados*.

Nota 1.1.1. En la práctica, es común referirse a un espacio topológico $(X, \mathcal{T})$ simplemente como X; la topología en cuestión se aclara al inicio de la discusión o bien, los conceptos definidos y resultados obtenidos son independientes de la topología.

En $\mathbb{R}$, la topología natural donde las vecindades son intervalos abiertos da lugar a la formalización de dos conceptos importantes; la definición ε-δ de funciones

continuas y la definición ε-N de convergencia de sucesiones. En espacios topológicos, los conjuntos abiertos son el fundamento que generaliza estas ideas.

Ambos conceptos son muy importantes en $\mathbb{R}$; celebrados teoremas de existencia y de optimización, como el teorema del valor intermedio o el teorema de Weierstrass, dependen de la continuidad de una función o involucran convergencia de sucesiones. Más aún, en ambos es crucial que el dominio en el cual se está trabajando sea cierto tipo especial de conjunto llamado compacto.

1.2. Funciones continuas

Al haber dotado a un conjunto X de una topología le hemos brindado una estructura, similar a como en álgebra se dota a un conjunto de una operación especial entre elementos de este. Tener una estructura vuelve relevante considerar funciones que la preserven; en álgebra tales funciones son los homomorfismos, en topología las funciones continuas.

Una función $f\colon X \to Y$ es *continua* si para todo abierto B en Y se tiene que $f^{-1}(B)$ es abierto en X. Hay que resaltar que esta definición no sólo depende de la función sino de la topología escogida. Por ejemplo, en la topología trivial $\mathcal{T} = \{\emptyset, X\}$ solo las funciones constantes son continuas, mientras que en la topología discreta $\mathcal{T} = \mathcal{P}(X)$ toda función es continua. El último ejemplo podría sugerir que la topología discreta es siempre la «óptima» pues todas las funciones preservan estructura topológica. Sin embargo, esto resulta en que se pierdan ciertas nociones importantes, como los conjuntos compactos.

Entonces, ¿cómo comparamos dos topologías y qué hace a una más favorable que la otra? La primera pregunta se resuelve a través de la relación de contención. Sean $\mathcal{T}$ y $\mathcal{T}'$ dos topologías en X, decimos que $\mathcal{T}$ es *más débil* que $\mathcal{T}'$ si $\mathcal{T} \subset \mathcal{T}'$. En nuestro caso, la respuesta de la segunda está asociada a la de la primera y a los conjuntos compactos. Resulta que una topología más débil tiende a tener una mayor cantidad de conjuntos compactos (ver Nota 1.3.3)

Entonces, si tenemos un conjunto X y una colección de funciones $(f_i)_{i\in I}$ tal que para todo $i \in I$, $f_i\colon X \to Y_i$, donde Y_i tienen topologías $\mathcal{T}_i$, nos puede interesar dotar a X de un topología de tal manera que cada una de estas funciones sea continua. Más aún, por la observación dada anteriormente, lo más conveniente es considerar la topología más débil posible, a esta se le conoce como la *topología inicial* inducida

por las $(f_i)_{i\in I}$.

Un caso particular de la topología inicial es la topología producto. Sean $(Y_i, \mathcal{T}_i)$, $i \in \mathbb{N}$ una colección de espacios topológicos. Sea

$$X := \prod_{i\in\mathbb{N}} Y_i = \{(y_1, y_2, y_3, \cdots) \mid y_i \in Y_i\},$$

la *topología producto* en X es la topología inicial inducida por los mapeos de proyección

$$\pi_j \colon (y_i)_{i\in\mathbb{N}} \mapsto y_j$$

de X a Y_j para todo $j \in I$.

Más adelante mencionaremos otro ejemplo de la topología inicial, conocido como la topología débil*.

1.3. Conjuntos compactos

La base del análisis son las sucesiones y la noción de «límite». En $\mathbb{R}$ la convergencia de una sucesión se define a partir de las vecindades. De una manera análoga a la definición usual, decimos que una sucesión $(x_n)_{n\geq 1}$ de elementos en un espacio topológico X, *converge* a $x \in X$, escrito $x_n \to x$ si para cualquier abierto U tal que $x \in U$, existe $N \in \mathbb{N}$ tal que $x_n \in U$ para todo $n \geq N$. Se dice también que x es un *límite* de la sucesión.

Nota 1.3.1. Hay topologías donde los límites no son únicos; e.g. en la topología trivial, todos los $x \in X$ son límites de cualquier sucesión. Pero si la topología es Hausdorff, es decir, que para cada par de puntos $x, y \in X$ existen abiertos U y V tales que $x \in V$, $y \in V$ y $U \cap V = \emptyset$ (que es el caso de los espacios métricos), entonces el límite es único.

Esta perspectiva analítica se extiende a definir equivalentemente a un conjunto cerrado $C \subset X$, al menos en un espacio de Hausdorff (ver Nota 1.3.1) como un conjunto que posee los límites de todas las sucesiones convergentes $(x_n)_{n\geq 1}$ de elementos en C. Esto nos permite estudiar la estructura topológica desde un punto de vista análitico, por ejemplo:

Teorema 1.3.2. *Sea $f \colon X \to Y$ una función continua. Si $(b_n)_{n\geq 1}$ es una sucesión de elementos en X tal que $b_n \to b \in X$ entonces $f(b_n) \to f(b)$.*

Demostración. Sea U un abierto en Y que contiene a $f(b)$. Como f es continua, $f^{-1}(U)$ es abierto en X y contiene a b. Al ser b límite de la sucesión $(b_n)_{n\geq 1}$, existe un entero N a partir del cual $b_n \in f^{-1}(U)$ para todo $n \geq N$. Por lo tanto, $f(b_n) \in U$ para todo $n \geq N$. Como U era arbitrario, $f(b_n) \to f(b)$. □

En $\mathbb{R}$ los intervalos cerrados $[a, b]$ permiten un buen comportamiento de sucesiones y funciones definidas en ellos. Una de las razones principales es que este tipo de conjuntos son a la vez acotados (impide divergencia al infinito) y cerrados bajo límites (si una sucesión es convergente, su límite debe pertenecer al intervalo). Esto impide que la sucesión esté muy dispersa. En el peor de los casos, para una sucesión en un intervalo, tendríamos que al menos una subsucesión converge, esto es lo que da vida al concepto de conjunto compacto.

Decimos que un conjunto M es *compacto* si toda sucesión $(x_n)_{n\geq 1} \in M$ posee una subsucesión convergente $(x_{n_j})_{j\geq 1}$ cuyo límite está en M.

Nota 1.3.3. Sean dos topologías $\mathcal{T} \subset \mathcal{T}'$ definidas en un espacio X. Notemos que si p es un límite de alguna sucesión en X respecto a la topología $\mathcal{T}'$ entonces también es límite de esa sucesión respecto a $\mathcal{T}$. Por ello, si $C \subset X$ es un conjunto compacto respecto a la topología $\mathcal{T}'$ al tomar una subsucesión convergente $(x_{n_k})_{k\geq 1}$ de una sucesión $(x_n)_{n\geq 1}$ en C, esta subsucesión también será convergente en la topología $\mathcal{T}$. En otras palabras, si C es compacto en $\mathcal{T}'$, también lo es en $\mathcal{T}$.

La compacidad permite resultados muy interesantes; converso a la discusión realizada anteriormente por ejemplo,

Teorema 1.3.4. *Todo conjunto compacto es cerrado y acotado.*

Demostración. Para una demostración, ver lema 2.5-2 de [13]. □

Otro gran ejemplo que se discutió en la sección de topología es el teorema de Weierstrass.

Teorema 1.3.5 (Weierstrass)**.** *Toda función continua de valor real asume, en un conjunto compacto, un máximo y un mínimo absolutos.*

Demostración. Sea $f\colon X \to \mathbb{R}$ una función continua, $M \subset X$ compacto y $L = \sup_{m\in M} f(m)$. Construimos una sucesión $y_1 < y_2 < y_3 < \cdots$ tal que $y_n \to L$, donde $y_n \in f(M)$ para todo $n \in \mathbb{N}$. Así que existe una sucesión (x_n) de elementos en M

tales que $f(x_n) = y_n$ y al ser M un conjunto compacto, esta sucesión posee una subsucesión $(x_{n_j}) \to x \in M$. Finalmente, por el Teorema 1.3.2 la continuidad de f nos asegura que

$$f(x) = f\big(\lim_{j\to\infty} x_{n_j}\big) = \lim_{j\to\infty} f(x_{n_j}) = L < +\infty$$

por lo que f alcanza un máximo en M. De forma análoga, se demuestra que f alcanza un mínimo en M. □

Un concepto importante relacionado con límites de sucesiones es la *densidad*. Así como todo número $\xi \in \mathbb{R}$ puede aproximarse por medio de una sucesión de elementos $(q_n)_{n\geq 1}$ en $\mathbb{Q}$, o formalmente $q_n \to \xi$, existen contextos más generales en los que un conjunto puede ser descrito como límites de sucesiones de otro conjunto. Decimos que B es *denso* en X si para todo $x \in X$ existe una sucesión de elementos en B, $(b_n)_{n\geq 1}$ tal que $b_n \to x$.

La principal razón por la que la densidad es relevante es que para demostrar un resultado que concierne a un conjunto X, en varias ocasiones basta con probar que dicho resultado se cumple en un conjunto B denso en X y «llevar el argumento al límite». Usualmente para realizar formalmente este último paso se utiliza el Teorema 1.3.2 o bien, directamente el siguiente corolario:

Corolario 1.3.6. *Sea $f\colon X \to Y$ continua y B denso en X, entonces $f(B)$ es denso en $f(X)$.*

Demostración. Todo elemento en $f(X)$ puede escribirse como $f(x)$ para algún $x \in X$. Como B es denso en X, existe una sucesión $(b_n)_{n\geq 1}$ de elementos de B tal que $b_n \to x$. Luego, por la continuidad de f y el Teorema 1.3.2, $f(b_n) \to f(x)$. □

Existe también una definición conjuntista de conjuntos compactos, la cual es equivalente a la mencionada anteriormente en espacios métricos.

Teorema 1.3.7. *Un conjunto compacto $M \subset X$ cumple que para cualquier colección de abiertos U tal que $M \subset \bigcup U$, existe una subcolección finita $U_1, \cdots, U_N$ tal que $M \subset \bigcup_{n=1}^{N} U_n$.*

Demostración. Para una demostración, considerar la definición de conjunto compacto, el Teorema 7.19 y el Corolario 7.20 de [9]. □

Nota 1.3.8. En realidad, la condición en el Teorema 1.3.7 es la definición común de conjunto compacto. La definición dada por nosotros recibe en general el nombre de *compacidad por secuencias*.

Finalmente, presentamos un ejemplo importante de la topología inicial, en la que se «preserva» la calidad de compacidad.

Teorema 1.3.9 (Tychonoff)**.** *Sean X_i con $i \in I$, espacios topológicos compactos. Entonces $\prod_{i\in I} X_i$ es compacto respecto a la topología producto.*

Demostración. Para una demostración, ver el Teorema 7.25 de [9]. □

1.4. Conjuntos y funciones medibles

La integral de Lebesgue busca generalizar la integral de Riemann. La diferencia clave en la definición es la forma en la que se particiona el dominio respecto del cual se integra; en Riemann se toman uniones de intervalos, mientras que en Lebesgue se toman las preimágenes del integrando.

Para tal efecto, se necesita primero establecer un concepto de «tamaño» o «medida», pues las preimágenes de una función pueden variar drásticamente. Dado un conjunto X, resulta ser imposible asignar una «medida» a todos los subconjuntos de X, de una manera coherente con la noción intuitiva de «medida». Pensar, por ejemplo en la «paradoja» de Banach-Tarski. Por ello, pediremos condiciones mínimas sobre una colección de conjuntos «medibles».

Sea X un conjunto. Una colección de conjuntos $\mathcal{X} \subset \mathcal{P}(X)$ es llamada una *σ-álgebra* en X si cumple:

a) $\emptyset \in \mathcal{X}$.

b) Es cerrada bajo complementos; si $A \in \mathcal{X}$, entonces $A^c \in \mathcal{X}$.

c) Es cerrada bajo uniones contables; si $A_n \in \mathcal{X}$ con $n \in \mathbb{N}$, entonces $\bigcup_{n=1}^{\infty} A_n \in \mathcal{X}$.

A sus elementos se les llama *conjuntos medibles* y a la pareja $(X, \mathcal{X})$ se le conoce como *espacio medible*.

Dependiendo de X, pueden existir muchas familias de subconjuntos que cumplen con las condiciones dadas, por ejemplo el conjunto potencia $\mathcal{P}(X)$ claramente lo hace. Sin embargo, sí existen σ-álgebras distintas de $\mathcal{P}(X)$ y más aún, es posible "completar" una colección arbitraria de subconjuntos y convertirla en una σ-álgebra. Obviamente, la completación más interesante es la que requiera agregar menos conjuntos. De la definición, no es dificil ver que la intersección arbitraria de σ-álgebras es también una σ-álgebra. Utilizando esta propiedad, conseguimos minimalidad considerando lo siguiente: dada una colección $\mathcal{C} \subset \mathcal{P}(X)$, llamamos a la σ-álgebra *generada por* $\mathcal{C}$ a la intersección de todas las σ-álgebras que contengan a $\mathcal{C}$,

$$\sigma(C) := \bigcap_{\substack{\mathcal{X} \ \sigma\text{-álgebra} \\ \mathcal{X} \supset \mathcal{C}}} \mathcal{X}.$$

En particular, para un espacio donde esté definida una topología nos interesará la *σ-álgebra de Borel* $\mathcal{B}$, cuyos elementos se denominan *conjuntos Borel medibles* y es la generada por la topología. La razón por la cual esta σ-álgebra es importante es porque conecta la estructura topológica del espacio junto con el concepto de medida. En el sentido Riemann, toda función continua en un intervalo es integrable en ese intervalo, así que en el sentido Lebesgue debería serlo también y provisto que las preimágenes de una función son las que dictan la partición del dominio sobre el cual se integra, como mínimo los conjuntos abiertos y cerrados deben ser medibles. En el Teorema 1.5.1 a continuación, vemos que algunos conjuntos en $\mathcal{B}$ no son ni abiertos ni cerrados. A pesar de esto, podemos interpretar intuitivamente a $\mathcal{B}$ como una colección de conjuntos razonables desde el punto de vista topológico.

De manera análoga a cómo una función continua preserva una estructura topológica, introducimos el concepto de función medible, una función que preserve la estructura de la σ-álgebra. Sean $(X, \mathcal{X})$ y $(\mathbb{C}, \mathcal{Y})$ espacios medibles, se dice que una función $f\colon X \mapsto \mathbb{C}$ es *medible* si para todo $M \in \mathcal{Y}$, $f^{-1}(M) \in \mathcal{X}$. Al conjunto de funciones medibles de un espacio medible $(X, \mathcal{X})$ se le denota como $M(X, \mathcal{X})$ y como $M^+(X, \mathcal{X})$ al conjunto de funciones medibles que toman valores reales no negativos. Si $\mathcal{X} = \mathcal{B}$, decimos que f es *Borel medible* y en efecto, la definición de $\mathcal{B}$ implica que funciones continuas, relevantes para la topología, sean también Borel medibles.

1.5. Medida

Ya establecimos los requisitos mínimos que debe cumplir una familia $\mathcal{X}$ de conjuntos medibles: ser una σ-álgebra. Ahora, definimos *medida* como una función $\mu\colon \mathcal{X} \to [0,\infty]$ que cumple:

1. $\mu(\emptyset) = 0$,
2. $\mu(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu(E_n)$, si $E_i \cap E_j = \emptyset$ para $i \neq j$. (σ-finitud)

y a la terna $(X, \mathcal{X}, \mu)$ como un *espacio de medida*. El ejemplo de medida más relevante para nuestra discusión será la *medida de Lebesgue* μ_ℓ. Esta es la única medida definida en $(\mathbb{R}, \mathcal{B})$ que coincide con la longitud de los intervalos abiertos; $\mu_\ell((a,b)) = b - a$ con $a, b \in \mathbb{R}$ (de hecho coincide con cualquier tipo de intervalo). Para detalles en su existencia y construcción referimos al lector a [3]. A toda medida que esté definida en $\mathcal{B}$ se le llama una *medida de Borel*.

Para ilustrar cómo el concepto de medida puede abarcar conjuntos «más arbitrarios» que los intervalos, veamos lo siguiente

Teorema 1.5.1. Los números racionales son Borel medibles y tienen medida de Lebesgue 0.

Demostración. Es sencillo demostrar que todo intervalo de la forma $[a,b]$ con $a, b \in \mathbb{R}$ es Borel-medible. En particular, si $a = b = q \in \mathbb{Q}$, $[a,b] = \{q\}$ es medible y puesto que los racionales son contables $\bigcup_{q\in\mathbb{Q}}\{q\} = \mathbb{Q} \in \mathcal{B}$. Usando el axioma de cerradura bajo uniones contables, se demuestra que $\mu_\ell([q,q]) = 0$. Entonces

$$\mu_\ell(\mathbb{Q}) = \sum_{q\in\mathbb{Q}} \mu_\ell(\{q\}) = 0. \qquad \square$$

A pesar de que el conjunto de racionales sea infinito y no acotado, la medida de Lebesgue logra capturar la noción de que "ocupan poco espacio" en $\mathbb{R}$ (al menos en la intuición de medida). De hecho, para una medida arbitraria μ, si alguna afirmación se cumple en todo el espacio, excepto en un conjunto E con $\mu(E) = 0$, se dice que la afirmación se cumple *μ-casi en todas partes*. Si en dado caso la medida μ se entiende del contexto simplemente utilizamos el término *casi en todas partes*.

Por otro lado, es sencillo ver que la medida de Lebesgue de todo conjunto acotado debe ser finita, lo cual es razonable en este caso. Sin embargo, existen otras

medidas como la inducida por la cardinalidad, que no cumplen esta propiedad. En búsqueda de mantener la conexión entre medida, topología y esta intuición, si una medida cumple que todo conjunto compacto tiene medida finita entonces se dice que es *localmente finita*. Si una medida localmente finita además es de Borel, se dice que es una medida de *Radon*. Hablaremos de la medida de Radon en mayor detalle en el Capítulo 5.

Finalmente, en un espacio de medida $(X, \mathcal{X}, \mu)$, si X es una unión de conjuntos de medida μ finita, entonces se dice que μ es σ-*finita*. La medida de Lebesgue μ_ℓ es σ-finita y esto nos permitirá usar dos teoremas importantes de la teoría de integración, mencionados más adelante.

1.6. Integración

Sea $(X, \mathcal{X}, \mu)$ un espacio de medida y $E \in \mathcal{X}$. Consideremos la función característica de E, $\chi_E \colon X \to \mathbb{R}$, definida por

$$\chi_E(x) := \begin{cases} 1, & x \in E, \\ 0, & x \notin E. \end{cases}$$

Este es el caso más básico de una función medible, y al ser nula en todas partes excepto en E es natural definir

$$\int \chi_E \, d\mu := \mu(E)$$

y así mismo, extender esta definición a una función de la forma $\varphi := \sum_{n=1}^{N} a_n \chi_{E_n}$, donde $a_n \in \mathbb{R}^+$ para todo n y $E_i \cap E_j = \emptyset$ si $i \neq j$, como

$$\int \varphi \, d\mu = \sum_{n=1}^{N} a_n \mu(E).$$

A estas funciones que obtienen una cantidad finita de valores distintos en su rango se les conoce como *funciones simples*.

De esta manera, podemos hacer un razonamiento análogo al que se utiliza para definir la integral de Riemann: aproximar el área bajo la curva por rectángulos acotados por la misma. Motivados por esto, se define para una función $f \colon X \to \mathbb{R}_{\geq 0}$

medible,

$$\int f\, d\mu := \sup\left\{ \int \varphi\, d\mu \;\middle|\; \varphi \in M^+(X, \mathcal{X}) \text{ simple y tal que } f \geq \varphi \right\}.$$

Nota 1.6.1. Para dos funciones f y g con el mismo dominio, $f \geq g$ denota $f(x) \geq g(x)$ para todo x en el dominio.

Esta definición se extiende naturalmente para una función $f\colon X \to \mathbb{R}$ separándola en sus partes no negativa y negativa $f = f^+ - f^-$,

$$\int f\, d\mu := \int f^+\, d\mu - \int f^-\, d\mu.$$

Finalmente, para una función $f\colon X \to \mathbb{C}$,

$$\int f\, d\mu := \int \mathrm{Re}(f)\, d\mu + i \int \mathrm{Im}(f)\, d\mu.$$

En general, las funciones cuyas integrales son finitas son las más interesantes. Formalmente, si la integral de $|f|$ es finita, decimos que f es *integrable*. Es sencillo demostrar que el conjunto de funciones integrables es un espacio vectorial sobre los complejos. Más aún, con la intención de dotar de una norma a este conjunto, se define

$$||f||_1 := \int |f|\, d\mu.$$

Sin embargo, esta no es una norma, sino una seminorma. Para obtener un espacio normado íntimamente relacionado con estas funciones, se considera la relación de equivalencia $\sim$. Dos funciones f y g están relacionadas, escrito $f \sim g$, si y solo si $f = g$ μ-casi en todas partes.

Al espacio cociente generado por esta relación de equivalencia, se le conoce como $L_1(X)$ y es un espacio normado con la norma definida anteriormente.

Nota 1.6.2. A pesar de que se trabaja con el conjunto de clases de equivalencia, es común tratarlas como funciones, pues en la mayoría de situaciones el comportamiento del espacio cociente es análogo al del espacio original.

Aún más interesante, $L_1(X)$ es *completo*, es decir, toda sucesión de Cauchy en la métrica inducida por la norma descrita anteriormente es convergente. También existe una relación útil entre convergencia en el sentido $L_1(X)$ y el sentido puntual,

Teorema 1.6.3. *Si una sucesión de funciones medibles (f_n) converge a f en $L_1(X)$, es decir $||f_n - f||_1 \to 0$, entonces existe una subsucesión (f_{n_j}) que converge puntualmente casi en todas partes a f.*

Demostración. Para una demostración, ver el Teorema 7.23 de [2]. □

Como se mencionó anteriormente, esta forma de integrar surgió como una búsqueda por generalizar la integral de Riemann. En particular si $\mu = \mu_\ell$, la medida de Lebesgue, tenemos que para toda función f Riemann-integrable, f es Lebesgue medible y

$$\int_a^b f(x)dx = \int_{[a,b]} f \, d\mu_\ell. \tag{1.1}$$

Esto surge como consecuencia de, entre otros resultados, el siguiente teorema recurrente en la teoría de integración.

Teorema 1.6.4 (Teorema de Convergencia Dominada)**.** *Sea $(X, \mathcal{X}, \mu)$ un espacio de medida y (f_n) una sucesión de funciones complejas medibles con $f(x) = \lim_{n\to\infty} f_n(x)$ para todo $x \in X$. Suponiendo que existe una función integrable $g\colon X \to \mathbb{R}_{\geq 0}$ con $|f_n| \leq g$ para $n \in \mathbb{N}$. Entonces f es integrable y*

$$\int f \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu.$$

Demostración. Para una demostración, ver Teorema 3.26 de [2]. □

De igual manera que en la integral de Riemann, es posible considerar integrales de funciones de varias variables. El teorema más importante en este ámbito es el de Fubini:

Teorema 1.6.5 (Fubini)**.** *Para espacios de medida σ-finita $(X, \mathcal{X}, \mu)$ y $(Y, \mathcal{Y}, \nu)$ existe una única medida σ-finita $\mu \times \nu$ en $X \times Y$ tal que*

$$(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$$

para todos $A \in \mathcal{X}$ y $B \in \mathcal{Y}$. Si además, f es una función medible en $X \times Y$ y

$$\int_{X\times Y} |f(x, y)|d(\mu \times \nu) < \infty$$

entonces para casi toda $y \in Y$ y $x \in X$ se sigue

$$\int_{X\times Y} |f(x,y)|d(\mu\times\nu) = \int_X \left(\int_Y f(x,y)d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X f(x,y)d\mu(x) \right) d\nu(y).$$

Demostración. Para una demostración, ver Teorema 5.32 de [2]. □

1.7. Derivada de Radon-Nikodym

Una propiedad interesante de las integrales es que podemos construir medidas a partir de ellas. Si $(X, \mathcal{X})$ es un espacio medible, f es una función medible y μ una medida para este espacio, entonces

$$\nu(E) := \int_E f \, d\mu \tag{1.2}$$

también define una medida para el mismo. Esto nos hace considerar el converso, ¿será posible para cada par de medidas, hallar una función f de tal manera que una se pueda escribir como la integral de la otra, como en (1.2)?

Podemos observar en la definición de esta medida ν, que si $\mu(E) = 0$ entonces $\nu(E) = 0$. Dadas dos medidas ν y μ, decimos que ν es *absolutamente continua* respecto a μ, escrito $\nu \ll \mu$, si $\mu(E) = 0$ implica $\nu(E) = 0$ para todo E medible. Resulta que esto en conjunto con σ-finitud es suficiente para garantizar una representación de la forma (1.2).

Teorema 1.7.1 (Radon-Nikodym)**.** *Sean ν y μ medidas σ-finitas definidas en $\mathcal{X}$ tales que $\nu \ll \mu$. Entonces existe una función $f \in M^+(X, \mathcal{X})$ tal que*

$$\nu(E) = \int_E f \, d\mu, \quad E \in \mathcal{X}.$$

Más aún, la función f, llamada la derivada de Radon-Nikodym de ν respecto a μ y denotada por $\frac{d\nu}{d\mu}$, está determinada de manera única μ-casi en todas partes.

Demostración. Para una demostración, ver Teorema 8.9 de [3]. □

2. Tópicos de teoría de números y análisis

2.1. Propiedades básicas de números primos

Un entero positivo $n > 1$ que posee como divisores positivos únicamente a 1 y a él mismo es llamado *número primo* (o simplemente *primo*). Un entero positivo mayor que 1 que no es primo se dice que es *compuesto.*

Como se mencionó inicialmente, existen resultados interesantes concernientes al estudio de los números primos que se conocen desde la antigüedad. Naturalmente, es importante para la discusión que

Teorema 2.1.1 (Euclides)**.** *Existen infinitos números primos.*

Demostración. Sea $\{p_1, p_2, \cdots, p_n\}$ un conjunto de primos y

$$N = p_1 p_2 \cdots p_n + 1.$$

Notemos que N no es divisible por ninguno de los primos en el conjunto inicial, entonces existe un primo p_{n+1} que no está en lista. Como n es arbitrario, podemos repetir el procedimiento agregando p_{n+1} a la lista para obtener infinitos primos. □

Dado un número n, es posible determinar si este es primo diviéndolo por primos menores que él. Es posible acotar la cantidad de operaciones que hay que realizar.

Teorema 2.1.2. *Si n es un entero positivo tal que para todo primo $p \leq \sqrt{n}$ se cumple que p no es divisor de n, entonces n es primo.*

Demostración. Supongamos lo contrario, entonces existe un número compuesto n tal que su menor divisor primo p es mayor que $\sqrt{n}$. Entonces existe un entero $k > p$ tal que

$$n = kp \geq p^2 > n,$$

una contradicción. □

Para el estudio de la distribución de primos se introdujo la función $\pi(x)$, que denota la cantidad de primos p que sean menores o iguales a x para un número real positivo dado x:

$$\pi(x) = \sum_{p \leq x} 1, \qquad \text{donde } p \text{ denota un número primo.}$$

Trabajar con el conjunto de primos es complicado debido a su distribución. Como el resultado que nos concierne es asintótico, es comúnmente provechoso considerar un conjunto que sea más sencillo de manejar, aunque sea más grande, pero del cual aún podamos obtener información útil para el problema original. Basado en esta idea, introducimos la función Λ de *von Mangoldt*:

$$\Lambda(n) = \begin{cases} \log p, & \text{si } n = p^k \text{ para un primo } p \text{ y algún entero positivo } k. \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Como se puede ver, Λ es una función de conteo con peso sobre las potencias de primos. Por otro lado, si se tiene en mente la conjetura de Gauss y el hecho de que el logaritmo separa estructuras multiplicativas en sumas, es perfectamente razonable que esta función auxiliar involucre al logaritmo. Ahora análogo a π, consideramos la suma

$$\sum_{1 \leq n \leq x} \Lambda(n)$$

y al mismo tiempo, puede resultar interesante considerar la subsuma

$$\sum_{d | \lfloor x \rfloor} \Lambda(d).$$

En efecto, es fácil ver que la suma de arriba es igual a $\log \lfloor x \rfloor$, esto nos sugiere utilizar las herramientas teórico-numéricas presentadas a continuación.

2.2. Inversión de Möbius

Dadas dos funciones $f_1, f_2 \colon \mathbb{N} \to \mathbb{R}$ se define la *convolución de Dirichlet* como

$$(f_1 \underset{D}{*} f_2)(n) = \sum_{d|n} f_1(d) f_2\left(\frac{n}{d}\right).$$

Notemos que si d es un divisor de n, entonces $\frac{n}{d}$ también lo es y $n/(\frac{n}{d}) = d$, que implica

$$(f_1 \underset{D}{*} f_2)(n) = \sum_{d|n} f_1(d) f_2\left(\frac{n}{d}\right) = \sum_{d|n} f_1\left(\frac{n}{d}\right) f_2(d) = (f_2 \underset{D}{*} f_1)(n),$$

así que la convolución de Dirichlet es conmutativa. También, si $d_1|d$ y $d|n$ entonces $d_1|n$ y $n = d_1 \times d/d_1 \times n/d$. De esto,

$$\begin{aligned}
(f_1 \underset{D}{*} f_2) \underset{D}{*} f_3(n) &= \sum_{d|n} \left[\sum_{d_1|d} f_1(d_1) f_2\left(\frac{d}{d_1}\right) \right] f_3\left(\frac{n}{d}\right) \\
&= \sum_{d|n} f_1(d) \left[\sum_{d_1|d} f_2(d_1) f_3\left(\frac{d}{d_1}\right) \right] \\
&= f_1 \underset{D}{*} (f_2 \underset{D}{*} f_3)(n),
\end{aligned}$$

la convolución de Dirichlet es entonces asociativa.

Una propiedad interesante del caso $f \underset{D}{*} \mathbb{1}$, donde $\mathbb{1}$ representa la función constante 1, es que el mapeo $f \longmapsto f \underset{D}{*} \mathbb{1}$ es invertible. Para hallar el mapeo inverso, definimos la *función de Möbius* $\mu \colon \mathbb{N} \to \mathbb{Z}$ como

$$\mu(n) = \begin{cases} 0, & \text{si } p^2|n \text{ para algún primo } p \\ (-1)^k, & \text{si } n = p_1 p_2 \cdots p_k \end{cases}$$

donde $p_1, p_2, \cdots, p_k$ denotan distintos primos.

Teorema 2.2.1 (Inversión de Möbius)**.** *Si* $f \colon \mathbb{N} \to \mathbb{R}$ *entonces* $(f \underset{D}{*} \mathbb{1}) \underset{D}{*} \mu = f$.

Demostración. Por asociatividad, $(f \underset{D}{*} \mathbb{1}) \underset{D}{*} \mu = f \underset{D}{*} (\mathbb{1} \underset{D}{*} \mu)$ y por conmutatividad

$$(\mathbb{1} \underset{D}{*} \mu)(n) = \sum_{d|n} \mu(d).$$

Si $n = 1$, claramente el resultado de la suma de arriba es 1. De otra forma, sea $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$, entonces todo divisor d se anula bajo μ a menos que sea el producto de $0 \leq r \leq k$ primos distintos del conjunto $\{p_1, p_2, \cdots, p_k\}$. Entonces por la definición de μ se tiene

$$\sum_{d|n} \mu(d) = \sum_{r=0}^{k} \binom{k}{r} (-1)^r = (1-1)^k = 0.$$

De esto, $f \underset{D}{*} (\mathbb{1} \underset{D}{*} \mu)(n) = f(n)$ para toda n. □

No es dificil ver que $(\Lambda \underset{D}{*} \mathbb{1})(n) = \sum_{d|n} \Lambda(d) = \log n$ así que el teorema anterior nos permite reescribir $\Lambda = (\Lambda(n) \underset{D}{*} \mathbb{1}) \underset{D}{*} \mu = \mu \underset{D}{*} \log$. La función Λ, tiene una sencilla representación en funciones más simples que interactúan por medio de la convolución de Dirichlet, esto motiva a explorar el comportamiento entre otras funciones similares con esta misma operación. Con algo de suerte y visión se halló el siguiente resultado

$$\Lambda_2(n) := (\mu \underset{D}{*} \log^2)(n) \equiv (\Lambda \underset{D}{*} \Lambda)(n) + \Lambda(n) \log n. \tag{2.1}$$

Demostración.

$$\begin{aligned}
\log^2 n &= \log n \sum_{d|n} \Lambda(d) \\
&= \sum_{d|n} \Lambda(d) \log\left(\frac{n}{d} d\right) \\
&= \sum_{d|n} \Lambda(d) \log \frac{n}{d} + \sum_{d|n} \Lambda(d) \log d \\
&= (\Lambda \underset{D}{*} \log)(n) + ([\Lambda \log] \underset{D}{*} \mathbb{1})(n)
\end{aligned}$$

Entonces, por la distributividad de $\underset{D}{*}$ respecto a la suma

$$(\mu \underset{D}{*} \log^2)(n) = \mu \underset{D}{*} (\Lambda \underset{D}{*} \log)(n) + \mu \underset{D}{*} ([\Lambda \log] \underset{D}{*} \mathbb{1})(n)$$

pero

$$\mu \underset{D}{*} (\Lambda \underset{D}{*} \log)(n) = \Lambda \underset{D}{*} (\mu \underset{D}{*} \log)(n) = (\Lambda \underset{D}{*} \Lambda)(n)$$

y por inversión de Möbius,

$$\mu \underset{D}{*} ([\Lambda \log] \underset{D}{*} \mathbb{1})(n) = \Lambda(n) \log(n).$$ □

Con esto cubierto, tomamos un pequeño desvío hacía el análisis real. Veremos definiciones que nos servirán para enunciar el teorema de los números primos de manera compacta pero formal. Así como también ciertas técnicas y resultados que nos serán útiles para su deducción.

2.3. Notación de Landau

Es común que en matemática de pregrado exista una inclinación hacia resultados en los cuales estén involucrados fórmulas exactas o identidades, como por ejemplo, en la resolución de ecuaciones diferenciales y de recurrencia por medio de métodos que proveen explícitamente la solución. Sin embargo, fuera del libro de texto, las situaciones en las cuales el objeto de estudio se comporta de esta manera no son la mayoría, por lo cual en ciertos casos se opta por derivar *estimaciones asintóticas* para las funciones o cantidades en cuestión. Es decir, aproximaciones de estas por medio de funciones «más simples», que permiten describir de cierta manera su comportamiento en un intervalo del dominio.

La notación asintótica más común es la de *O-grande*. Dadas dos funciones f y g de variable real o entera, la notación

$$f(x) = O(g(x)),$$

expresa que $g(x)$ acota asíntóticamente a $f(x)$. Es decir, existen constantes x_0 y c tales que

$$|f(x)| \leq c|g(x)| \quad (x \geq x_0).$$

Se dice que $f(x)$ es *de orden* $O(g(x))$. Por ejemplo $x = O(e^x)$ ya que $e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots$. Otro ejemplo, que será relevante en la siguiente sección es

$$\log^2(1+x) = O(\sqrt{x}) \tag{2.2}$$

que viene de

$$\log^2(1+x) \leq 9\sqrt{x}$$

y esto a su vez de

$$1 + x \leq e^{3\sqrt[4]{x}} = 1 + 3\sqrt[4]{x} + \frac{9\sqrt[2]{x}}{2} + \frac{27\sqrt[4]{x^3}}{6} + \frac{81x}{24} + \cdots$$

Una notación más fuerte es la de *o-pequeña*. Dadas dos funciones f y g de variable real o entera, la notación

$$f(x) = o(g(x)) \quad (x \to \infty),$$

significa que $g(x) \neq 0$ para x suficientemente grande y

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$$

o equivalentemente,

$$|f(x)| \leq c|g(x)| \quad (x \geq x_0),$$

para todo $c > 0$ arbitrariamente pequeña y x_0 dependiente de c. Se dice que $f(x)$ es de *menor orden que* $g(x)$. Por ejemplo $\log x = o(x)$ ya que por L'Hopital,

$$\lim_{x\to\infty} \frac{\log x}{x} = \frac{\frac{1}{x}}{1} = 0.$$

Relacionada íntimamente con la anterior, tenemos la equivalencia asintótica. Dos funciones son *asintóticamente equivalentes*, escrito $f \sim g$, si

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1.$$

Esto sucede sí y solo sí $f(x)$ puede escribirse como $g(x)$ más una función de orden menor que $g(x)$, simbólicamente representado como $f(x) = g(x) + o(g(x))$. [11]

Teorema 2.3.1 (Teorema de los Números Primos)**.** *Las funciones* $\pi(x)$ *y* $\frac{x}{\log x}$ *son asintóticamente equivalentes. Es decir,*

$$\pi(x) \sim \frac{x}{\log x}.$$

2.4. Estimaciones importantes

Ya presentadas las notaciones de O-grande y o-pequeña, expondremos algunas estimaciones que nos permitirán derivar la fórmula de simetría de Selberg.

Teorema 2.4.1. *Se tiene que*

$$\frac{1}{y}\sum_{m\leq y}\log^2 m = \sum_{m\leq y}\frac{2\log(\frac{y}{m}) - c_3 + c_4\frac{m}{y}}{m} + O\bigg(\frac{\log^2(1+y)}{y}\bigg).$$

Demostración. Comencemos con la función $f(t) = \frac{1}{\lfloor t\rfloor}$ y la siguiente observación,

$$\int_1^{\lfloor y\rfloor+1} f(t)dt = \sum_{m=1}^{\lfloor y\rfloor}\bigg(\int_m^{m+1}\frac{1}{m}dt\bigg) = \sum_{m\leq y}\frac{1}{m}.$$

Usando esto,

$$\begin{aligned}\sum_{m\leq y}\frac{1}{m} &= \int_1^{\lfloor y\rfloor+1} f(t)dt + \log y - \int_1^y \frac{dt}{t}\\ &= \log y + \int_1^y f(t) - \frac{1}{t}\,dt + \int_y^{\lfloor y\rfloor+1} f(t)dt\\ &= \log y + \int_1^\infty f(t) - \frac{1}{t}\,dt - \int_y^\infty f(t) - \frac{1}{t}\,dt + \int_y^{\lfloor y\rfloor+1} f(t)dt\end{aligned}$$

pero

$$f(t) - \frac{1}{t} = \frac{t - \lfloor t\rfloor}{t^2 - t(t-\lfloor t\rfloor)} \leq \frac{1}{t^2-t} \leq \frac{2}{t^2} \quad \text{para } t\geq 2,$$

así que la segunda integral converge, la tercera es de orden $\frac{1}{y}$ y entonces

$$\sum_{m\leq y}\frac{1}{m} = \log y + c_1 + O\bigg(\frac{1}{y}\bigg). \tag{2.3}$$

Utilizando argumentos similares en la función $f(t) = \frac{\log\lfloor t\rfloor}{\lfloor t\rfloor}$ llegamos a

$$\sum_{m\leq y}\frac{\log m}{m} = \frac{\log^2 y}{2} - c_2 + O\bigg(\frac{\log(1+y)}{y}\bigg),$$

que en conjunto con el resultado anterior nos lleva a

$$\sum_{m\leq y}\frac{\log(y/m)}{m} = \frac{\log^2 y}{2} + c_1\log y + c_2 + O\bigg(\frac{\log(1+y)}{y}\bigg). \tag{2.4}$$

Por otro lado, si $f\colon [1,\infty)\to\mathbb{R}$ es una función no decreciente deducimos:

a) $\int_1^y f(t)dt = \int_1^{\lfloor y \rfloor} f(t)dt + \int_{\lfloor y \rfloor}^y f(t)dt \leq \int_1^{\lfloor y \rfloor} f(t)dt + f(y)$.

b) $\int_{\lfloor y \rfloor}^y f(t)dt = \sum_{m=2}^{\lfloor y \rfloor} \int_{m-1}^m f(t)dt \leq \sum_{m=2}^{\lfloor y \rfloor} f(m)$.

c) Si $g(t) := f(\lfloor t \rfloor)$, la monotonía de f nos garantiza $g(t) \leq f(t)$ para todo t y por lo tanto

$$\int_2^{\lfloor y \rfloor + 1} g(t)dt = \sum_{m=2}^{\lfloor y \rfloor} \int_m^{m+1} g(t)dt = \sum_{m=2}^{\lfloor y \rfloor} f(m) \leq \int_2^{\lfloor y \rfloor + 1} f(t)dt.$$

d) $\int_2^{\lfloor y \rfloor + 1} f(t)dt = \int_2^y f(t)dt + \int_y^{\lfloor y \rfloor + 1} f(t)dt \leq \int_2^y f(t)dt + f(y+1)$.

De (a), (b),(c) y (d) obtenemos la desigualdad

$$\left| \sum_{m=1}^{\lfloor y \rfloor} f(m) - \int_1^y f(t)dt \right| \leq f(1) + f(y+1) \tag{2.5}$$

y en particular tomando $f(t) = \log^2 t$, haciendo uso de $\log^2 t = \frac{d}{dt}[t \log^2 t - 2t \log t + 2t]$, conseguimos la estimación:

$$\frac{1}{y} \sum_{m \leq y} \log^2 m = \log^2 y - 2 \log y + 2 + O\left(\frac{\log^2(1+y)}{y} \right). \tag{2.6}$$

Notemos que si multiplicamos (2.3) por $-2 - c_1$ y (2.4) por 2, al sumar se obtiene

$$\sum_{m \leq y} \frac{2 \log(\frac{y}{m}) - 2 - 2c_1}{m} = \log^2 y - 2 \log y + 2c_2 - (2 + 2c_1)c_1 + O\left(\frac{\log(1+y)}{y} \right)$$

y de la siguiente observación,

$$\sum_{m \leq y} \frac{\frac{m}{y}}{m} = \frac{\lfloor y \rfloor}{y} = 1 + O\left(\frac{1}{y} \right)$$

haciendo $c_3 = 2 + 2c_1$ y $c_4 = 2 - 2c_2 + c_3 c_1$,

$$\sum_{m \leq y} \frac{2 \log(\frac{y}{m}) - c_3 + c_4 \frac{m}{y}}{m} = \log^2 y - 2 \log y + 2 + O\left(\frac{\log(1+y)}{y} \right).$$

De lo anterior, utilizando (2.6) llegamos a

$$\frac{1}{y}\sum_{m\leq y}\log^2 m = \sum_{m\leq y}\frac{2\log(\frac{y}{m}) - c_3 + c_4\frac{m}{y}}{m} + O\bigg(\frac{\log^2(1+y)}{y}\bigg). \qquad \square$$

2.5. Fórmula de simetría de Selberg

En su demostración elemental, la visión más crucial de Selberg era construir una fórmula asintótica que relacionara $\sum_{p\leq x}\log p$ a una función que involucrara únicamente a $\log x$. De esto nació su famosa Fórmula de Simetría, que podemos reescribir gracias a la Ecuación (2.1) como

$$\sum_{n\leq x}\Lambda_2(n) = 2x\log x + O(x) \tag{2.7}$$

Demostración. De la ecuación 2.1

$$\sum_{n\leq x}\Lambda_2(n) = \sum_{n\leq x}\sum_{d|n}\mu(d)\log^2\frac{n}{d},$$

intercambiamos los órdenes de suma

$$\sum_{n\leq x}\Lambda_2(n) = \sum_{d\leq x}\mu(d)\sum_{\substack{n\leq x\\ d|n}}\log^2\frac{n}{d}$$

escribimos $m = n/d$,

$$\sum_{\substack{n\leq x\\ d|n}}\log^2\frac{n}{d} = \sum_{m\leq\frac{x}{d}}\log^2 m$$

y de esto

$$\frac{1}{x}\sum_{n\leq x}\Lambda_2(n) = \sum_{d\leq x}\frac{\mu(d)}{d}\frac{1}{x/d}\sum_{m\leq\frac{x}{d}}\log^2 m.$$

Usando el Teorema 2.4.1 con $y = x/d$ y escribiendo $F(t) = 2\log(t) - c_3 + c_4 t^{-1}$

para simplificar la notación,

$$\begin{aligned}\frac{1}{x}\sum_{n\leq x}\Lambda_2(n) &= \sum_{d\leq x}\frac{\mu(d)}{d}\sum_{m\leq\frac{x}{d}}\frac{F(\frac{x}{dm})}{m}+\sum_{d\leq x}\frac{\mu(d)}{d}O\bigg(\frac{\log^2(1+\frac{x}{d})}{x/d}\bigg)\\ &= \sum_{d\leq x}\mu(d)\sum_{m\leq\frac{x}{d}}\frac{F(\frac{x}{dm})}{dm}+O\bigg(\sum_{d\leq x}\frac{\log^2(1+\frac{x}{d})}{x}\bigg)\\ &= \sum_{d\leq x}\mu(d)\sum_{\substack{n\leq x\\ d|n}}\frac{F(\frac{x}{n})}{n}+O\bigg(\sum_{d\leq x}\frac{\log^2(1+\frac{x}{d})}{x}\bigg)\\ &= \sum_{n\leq x}\frac{F(\frac{x}{n})}{n}\sum_{d|n}\mu(d)+O\bigg(\sum_{d\leq x}\frac{\log^2(1+\frac{x}{d})}{x}\bigg)\end{aligned}$$

pero como vimos en la demostración de la inversión de Möbius, $\sum_{d|n}\mu(d)$ es igual a 1 si $n=1$ e igual 0 en cualquier otro caso, entonces

$$\frac{1}{x}\sum_{n\leq x}\Lambda_2(n) = F(x)+O\bigg(\sum_{d\leq x}\frac{\log^2(1+\frac{x}{d})}{x}\bigg).$$

Además (2.2) nos garantiza que $\log^2(1+x)=O(\sqrt{x})$ de donde es sencillo ver que

$$\sum_{n\leq x}\Lambda_2(n) = 2x\log x+O(x). \qquad \square$$

3. Tópicos de análisis funcional

El análisis funcional es una rama de la matemática originada del álgebra lineal y áreas del análisis clásico como ecuaciones diferenciales, teoría de aproximación y cálculo de variaciones. Se centra primordialmente en espacios abstractos como los de Hilbert o Banach: espacios vectoriales dotados de una métrica que puede ser inducida por un producto interno o una norma.

Sea V un espacio vectorial real o complejo. Una función $||\cdot||: V \to \mathbb{R}$ es una *norma* si cumple las siguientes propiedades:

a) $||v|| \geq 0$ para todo $v \in V$ y $||v|| = 0$ si y solo si $v = 0$, (positividad)

b) $||\alpha v|| = |\alpha|\,||v||$ para todo $v \in V$ y para todo α escalar, (homogeneidad)

c) $||v + w|| \leq ||v|| + ||w||$ para todo $v, w \in V$ (Desigualdad del triángulo)

Si $||\cdot||$ es una norma en V, entonces $(V, ||\cdot||)$ es un *espacio normado*. En ocasiones surgen funciones $||\cdot||$ que cumplen homogeneidad, desigualdad del triángulo y toman valores no negativos, pero el conjunto de elementos en V cuya imagen es cero, resulta ser un subespacio lineal no nulo llamado *kernel* de $||\cdot||$. En dado caso, la función $||\cdot||$ se dice que es una *seminorma*.

Una norma $||\cdot||$ induce la métrica $d(u, v) := ||u - v||$, dentro del espacio normado. Análogamente, una seminorma induce una pseudométrica. Esto nos permite considerar convergencia de sucesiones, sucesiones de Cauchy y más aún, permite la introducción de una topología mediante la definición usual de bolas abiertas. En el caso de la seminorma, su kernel es cerrado en esta topología.

Teorema 3.0.1. *El kernel*

$$V_0 := \{v \in V \mid ||v|| = 0\}$$

de una seminorma $||\cdot||$ en un espacio vectorial es un subespacio cerrado en la topología inducida por la seminorma.

Demostración. Por la desigualdad del triángulo, la seminorma es continua respecto a la topología inducida por ella misma. De modo que $V_0 = (||\cdot||)^{-1}\{0\}$ debe ser cerrado. □

Nota 3.0.2. En un espacio infinitodimensional V es posible tener un subespacio lineal W no cerrado. Consideremos por ejemplo el espacio normado de funciones continuas definidas en $[0,1]$ con la norma del máximo $V = (C[0,1], \text{máx}\,|\cdot|)$ y el subespacio de todos los polinomios definidos en el mismo intervalo $W = (P[0,1], \text{máx}\,|\cdot|)$. Por el Teorema de Stone-Weierstrass, W es denso en V así que W no es cerrado.

Gracias a esto, es posible construir una norma a partir de una seminorma, que aunque definida en otro espacio, posee una íntima relación con el espacio original y la seminorma definida en este.

Teorema 3.0.3. *Para todo espacio vectorial V equipado con una seminorma $||\cdot||$ y todo subespacio cerrado $W \subset V$, la expresión*

$$||v + W||_{V/W} := \inf_{w \in W} ||v + w||$$

para $v \in V$ define una norma en el espacio cociente $V/W = \{v + W \mid v \in V\}$. Más aún, si $V_0 := \{v \in V \mid ||v|| = 0\}$, entonces $||v + V_0||_{V/V_0} = ||v||$.

Demostración. Para una demostración, ver lema 2.15 de [8]. □

Si un espacio normado es *completo*, es decir que toda sucesión de Cauchy en V converge en V (bajo la métrica inducida por $||\cdot||$), decimos que $(V, ||\cdot||)$ es un *espacio de Banach*. Esta definición es sensata pues no todo espacio normado es de Banach, basta notar que $\mathbb{Q}$ como espacio vectorial sobre sí mismo con la norma del valor absoluto no es completo. Sin embargo, $\overline{\mathbb{Q}} = \mathbb{R}$ que también es un espacio normado sobre $\mathbb{Q}$ sí lo es, lo que sugiere la posiblidad de «completar» un espacio normado dado.

3.1. Completación de un espacio normado

Sea $(V, ||\cdot||)$ un espacio normado. Puesto que las sucesiones de Cauchy en V están dictadas por la métrica inducida por $||\cdot||$, haremos la primera parte de la discusión haciendo enfásis en un espacio métrico (V, d) y luego utilizaremos los

resultados para extender la hipótesis a un espacio normado. La idea principal es «agregar» los límites de las sucesiones de Cauchy a V, tomando en cuenta que pueden haber sucesiones que converjan al mismo límite. Estas dos consideraciones nos llevan a definir

$$W := \{(v_n) \mid (v_n) \text{ es de Cauchy en } V\}$$

y la relación

$$(v_n) \sim (v_n') \text{ si y solo si } \lim_{n\to\infty} d(v_n, v_n') = 0,$$

que resulta ser de equivalencia en W. Entonces de manera natural, tornamos nuestra atención al conjunto de clases de equivalencia $\hat{V}$, al cual podemos dotar de la métrica

$$\hat{d}(\hat{v}, \hat{u}) = \lim_{n\to\infty} d(v_n, u_n)$$

donde (v_n) y (u_n) son representantes de $\hat{v}$ y $\hat{u}$, respectivamente.

Lo construido hasta el momento parece prometedor, inmediatamente vemos que una propiedad conveniente de $\hat{V}$ es la posibilidad de asociar de forma única a cualquier vector $v \in V$ con la clase de equivalencia a la que pertenece la sucesión constante $(v, v, v, \cdots)$. Es decir, V tiene una «copia» X en $\hat{V}$, en la cual $\hat{d}$ se comporta como d. Sin embargo, al haber más elementos en $\hat{V}$, es posible «asignar un límite» a toda sucesión en W.

Sea $(\hat{x}_n)$ una sucesión en $X \subset \hat{V}$ (una sucesión de clases de equivalencia), cada elemento $\hat{x}_n$ en la sucesión posee un representante $(x_n) = (v_n, v_n, v_n, \cdots)$, entonces intuitivamente $(\hat{x}_n)$ representa a la sucesión (v_n) en V. Si $(v_n) \in W$, esta pertenece a una clase de equivalencia $\hat{v}$ que resulta ser el límite de $(\hat{x}_n)$. En otras palabras, toda sucesión en W «converge en $\hat{V}$».

Por otro lado, se puede demostrar que $\overline{X} = \hat{V}$ y esto nos permite asociar a toda sucesión de Cauchy en $\hat{V}$ con una sucesión $(v_n) \in W$ y esto implica que $\hat{V}$ es completo. Este es un bosquejo de la demostración de

Teorema 3.1.1 (Completación de un Espacio)**.** *Para un espacio métrico (V, d) existe un espacio métrico completo $(\hat{V}, \hat{d})$ que tiene un subespacio X isométrico a V y denso en $\hat{V}$.*

Demostración. Para una demostración a detalle, ver Teorema 1.6-2 de [13]. □

Ahora, si V fuera un espacio normado, podemos considerar la métrica inducida

por norma y construir un espacio de Banach $\hat{V}$ con propiedades análogas a las de este teorema. Basta con definir adecuadamente las operaciones algebraicas en $\hat{V}$ y una norma que coincida con la métrica.

Si $\hat{v}$ y $\hat{u}$ son elementos de $\hat{V}$, con representantes (v_n) y (u_n), entonces la sucesión (z_n) definida por $z_n := v_n + u_n$ para $n \geq 1$, es también de Cauchy y por lo tanto pertenece a una clase de equivalencia $\hat{z}$, que definimos de manera natural como el resultado de $\hat{u} + \hat{v}$. Análogamente, se define el producto por escalar. Finalmente, para todo $\hat{v} \in V$ podemos definir $||v||_{\hat{V}} := \hat{d}(\hat{0}, \hat{v})$ que claramente coincide con la métrica y en X imita el comportamiento de la norma en V.

Teorema 3.1.2. *Si $(V, ||\cdot||)$ es un espacio normado. Existe un espacio de Banach $\hat{V}$ que tiene un subespacio X isométrico a V y denso en $\hat{V}$.*

Demostración. Para una demostración a detalle ver Teorema 2.3-2 de [13]. □

3.2. Topología débil*

En el estudio de espacios vectoriales, uno de los objetos más importantes son las transformaciones lineales y más aún, los funcionales lineales, es decir transformaciónes lineales del espacio vectorial al campo. Si se trabaja en un espacio normado se puede tomar en cuenta la noción de continuidad de los funcionales. Si X es un espacio normado, el *dual de* X, denotado por X^*, consiste de todos los funcionales lineales continuos definidos en X. El espacio X^* es un espacio normado, bajo la *norma operador* $||\varphi|| := \sup_{||x||=1} ||\varphi(x)||$ para $\varphi \in X^*$. La norma operador de un funcional φ es el menor número real tal que

$$\frac{||\varphi(x)||}{||x||} \leq ||\varphi|| \quad \text{para todo } x \in X. \tag{3.1}$$

Para un funcional φ, si la expresión de la izquierda en (3.1) está acotada decimos que φ es un *funcional acotado*. Por lo mencionado arriba, todo funcional continuo es acotado, pero también todo funcional acotado es continuo.

El uso de teoremas relacionados a continuidad es una práctica común. Para deducir propiedades de un objeto concreto, se parte de alguna sucesión de elementos que convergen a él y que se conocen de mejor manera. El sentido de convergencia depende de la topología utilizada y una topología de particular interés, definida en el dual X^* de un espacio normado X, es la topología débil*.

Dado X un espacio normado, consideremos para cada $x \in X$, el funcional lineal $e_x \colon X^* \to \mathbb{R}$ definido como $\varphi \mapsto e_x(\varphi) = \varphi(x)$ (también llamado mapeo de evaluación). Conforme x varía en X, se obtiene una colección de funcionales $(e_x)_{x \in X}$ en X^*. La *topología débil** es la topología inducida por estos mapeos. Es decir, es la más débil definida en X^*, para la cual todo mapeo de la colección $(e_x)_{x \in X}$ es continuo.

Nota 3.2.1. En espacios finitodimensionales el mapeo $x \mapsto e_x$ para todo $x \in X$ es un isomorfismo.

La topología debil* depende de los mapeos de evaluación y estos a su vez dependen de elementos de X. Esto sugiere la existencia de una conexión entre convergencia en la topología débil* y los elementos de X. El siguiente teorema nos muestra tal conexión:

Teorema 3.2.2. *Sea (φ_n) una sucesión de elementos en X^*. Entonces φ_n converge a φ en la topología débil*, denotado como $\varphi_n \overset{*}{\rightharpoonup} \varphi$, si y solo si $\varphi_n(x) \to \varphi(x)$ para todo $x \in X$.*

Demostración. Para una demostración ver proposición 3.13 de [4]. □

Como la topología debil* es más débil que otras, posee más conjuntos compactos, lo que la hace interesante. De hecho, la propiedad más importante de esta topología es que, en ella, la bola unitaria cerrada es siempre compacta. Este resultado se conoce como el Teorema de Banach-Alaoglu-Bourbaki:

Teorema 3.2.3 (Banach-Alaoglu-Bourbaki)**.** *La bola cerrada*

$$B_{X^*} = \{\varphi \in X^* \mid ||\varphi|| \leq 1\}$$

es compacta en la topología debil.*

Demostración. Para una demostración ver Teorema 3.16 de [4]. □

Este teorema nos permite asegurar la existencia de un funcional al cual converge puntualmente una subsucesión de una sucesión acotada de funcionales. Tal existencia será utilizada en la demostración del Teorema 5.2.2.

3.3. El espacio $C_c(\mathbb{R})$

En la demostración de TNP, será de vital importancia el espacio vectorial de funciones complejas definidas en los reales, continuas y de soporte compacto, denotado por $C_c(\mathbb{R})$. El *soporte de una función* f, $\text{supp}(f)$ se define como

$$\text{supp}(f) = \overline{\{x \in \mathbb{R} \mid f(x) \neq 0\}}.$$

El espacio $C_c(\mathbb{R})$ tiene algunas propiedades interesantes, entre ellas podemos resaltar que la integral de Lebesgue de una de estas funciones es relativamente simple. Si $f \in C_c(\mathbb{R})$, por ser una función continua es Riemann integrable y su integral de Lebesgue coincide con su integral de Riemann. Además por su soporte, dicha integral de Riemann es propia. Más aún, la continuidad en el compacto nos garantiza un máximo y un mínimo, por lo que cada una de estas funciones está en $L_1(\mathbb{R})$ con la medida de Lebesgue. Es más, cualquier función en $L_1(\mathbb{R})$ puede ser aproximada por funciones en $C_c(\mathbb{R})$, este es un hecho crucial para la demostración del TNP.

Teorema 3.3.1. *El espacio $C_c(\mathbb{R})$ es denso en $L_1(\mathbb{R})$ con una medida μ σ-finita.*

Demostración. Para una demostración ver Proposición 2.51 de [8]. □

De hecho podemos ir más alla y considerar el espacio $C_c^\infty(\mathbb{R})$ de *funciones suaves de soporte compacto* para el cual también se verifica

Teorema 3.3.2. *El espacio $C_c^\infty(\mathbb{R})$ es denso en $L_1(\mathbb{R})$ con una medida μ σ-finita.*

Demostración. Para una demostración ver Corolario 4.23 de [4]. □

El buen comportamiento de $C_c(\mathbb{R})$ respecto a las integrales permite caracterizar a sus funcionales lineales como mapeos de evaluación respecto a una integral con una medida localmente finita. En el fondo, esto representa una dualidad entre funcionales y medidas, de una manera similar a como existe una dualidad entre X y $(X^*)^*$ para espacios vectoriales finitodimensionales (ver Nota 3.2.1). Las dualidades nos permiten estudiar un objeto desde distintas perspectivas en las que se conserva su estructura esencial y por ello, intercambiarlas según sea conveniente para nuestro objetivo.

Teorema 3.3.3 (Teorema de Representacion de Riesz)**.** *Sea $\varphi\colon C_c(\mathbb{R}) \to \mathbb{C}$ un funcional lineal que cumple $\varphi(f) \geq 0$ si $f \geq 0$. Entonces existe una única medida*

de Borel localmente finita μ tal que

$$\varphi(f) = \int f \, d\mu$$

para todo $f \in C_c(\mathbb{R})$.

Demostración. Para una demostración ver Teorema 7.44 de [8]. □

3.4. Álgebras de Banach

Recordemos que un álgebra $\mathcal{A}$ sobre un campo $\mathbb{F}$ es un espacio vectorial $\mathcal{A}$ sobre $\mathbb{F}$ tal que para cada pareja $(x, y) \in \mathcal{A}^2$ existe un único producto $xy \in \mathcal{A}$ que cumple las siguientes propiedades:

a) $(xy)z = x(yz)$, (asociatividad)

b) $x(y + z) = xy + xz$, (distributividad izquierda respecto a la suma)

c) $(x + y)z = xz + yz$, (distributividad derecha respecto a la suma)

d) $\alpha(xy) = (\alpha x)y = x(\alpha y)$ (distributividad respecto al producto escalar)

para cualesquiera $x, y, z \in \mathcal{A}$ y escalares α.

Un *álgebra normada* $\mathcal{A}$, es un espacio normado que es a la vez un álgebra, tal que para todo $x, y \in \mathcal{A}$

$$||xy|| \leq ||x|| \, ||y||. \tag{3.2}$$

La desigualdad (3.2) implica que la operación producto es continua.

Teorema 3.4.1. *El producto definido en un álgebra normada $\mathcal{A}$ es continuo.*

Demostración. Sean $a, b \in \mathcal{A}$ y $0 < \varepsilon < 1$. Tomando $\delta < \frac{\varepsilon}{||a||+1+||b||}$, se tiene que para todo x, y con $||x - a|| < \delta$ y $||y - b|| < \delta$,

$$\begin{aligned} ||xy - ab|| &\leq ||x(y - b)|| + ||(x - a)b|| \\ &< (||x|| + ||b||)\delta \\ &< (||a|| + 1 + ||b||)\delta \\ &< \varepsilon. \end{aligned}$$

□

Si $\mathcal{A}$ tiene una identidad multiplicativa e, se dice que es un *álgebra unitaria* y podemos asumir, reescalando la norma de ser necesario, que

$$||e|| = 1.$$

Finalmente, un *álgebra de Banach* es un álgebra normada que es completa considerada como un espacio normado. Durante esta discusión las álgebras de Banach serán sobre $\mathbb{C}$ a menos que sea especificado de otra manera.

Las álgebras de Banach unitarias dan lugar a la existencia de elementos *invertibles*, es decir $x \in \mathcal{A}$ tales que existe x^{-1} que cumple $xx^{-1} = x^{-1}x = e$. Los elementos invertibles a su vez, dan lugar al concepto de espectro y resolvente. Sea $\mathcal{A}$ un álgebra de Banach con identidad, el *espectro* $\sigma(x)$, de un elemento $x \in \mathcal{A}$, se define como

$$\sigma(x) := \{\lambda \in \mathbb{C} \mid x - \lambda e \text{ no es invertible}\},$$

y el *conjunto resolvente* $\rho(x)$ como

$$\rho(x) := \{\lambda \in \mathbb{C} \mid x - \lambda e \text{ es invertible}\}.$$

El espectro generaliza la noción de valor propio: para un álgebra $\mathcal{A}$ de mapeos lineales, el espectro de un elemento $T \in \mathcal{A}$ es igual al conjunto de valores propios de T. El teorema fundamental del álgebra nos asegura que todo operador lineal definido en un espacio finitodimensional, posee al menos un valor propio. Esto nos motiva a preguntarnos si para todo elemento x en un álgebra de Banach unitaria, su espectro es no vacío y en efecto,

Teorema 3.4.2. *Sea $\mathcal{A}$ un álgebra de Banach unitaria. Para todo $x \in \mathcal{A}$ se tiene que $\sigma(x) \neq \emptyset$. Además, $\sigma(x) \subset \{\lambda \in \mathbb{C} \mid |\lambda| \leq ||x||\}$.*

Demostración. Para una demostración ver Teoremas 7.7-3 y 7.7-4 de [13]. □

Podemos estudiar geométricamente este conjunto midiendo «qué tan grandes» son sus elementos, por lo que consideramos el radio espectral. Para x un elemento de un álgebra de Banach unitaria $\mathcal{A}$, definimos su *radio espectral* $r_\sigma(x) := \sup_{\lambda \in \sigma(x)} |\lambda|$. El siguiente teorema relaciona la estructura algebraica de un álgebra de Banach con su estructura topológica.

Teorema 3.4.3 (Fórmula del Radio Espectral)**.** *Sea $\mathcal{A}$ un álgebra de Banach unitaria, entonces para cada $x \in \mathcal{A}$ se tiene*

$$r_\sigma(x) = \lim_{n\to\infty} \sqrt[n]{||x^n||}. \tag{3.3}$$

Demostración. Para una demostración ver Teorema 11.6 de [8]. □

3.5. Caracteres

En el caso de que un álgebra de Banach $\mathcal{A}$ sea conmutativa, es decir que $ab = ba$ para cualesquiera $a, b \in \mathcal{A}$, es útil estudiar funciones definidas en $\mathcal{A}$ que preserven esta propiedad (llamados también homomorfismos). En particular, nos centraremos en funcionales en $\mathcal{A}$.

Se define el *dual de Gelfand* $\sigma(\mathcal{A})$, como el conjunto de todos los homomorfismos continuos $\psi\colon \mathcal{A} \mapsto \mathbb{C}$, a los cuales nos referiremos como *caracteres*. Es decir,

$$\sigma(\mathcal{A}) := \{\psi \in \mathcal{A}^* \backslash \{0\} \mid \psi(ab) = \psi(a)\psi(b), \text{ para todo } a, b \in \mathcal{A}\}.$$

Nota 3.5.1. En realidad todo homomorfismo algebraico $\psi\colon \mathcal{A} \to \mathbb{C}$ es continuo, ver lema 11.24 de [8]. Por lo que técnicamente, la hipótesis de continuidad en la definición anterior puede ser excluida.

Como primera propiedad interesante tenemos lo siguiente:

Teorema 3.5.2. *Sea $\mathcal{A}$ un álgebra de Banach unitaria y conmutativa. Todos los caracteres en $\mathcal{A}$ tienen norma (de operador) menor o igual a 1.*

Demostración. Sea ψ un caracter en $\mathcal{A}$. Consideremos $x \in \mathcal{A}$ tal que $\psi(x) \neq 0$. Notemos que $y := x - \psi(x)e$, pertenece al kernel de ψ, por lo que no puede ser invertible. De modo que $\psi(x) \in \sigma(x)$ y $|\psi(x)| \leq ||x||$ por el Teorema 3.4.2. □

Nota 3.5.3. De hecho, hemos demostrado que $\{\psi(x) \mid \psi \in \sigma(\mathcal{A})\} \subset \sigma(x)$.

Esta demostración es fascinante porque a partir de un caracter, obtuvimos un número del espectro $\sigma(x)$. Además, este proceso no fue dependiente del caracter elegido, ni del elemento x. Así que surge la pregunta, ¿será posible para todo elemento $x \in \mathcal{A}$, asociar cada $\lambda \in \sigma(x)$ con $\psi(x)$ para algún $\psi \in \sigma(\mathcal{A})$? La respuesta es sí.

Antes de proseguir, definimos algunos términos básicos de álgebra. Un *ideal* $\mathcal{J}$ de un álgebra conmutativa $\mathcal{A}$ es un subespacio tal que

$$\mathcal{A}\mathcal{J} = \{xj \mid x \in \mathcal{A}, j \in \mathcal{J}\} \subset \mathcal{J}$$

y para todo ideal $\mathcal{J}$ el cociente $A/\mathcal{J}$ es también un álgebra conmutativa donde la multiplicación está dada por

$$(x + \mathcal{J})(y + \mathcal{J}) = xy + \mathcal{J}$$

para cualesquiera $x, y \in \mathcal{A}$. Un ideal $\mathcal{J} \subset \mathcal{A}$ se dice que es *propio* si $\mathcal{J} \neq \mathcal{A}$. Un *ideal maximal* $\mathcal{M}$ en un álgebra unitaria conmutativa $\mathcal{A}$ es un ideal propio que cumple que si $\mathcal{J}$ es un ideal de $\mathcal{A}$ tal que $\mathcal{M} \subset \mathcal{J} \subset \mathcal{A}$, entonces $\mathcal{J} = \mathcal{M}$ o $\mathcal{J} = \mathcal{A}$.

Además, presentamos los siguientes resultados,

Proposición 3.5.4. *El cociente $\mathcal{A}/\mathcal{M}$, donde $\mathcal{A}$ es un álgebra unitaria y $\mathcal{M}$ un ideal maximal de $\mathcal{A}$, es un campo.*

Demostración. Sea $x + \mathcal{M} \in \mathcal{A}/\mathcal{M}\backslash\{0\}$. El ideal $\mathcal{J} := \mathcal{A}x + \mathcal{M}$ es estrictamente mayor a $\mathcal{M}$, entonces $\mathcal{J} = \mathcal{A}$. Como $\mathcal{A}$ es unitaria, existe un elemento $a \in \mathcal{A}$ y $m \in \mathcal{M}$ tal que $ax + m = e$. Así que todo elemento no nulo en $\mathcal{A}/\mathcal{M}$ tiene un inverso. □

Proposición 3.5.5. *Sea $\mathcal{A}$ un álgebra de Banach unitaria y conmutativa. En la topología inducida por norma, la clausura de cualquier ideal en $\mathcal{A}$ es un ideal.*

Demostración. Para una demostración ver lema 11.25 de [8]. □

Proposición 3.5.6. *Sea $\mathcal{A}$ un álgebra de Banach unitaria y conmutativa. Todo ideal maximal $\mathcal{M}$ es cerrado en la topología inducida por norma.*

Demostración. Para una demostración ver lema 11.25 de [8]. □

Proposición 3.5.7. *Sea R un anillo conmutativo con unidad y $\mathcal{J}_0 \subset R$ un ideal propio. Entonces existe un ideal maximal $\mathcal{M} \subset R$ tal que $\mathcal{J}_0 \subset \mathcal{M}$.*

Demostración. Para una demostración ver lema 11.26 de [8]. □

Con esto cubierto, procedemos a mostrar la asociación entre el espectro de un elemento x y sus imágenes respecto a los caracteres.

Teorema 3.5.8. *Sea $\mathcal{A}$ un álgebra de Banach unitaria y conmutativa. Para todo $x \in \mathcal{A}$ se cumple*

$$\sigma(x) = \{\psi(x) \mid \psi \in \sigma(\mathcal{A})\}.$$

Demostración. Por la Nota 3.5.3 únicamente falta demostrar la contención

$$\sigma(x) \subset \{\psi(x) \mid \psi \in \sigma(\mathcal{A})\}.$$

Comencemos considerando un elemento $x_0 \in \mathcal{A}$ no invertible. Notemos que $\mathcal{J} := \mathcal{A}x_0$ es un ideal propio de $\mathcal{A}$, por la Proposición 3.5.7 existe un ideal maximal $\mathcal{M}$ que contiene a $\mathcal{J}$ y por la Proposición 3.5.6 $\mathcal{M}$ es cerrado.

Demostraremos que $\mathscr{B} := \mathcal{A}/\mathcal{M}$ también es un álgebra de Banach. En efecto, al equipar a $\mathscr{B}$ con la norma cociente, este se vuelve un espacio de Banach. Como $\mathcal{M}$ es un ideal, la multiplicación en $\mathscr{B}$ está bien definida y finalmente,

$$||xy + \mathcal{M}||_{\mathcal{A}/\mathcal{M}} \leq ||(x + m_1)(y + m_2)||_{\mathcal{A}} \leq ||(x + m_1)||_{\mathcal{A}} \leq ||(y + m_2)||_{\mathcal{A}}$$

para cualesquiera $x, y \in \mathcal{A}$ y $m_1, m_2 \in \mathcal{M}$. Esto implica, tomando el ínfimo respecto a m_1 y m_2, que

$$||(x + \mathcal{M})(y + \mathcal{M})||_{\mathcal{A}/\mathcal{M}} \leq ||x + \mathcal{M}||_{\mathcal{A}/\mathcal{M}} ||y + \mathcal{M}||_{\mathcal{A}/\mathcal{M}}$$

como se quería verificar.

Por otro lado, el Teorema 3.5.4 nos asegura que $\mathscr{B}$ también es un campo pues $\mathcal{M}$ es maximal y veremos que este es isomorfo a $\mathbb{C}$. Por el Teorema 3.4.2, para todo $x + \mathcal{M} \in \mathscr{B}$, existe $\lambda \in \mathbb{C}$ tal que $(x + \mathcal{M}) - \lambda(e + \mathcal{M})$ es no invertible, pero al ser $\mathscr{B}$ un campo esto es posible si y solo si $x + \mathcal{M} = \lambda e + \mathcal{M}$. Esto induce el mapeo $\Phi \colon \mathscr{B} \to \mathbb{C}$ definido por $x + \mathcal{M} \mapsto \lambda$. Para $x + \mathcal{M}$ y $y + \mathcal{M}$ en $\mathscr{B}$ denotamos $\alpha := \Phi(x + \mathcal{M})$, $\beta := \Phi(y + \mathcal{M})$ y $\lambda := \Phi(xy + \mathcal{M})$, entonces

$$\begin{aligned} \lambda(e + \mathcal{M}) &= xy + \mathcal{M} \\ &= (x + \mathcal{M})(y + \mathcal{M}) \\ &= (\alpha e + \mathcal{M})(\beta e + \mathcal{M}) \\ &= \alpha\beta(e + \mathcal{M}). \end{aligned}$$

Es decir, $\lambda = \alpha\beta$ y Φ es un homomorfismo. Si $\Phi(x + \mathcal{M}) = \Phi(y + \mathcal{M})$,

$$x + \mathcal{M} = \lambda e + \mathcal{M} = y + \mathcal{M},$$

entonces Φ es inyectivo y puesto que $\Phi(\lambda e + \mathcal{M}) = \lambda$ para todo $\lambda \in \mathbb{C}$, concluimos que Φ es un isomorfismo.

Por lo anterior, si x_0 es no invertible existe $\psi\colon \mathcal{A} \to \mathscr{B} \cong \mathbb{C}$ un homomorfismo no trivial con $\psi(x_0) = 0$. Si utilizamos este hecho para $x - \lambda e$ donde $\lambda \in \sigma(x)$, se concluye de la Nota 3.5.1 que existe $\psi \in \sigma(\mathcal{A})$ tal que $\psi(x) = \lambda$. Es decir, $\sigma(x) \subset \{\psi(x) \mid \psi \in \sigma(\mathcal{A})\}$. □

Este teorema tiene como corolario uno de los resultados más importantes, respecto a los caracteres, para la demostración de TNP.

Corolario 3.5.9 (Fórmula de Gelfand)**.**

$$\max_{\psi \in \sigma(\mathcal{A})} |\psi(x)| = \lim_{n \to \infty} \sqrt[n]{||x^n||}.$$

Demostración. Se sigue directamente del Teorema 3.4.3 y el Teorema 3.5.8. □

Es importante notar que estos últimos teoremas aplican a álgebras de Banach unitarias. Es posible «extender» su dominio a álgebras de Banach sin unidad, cambiando la perspectiva respecto a la cuál consideramos estas álgebras.

Toda álgebra de Banach sin unidad $\mathcal{A}$ puede considerarse como una subálgebra de Banach de un álgebra de Banach unitaria $\mathcal{A}_1$. De manera burda, esto quiere decir que podemos trabajar en una estructura que sí posee unidad, pero que opera de forma análoga al álgebra sin unidad. Para ver esto, basta tomar $\mathcal{A}_1 := \mathcal{A} \times \mathbb{C}$ (una extensión natural del espacio vectorial), definiendo el producto vectorial y la norma como

$$\begin{aligned}(x, \lambda) \cdot (y, \eta) &:= (xy + \lambda y + \eta x, \lambda\eta),\\ ||(a, \lambda)|| &:= ||a|| + |\lambda|.\end{aligned}$$

Se verifica fácilmente que la unidad en esta álgebra es el elemento $(0, 1)$ y el mapeo $x \mapsto (x, 0)$ de $\mathcal{A}$ hacia $\mathcal{A}_1$ es un homomorfismo isométrico.

4. Análisis de Fourier

4.1. Definición y propiedades de la transformada de Fourier

La *transformada de Fourier* $\widehat{f}$ de una función $f\colon \mathbb{R} \to \mathbb{C}$ se define como

$$\widehat{f}(t) = \int_{\mathbb{R}} f(x) e^{-2\pi i x t} dx, \tag{4.1}$$

y la *transformada inversa de Fourier* $\check{h}$ de una función $h\colon \mathbb{R} \to \mathbb{C}$ se define como

$$\check{h}(x) = \int_{\mathbb{R}} h(t) e^{2\pi i x t} dt. \tag{4.2}$$

Claramente, la transformada de Fourier no tiene por qué existir para una función f arbitraria, es decir si f no fuera integrable es posible que su transformada no esté definida. Afortunadamente, la integrabilidad (en el sentido de Lebesgue) es una condición suficiente de existencia. Si f está en $L_1(\mathbb{R})$, entonces

$$|\widehat{f}(t)| = \left| \int_{\mathbb{R}} f(x) e^{-2\pi i x t} dx \right| \leq \int_{\mathbb{R}} |f(x)| dx = ||f||_1.$$

La transformada de Fourier interactúa muy bien con ciertos *operadores de traslación* y de *convolución de funciones*. Para x_0 y t_0 reales, definimos el operador traslación T_{x_0} y el operador multiplicación M_{t_0} como

$$T_{x_0}(f)\colon x \longmapsto f(x - x_0)$$

$$M_{t_0}(f)\colon x \longmapsto e^{2\pi i x t_0} f(x).$$

De la definición se sigue fácilmente que $\widehat{T_{x_0}(f)} = M_{-x_0}(\widehat{f}\,)$ y $\widehat{M_{t_0}(f)} = T_{t_0}(\widehat{f}\,)$.

Por otro lado, si $f_1, f_2 \in L_1(\mathbb{R})$ definimos $f_1 * f_2$ en semejanza a la convolución de Dirichlet como

$$(f_1 * f_2)(x) = \int_{\mathbb{R}} f_1(y) f_2(x-y) dy.$$

El siguiente teorema nos expone algunas propiedades de la transformada de Fourier y la convolución que nos serán útiles para demostar TNP. Además ejemplifica la interacción mencionada entre la transformada de Fourier y la convolución de funciones.

Teorema 4.1.1. *Para $f_1, f_2 \in L_1(\mathbb{R})$, se tiene*

a) $||f_1 * f_2||_1 \leq ||f_1||_1 ||f_2||_1$ *y por lo tanto* $f_1 * f_2 \in L_1(\mathbb{R})$,

b) $f_1 * f_2 = f_2 * f_1$,

c) $\widehat{f_1 * f_2} = \widehat{f_1}\widehat{f_2}$,

d) $\int_{\mathbb{R}} \widehat{f} g \, dx = \int_{\mathbb{R}} f \widehat{g} \, dy$.

Demostración. a) Por desigualdad del triángulo,

$$\int_{\mathbb{R}} \left| \int_{\mathbb{R}} f_1(y) f_2(x-y) dy \right| dx \leq \int_{\mathbb{R}} \int_{\mathbb{R}} |f_1(y) f_2(x-y)| dy dx.$$

Por medio el Teorema de Fubini y el cambio de variable $u = x - y$ lo anterior se convierte en

$$\int_{\mathbb{R}} \int_{\mathbb{R}} |f_1(y) f_2(x-y)| dx dy = \int_{\mathbb{R}} |f_1(y)| dy \cdot \int_{\mathbb{R}} |f_2(u)| du = ||f_1||_1 ||f_2||_1.$$

Esto garantiza que $f_1 * f_2 \in L_1(\mathbb{R})$.

b) Haciendo $u = x - y$,

$$(f_1 * f_2)(x) = \int_{\mathbb{R}} f_1(y) f_2(x-y) dy = \int_{\mathbb{R}} f_1(x-u) f_2(u) du = (f_2 * f_1)(x)$$

c) Usando Fubini,

$$\widehat{f_1 * f_2}(t) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f_1(y) f_2(x-y) dy \right) e^{-2\pi i x t} dx$$

$$
\begin{aligned}
&= \int_{\mathbb{R}}\int_{\mathbb{R}} f_1(y)e^{-2\pi iyt}f_2(x-y)e^{-2\pi i(x-y)t}dydx \\
&= \int_{\mathbb{R}} e^{-2\pi iyt}f_1(y)\Bigg(\underbrace{\int_{\mathbb{R}} e^{-2\pi i(x-y)t}f_2(x-y)dx}_{\widehat{f_2}(t)}\Bigg)dy \\
&= \widehat{f_1}(t)\widehat{f_2}(t).
\end{aligned}
$$

d) De nuevo, utilizando Fubini

$$
\begin{aligned}
\int_{\mathbb{R}} \widehat{f}(x)g(x)\,dx &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(y)e^{-2\pi iyx}g(x)\,dydx \\
&= \int_{\mathbb{R}} f(y)\Bigg(\int_{\mathbb{R}} g(x)e^{-2\pi iyx}dx\Bigg)dy \\
&= \int_{\mathbb{R}} f(y)\widehat{g}(y)dy. \qquad \square
\end{aligned}
$$

A continuación presentamos un teorema que acredita sensatez a la definición y nombre de $\check{f}$.

Teorema 4.1.2 (Inversión de Fourier)**.** *Si $f \in L_1(\mathbb{R})$ tiene $\widehat{f} \in L_1(\mathbb{R})$, entonces f es igual a la transformada inversa de $\widehat{f}$, denotada por $(\widehat{f}\,)^{\vee}$, casi en todas partes.*

La idea de la demostración, es utilizar el inciso (d) del teorema anterior para definir una sucesión de funciones f_r que por un lado converge a f y por otro a $(\widehat{f}\,)^{\vee}$.

Demostración. Definamos $\phi_{r,x_0}(t) := e^{2\pi itx_0}\phi_r(t)$, donde $\phi_r(t) := e^{-\pi|rt|^2}$ para cada $x_0 \in \mathbb{R}$ y $r > 0$. A partir de esto consideremos

$$
f_r(x_0) = \int_{\mathbb{R}} \widehat{f}(t)\phi_{r,x_0}(t)dt = \int_{\mathbb{R}} f(x)\widehat{\phi}_{r,x_0}(x)dx. \tag{4.3}
$$

A continuación mostraremos que $f_r \to (\widehat{f}\,)^{\vee}$ puntualmente cuando $r \to 0$. Y que $f_r \to f$ en $L_1(\mathbb{R})$ cuando $r \to 0$.

a) $\big(f_r \to (\widehat{f}\,)^{\vee}\big)$: Notemos que

$$
\big|\widehat{f}(t)e^{2\pi itx_0}e^{-\pi|rt|^2}\big| = \big|\widehat{f}(t)e^{-\pi|rt|^2}\big| \le \big|\widehat{f}(t)\big| \in L_1.
$$

Además, $e^{-\pi|rt|^2} \to 1$ si $r \to 0$. Entonces por el Teorema de convergencia

dominada (Teorema 1.6.4)

$$f_r(x_0) = \int_{\mathbb{R}} \widehat{f}(t)\phi_{r,x_0}(t)dt \to \int_{\mathbb{R}} \widehat{f}(t)e^{2\pi i t x_0}dt = (\widehat{f}\,)^{\vee}(x_0),$$

para todo x_0.

b) $(f_r \to f)$: Primero,

$$\widehat{\phi}_r(x) = \int_{\mathbb{R}} e^{-\pi|ry|^2}e^{-2\pi i y x}dy,$$

al hacer la sustitución $u = ry$,

$$\widehat{\phi}_r(x) = \frac{1}{r}\int_{\mathbb{R}} e^{-\pi|u|^2}e^{-2\pi i u}du$$

y puesto que $\widehat{g}(t) = e^{-\pi t^2} = \phi_r$ si $g(x) = e^{-\pi x^2} = \phi_1$ (para una demostración de esto referimos al lector al ejemplo 9.27 de [8])

$$\widehat{\phi}_r(x) = \frac{e^{-\pi|\frac{x}{r}|^2}}{r},$$

más aún

$$\widehat{\phi}_{r,x_0}(x) = \frac{e^{-\pi\left|\frac{x-x_0}{r}\right|^2}}{r}.$$

De aquí concluimos que ϕ_r está en $L_1(\mathbb{R})$ para cada r. Con esto obtenemos de la Ecuación (4.3) que

$$f_r(x_0) = \int_{\mathbb{R}} f(x)\frac{e^{-\pi\left|\frac{x-x_0}{r}\right|^2}}{r}dx = (f * \widehat{\phi}_r)(x_0).$$

Por Teorema 4.1.1 tenemos que $f * \widehat{\phi}_r \in L_1(\mathbb{R})$ y por lo tanto f_r también.

Procedemos a examinar la diferencia $(f * \widehat{\phi})(x_0) - f(x_0)$. En el Ejemplo 9.27 de [8] también se demuestra que $\int_{\mathbb{R}} e^{-\pi z^2}dz = 1$, así que la conmutatividad de la convolución y la sustitución $z = x/r$ implican

$$\begin{aligned}(f * \widehat{\phi})(x_0) - f(x_0) &= \int_{\mathbb{R}} f(x_0 - x)\frac{e^{-\pi|\frac{x}{r}|^2}}{r}\,dx - f(x_0)\\ &= \int_{\mathbb{R}} \big(f(x_0 - rz) - f(x_0)\big)e^{-\pi|z|^2}\,dz.\end{aligned}$$

Al tomar normas, usar desigualdad triangular y Fubini,

$$\begin{aligned}||f * \widehat{\phi} - f||_1 &= \int_{\mathbb{R}} \left| \int_{\mathbb{R}} \big(f(x_0 - rz) - f(x_0)\big) e^{-\pi|z|^2}\, dz \right| dx_0 \\ &\leq \int_{\mathbb{R}} \int_{\mathbb{R}} \big|f(x_0 - rz) - f(x_0)\big| e^{-\pi|z|^2}\, dz dx_0 \\ &= \int_{\mathbb{R}} ||T_{rz}(f) - f||_1 e^{-\pi|z|^2} dz\end{aligned}$$

Ahora, por la densidad de $C_c(\mathbb{R})$ en $L_1(\mathbb{R})$, para cada $\varepsilon > 0$ podemos hallar una función $g \in C_c(\mathbb{R})$ con $||f - g||_1 < \varepsilon/3$. Por la continuidad uniforme de g, existe $\delta > 0$ tal que $||T_{rz}(g) - g||_1 < \varepsilon/3$ si $|rz| < \delta$. De ello,

$$||T_{rz}(f) - f||_1 \leq ||T_{rz}(f) - T_{rz}(g)||_1 + ||T_{rz}(g) - g||_1 + ||g - f||_1 < \varepsilon.$$

así que $||T_{rz}(f) - f||_1 \to 0$ si $r \to 0$ para $z \in \mathbb{R}$ fijo. Así que por el Teorema de convergencia dominada $||f * \widehat{\phi} - f||_1 \to 0$. Para concluir, por el Teorema 1.6.3 existe una sucesión de f_r que converge casi en todas partes tanto a f como a $(\widehat{f}\,)^\vee$, de donde se sigue el resultado. □

Una de las condiciones en el Teorema 4.1.2 es que $\widehat{f} \in L_1(\mathbb{R})$. De no ser así, no está claro cómo podría ser posible recuperar a f de $\widehat{f}$. Por lo tanto, es conveniente considerar clases de funciones para las cuales esa condición se cumpla. En particular una clase muy interesante son las funciones de Schwartz: el *Espacio de Schwartz* denotado por $\mathscr{S}(\mathbb{R})$ se define como

$$\mathscr{S}(\mathbb{R}) := \{f \colon \mathbb{R} \to \mathbb{C} \mid f \text{ es suave y } ||x^\alpha f^{(\beta)}||_\infty < \infty, \text{ para cualesquiera } \alpha, \beta \in \mathbb{N}\}.$$

donde

$$||f||_\infty := \text{ínf}\{M \in \mathbb{R} \mid |f| \leq M \textit{ casi en todas partes}\}.$$

Esto es, intuitivamente, que la derivada de cualquier orden de una función $f \in \mathscr{S}(\mathbb{R})$, decrece muy rápido; dicho de manera más precisa, posee un *decaimiento super-polinomial*. Esta propiedad le garantiza a cualquier función f en este espacio ser $L_1(\mathbb{R})$.

Teorema 4.1.3. *Si $f \in \mathscr{S}(\mathbb{R})$ entonces $f \in L_1(\mathbb{R})$.*

Demostración.

$$\begin{aligned}\int_{\mathbb{R}} |f|dx &= \int_{|x|\leq 1} |f|dx + \int_{|x|\geq 1} |f|dx \\ &\leq 2 \sup_{|x|\leq 1} |f| + \sup_{x\in\mathbb{R}} |x^N f| \int_{|x|\geq 1} \frac{dx}{|x^N|}\end{aligned}$$

y esto es menor a infinito para $N \geq 2$. □

Sin embargo, el decaimiento super-polinomial permite algo aún más interesante,

Teorema 4.1.4. *La transformada de Fourier de una función de Schwartz es también una función de Schwartz. Más aún, la transformada de Fourier de $\mathscr{S}(\mathbb{R})$ a $\mathscr{S}(\mathbb{R})$ es continua y posee a la transformada inversa de Fourier como inversa continua.*

Demostración. Sea $f \in \mathscr{S}(\mathbb{R})$. Utilizando diferenciación bajo el signo de integral (Corolario 5.9 de [3]), integración por partes, y desigualdad triangular, respectivamente

$$\begin{aligned}\left|\left|x^\alpha \frac{d^\beta}{dx^\beta}\widehat{f}(x)\right|\right|_\infty &= \left|\left|x^\alpha \int_{\mathbb{R}} (-2\pi it)^\beta f(t) e^{-2\pi ixt} dt\right|\right|_\infty \\ &= \frac{(2\pi)^\beta}{(2\pi)^\alpha}\left|\left|\int_{\mathbb{R}} \frac{d^\alpha}{dt^\alpha}[t^\beta f(t)] e^{-2\pi ixt} dt\right|\right|_\infty \\ &\leq \frac{(2\pi)^\beta}{(2\pi)^\alpha}\int_{\mathbb{R}} \left|\frac{d^\alpha}{dt^\alpha}[t^\beta f(t)]\right| dt.\end{aligned}$$

Es sencillo ver que el espacio de Schwartz es cerrado bajo suma, por lo tanto $\frac{d^\alpha}{dt^\alpha}[t^\beta f(t)]$ está en $\mathscr{S}(\mathbb{R}) \subset L_1(\mathbb{R})$ y la desigualdad concluye el resultado. □

Para finalizar la discusión de este espacio, presentamos el siguiente resultado que surge como corolario del Teorema 3.3.2.

Teorema 4.1.5. *El espacio de Schwartz $\mathscr{S}(\mathbb{R})$ es denso en $L_1(\mathbb{R})$ con la medida de Lebesgue.*

Demostración. Notemos que $C_c^\infty(\mathbb{R}) \subset \mathscr{S}(\mathbb{R})$ y puesto que $C_c^\infty(\mathbb{R})$ es denso en $L_1(\mathbb{R})$ se sigue el resultado. □

4.2. Caracteres de $L_1(\mathbb{R})$

Para concluir la sección estableceremos una dualidad entre los caracteres de $L_1(\mathbb{R})$ y mapeos de evaluación de la transformada de Fourier. Comenzamos definiendo el álgebra de Banach de interés.

Teorema 4.2.1. *El conjunto $L_1(\mathbb{R})$ junto con la convolución $*$ como producto y la norma usual $||\cdot||_1$, es un álgebra de Banach conmutativa.*

Demostración. Del Teorema 4.1.1, si $f_1, f_2 \in L_1(\mathbb{R})$ entonces $f_1 * f_2 \in L_1(\mathbb{R})$. La asociatividad de la convolución se sigue de Fubini y la distributividad respecto a la suma es consecuencia de la linealidad de la integral. Finalmente (a) y (b) del Teorema 4.1.1 prueban que $(L_1(\mathbb{R}), ||\cdot||_1)$ con $*$ como producto, es un álgebra de Banach conmutativa. □

Como se vio al final del Capítulo 3 , es posible agregar una unidad a cualquier álgebra de Banach. Por medio de tal construcción, consideramos a $(L_1(\mathbb{R}), ||\cdot||_1)$ como subálgebra de un álgebra Banach unitaria en las discusiones posteriores.

Para establecer la dualidad entre los caracteres de $L_1(\mathbb{R})$, denotados por $\sigma(L_1(\mathbb{R}))$ y la transformada de Fourier ocupamos el siguiente lema.

Lema 4.2.2. *Sea ψ en $\sigma(L_1(\mathbb{R}))$ y sea $u \in L_1(\mathbb{R})$ tal que $\psi(u) = 1$. Se define $\gamma_\psi \colon \mathbb{R} \to \mathbb{C}$ como $\gamma_\psi(\alpha) = \psi\big(u(t+\alpha)\big)$. Entonces existe un único número $\xi_\psi \in \mathbb{R}$ tal que $\gamma_\psi(\alpha) = e^{2\pi i \xi_\psi \alpha}$ para todo $\alpha \in \mathbb{R}$.*

Demostración. Tanto ψ como la traslación en el argumento son continuas en $L_1(\mathbb{R})$, de modo que γ_ψ es continua. Usando un cambio de variable, notamos

$$\begin{aligned}\gamma_\psi(\alpha+\beta) &= \psi\big(u(t+\alpha+\beta)\big)\psi(u)\\ &= \psi\big(u(t+\alpha+\beta) * u(t)\big)\\ &= \psi\big(u(t+\alpha) * u(t+\beta)\big)\\ &= \gamma_\psi(\alpha)\gamma_\psi(\beta).\end{aligned}$$

En particular,

$$\gamma_\psi(\alpha)\gamma_\psi(-\alpha) = \gamma_\psi(0) = 1, \tag{4.4}$$

$$\gamma_\psi(n\alpha) = \gamma_\psi(\alpha)^n \text{ para } n \in \mathbb{N}. \tag{4.5}$$

Si $\alpha_0 \in \mathbb{R}$ es tal que $|\gamma_\psi(\alpha_0)| \neq 1$, podemos suponer por (4.4) que $|\gamma_\psi(\alpha_0)| > 1$ y (4.5) garantiza la existencia de un natural n tal que $\gamma_\psi(n\alpha_0) > ||u||_1$. Sin embargo, el Teorema 3.5.2 nos asegura que para todo $\alpha \in \mathbb{R}$

$$|\gamma_\psi(\alpha)| \leq ||\psi||\, ||u(t+\alpha)||_1 \leq ||u||_1 \text{ para todo } \alpha \in \mathbb{R}.$$

Entonces $|\gamma_\psi(\alpha)| = 1$. Esto junto con el Teorema 7.6.2 de [6] implican la existencia de una única función continua $h\colon \mathbb{R} \to \mathbb{R}$ que satisface $h(0) = 0$ y

$$\gamma_\psi(\alpha) = e^{2\pi i h(\alpha)} \text{ para todo } \alpha \in \mathbb{R}.$$

Reescribimos la ecuación funcional como

$$e^{2\pi i(h(\alpha+\beta)-h(\alpha)-h(\beta))} = 1$$

que implica

$$h(\alpha+\beta) - h(\alpha) - h(\beta) = n$$

para algún entero n, pero al ser h continua, el Teorema del valor intermedio implica que este entero no depende de α y β. Además, puesto que $h(0) = 0$ tenemos $n = 0$, es decir

$$h(\alpha+\beta) = h(\alpha) + h(\beta).$$

Finalmente, las soluciones de esta ecuación son de la forma $h(\alpha) = \xi\alpha$ para algún $\xi \in \mathbb{R}$. □

Nota 4.2.3. Las soluciones continuas $f\colon \mathbb{R} \to \mathbb{R}$ de la ecuación

$$f(x+y) = f(x) + f(y)$$

(al igual que $f(x+y) = f(x)f(y)$) se obtienen describiendo su comportamiento en $\mathbb{Q}$ y utilizando la densidad de $\mathbb{Q}$ en $\mathbb{R}$. Sin embargo, este procedimiento no se sigue de forma análoga para hallar soluciones continuas $f\colon \mathbb{R} \to \mathbb{C}$ puesto que la ecuación $x^n = z$ donde $z \in \mathbb{C}$ es fijo posee n soluciones distintas en $\mathbb{C}$, de ahí surge la necesidad del uso del Teorema 7.6.2 de [6] en la demostración anterior.

El siguiente teorema nos asegura que todo caracter ψ en $\sigma(L_1(\mathbb{R}))$ puede ser interpretado como un mapeo $f \mapsto \widehat{f}(\xi)$ para algún real ξ.

Teorema 4.2.4. *Sea ψ un caracter en $L_1(\mathbb{R})$. Para todo $f \in L_1(\mathbb{R})$ se cumple*

$$\psi(f) = \int_{\mathbb{R}} f(\alpha) e^{-2\pi i \alpha \xi_\psi} d\alpha = \widehat{f}(\xi_\psi).$$

Demostración. Sea $u \in L_1(\mathbb{R})$ como en el Lema 4.2.2. Entonces $\psi(f) = \psi(f * u)$ y al ser ψ continua, el lema nos garantiza

$$\begin{aligned}
\psi(f * u) &= \psi\bigg(\int_{\mathbb{R}} f(\alpha) u(t-\alpha) d\alpha \bigg) \\
&= \int_{\mathbb{R}} \psi\big(f(\alpha) u(t-\alpha)\big) d\alpha \\
&= \int_{\mathbb{R}} f(\alpha) \psi\big(u(t-\alpha)\big) d\alpha \\
&= \int_{\mathbb{R}} f(\alpha) e^{-2\pi i \xi_\psi \alpha} d\alpha \\
&= \widehat{f}(\xi_\psi).
\end{aligned}$$

□

5. Convolución y convergencia de medidas

Parte clave de la conexión entre TNP y el análisis funcional es la consideración de medidas μ como funcionales en $C(\mathbb{R})$ via integración. Es decir,

$$f \longmapsto \int_{\mathbb{R}} f \, d\mu.$$

Esto nos permitirá utilizar la topología débil* para concluir varios resultados relacionados a medidas particulares, que serán de utilidad para la demostración de TNP.

5.1. Medidas de Radon

Para comenzar, recordemos que una medida de Radon es una medida de Borel que es localmente finita. Dada una función $f\colon \mathbb{N} \to \mathbb{R}$ le asociamos una medida $\nu_f\colon \mathcal{B} \to \mathbb{R}$ de la siguiente manera

$$\nu_f := \sum_{n=1}^{\infty} \frac{f(n)}{n} \delta_{\log n}, \tag{5.1}$$

donde $\delta_{\log n}\colon \mathcal{B} \to \mathbb{R}$ es llamada la *medida de Dirac en* $p = \log n$ y está definida para p en general como

$$\delta_p(B) := \begin{cases} 1 & \text{si } p \in B \\ 0 & \text{si } p \notin B \end{cases} \quad \text{para } B \in \mathcal{B}. \tag{5.2}$$

Para toda $f\colon \mathbb{N} \to \mathbb{R}$, esta es una medida de Radon. En efecto, si K es compacto podemos hallar un intervalo $[0, m]$ que contiene a K, de donde obtenemos:

$$\nu_f(K) \leq \nu_f([0,m]) = \sum_{n=1}^{N} \frac{f(n)}{n} < \infty,$$

donde N es el mayor entero con $\log N \leq m$.

Similar a la convolución de funciones, para dos medidas de Radon en $[0,\infty)$, definimos su convolución como

$$(\nu_1 * \nu_2)(B) = \iint \chi_B(x+y)\, d\nu_1(x) d\nu_2(y) \tag{5.3}$$

Esta es de nuevo una medida de Radon en $[0,\infty)$, porque si $B = [-x,x]$, vemos que el lado derecho de (5.3) está acotado por $\nu_1([0,x])\nu_2([0,x]) < \infty$.

El siguiente teorema, muestra que la asociación $f \mapsto \nu_f$ es una conexión entre dos estructuras algebraicas; G el monoide de funciones $f\colon \mathbb{N} \to \mathbb{R}$ con la convolución de Dirichlet y H el monoide de medidas de Radon con la convolución de medidas.

Teorema 5.1.1. *Para dos funciones* $f_1, f_2\colon \mathbb{N} \to \mathbb{R}$

$$\nu_{f_1} * \nu_{f_2} = \nu_{f_1 \underset{D}{*} f_2}$$

Demostración. Sea B un conjunto de Borel,

$$\nu_{f_1} * \nu_{f_2}(B) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \chi_B(\log n + \log m)\frac{f_1(n)}{n}\frac{f_2(m)}{m}$$

reescribiendo esta suma con índice $k = mn$,

$$\begin{aligned}
&= \sum_{k=1}^{\infty} \chi_B(\log k) \sum_{d|k} \frac{f_1(d)}{d}\frac{f_2(k/d)}{k/d} \\
&= \sum_{k=1}^{\infty} \frac{(f_1 \underset{D}{*} f_2)(k)}{k} \delta_{\log k}(B) \\
&= \nu_{f_1 \underset{D}{*} f_2}(B).
\end{aligned}$$

Por otro lado, el Teorema 2.2.1 nos asegura que la identidad en G es la función $(1 \underset{D}{*} \mu)$. Esta función es igual a 1 en $n = 1$ y 0 en los demás enteros. Luego, de la definición (5.1) se sigue que $\nu_{(1 \underset{D}{*} \mu)}(B)$ es 1 si $0 \in B$ y 0 en otro caso. Por último, de (5.3) se ve que $\nu * \nu_{(1 \underset{D}{*} \mu)} = \nu$ para toda ν medida de Radon en $[0,\infty)$. En otras palabras, las identidades de ambas estructuras están asociadas. □

5.2. Teoremas de Convergencia

Para usar la convergencia en la topología débil* introducimos el concepto de *medida desfasada*. Dada una medida de Radon ν en $\mathbb{R}$ y algún $h \in \mathbb{R}$ la medida desfasada $\tau_h\nu$ se define como

$$\int_{\mathbb{R}} f \, d\tau_h\nu = \int_{\mathbb{R}} f(t-h) d\nu(t).$$

Antes de continuar, recordemos que

$$\Lambda(n) = \begin{cases} \log p, & \text{si } n = p^k \text{ para un primo } p \text{ y algún entero positivo } k. \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

y que $\Lambda_2(n) = (\mu \underset{D}{*} \log^2)(n)$. La Ecuación (5.1) nos permite considerar una medida de Radon ν_{sym} relacionada a Λ y Λ_2, que por consiguiente posee una naturaleza teórico-númerica relevante para TNP.

El siguiente teorema establece que podemos utilizar una sucesión de medidas desfasadas $\tau_h\nu_{sym}$ para "aproximarnos" a un múltiplo de la medida de Lebesgue μ_ℓ. Esto debe interpretarse formalmente de la siguiente manera: por la discusión al inicio de este capítulo, podemos asociar a cada una de estas medidas un elemento en $C_c(\mathbb{R})^*$, entonces la sucesión de elementos asociados a las medidas desfasadas converge en la topología débil* al elemento asociado al múltiplo de μ_ℓ.

Proposición 5.2.1. *Definamos $\nu_{sym} := \nu_{\Lambda_2/\log}$. Entonces*

$$d\nu_{sym}(t) = d\nu_\Lambda(t) + \frac{1}{t} d(\nu_\Lambda * \nu_\Lambda)(t) \tag{5.4}$$

y $\tau_h\nu_{sym} \overset{*}{\rightharpoonup} 2\mu_\ell$ conforme $h \to \infty$.

Demostración. La Ecuación (5.4), se demuestra utilizando la Ecuación (2.1), el Teorema 5.1.1 y algunas manipulaciones algebraicas.

Recordemos que por el Teorema 3.2.2, la convergencia $\tau_h\nu_{sym} \overset{*}{\rightharpoonup} 2\mu_\ell$ en la topología débil* es equivalente a que

$$\int f d\tau_h\nu_{sym} \to 2 \int f d\mu_\ell \text{ cuando } h \to \infty, \tag{5.5}$$

para toda $f \in C_c(\mathbb{R})$.

Mostraremos que para toda $f(t) = \chi_I(t)e^t$, donde $I \subset \mathbb{R}$ es un intervalo arbitrario, se cumple la convergencia (5.5). Con esto demostrado, el resultado se sigue del hecho que toda $f \in C_c(\mathbb{R})$ puede aproximarse uniformemente, por arriba y abajo, por combinaciones lineales de funciones de la forma $\chi_I(t)e^t$.

Sea $I = (\log a, \log b]$ con $0 < a < b$ un intervalo arbitrario. Notemos que

$$2\int \chi_I(t)e^t dt = 2(b-a)$$

y como e^t es estrictamente creciente

$$\int \chi_I(t)e^t d\tau_h \nu_{sym} = \sum_{ae^h < n \leq be^h} e^{\log n - h} \frac{\Lambda_2(n)}{n \log n}. \tag{5.6}$$

Utilizaremos la Fórmula de simetría de Selberg (2.7) para ver que la expresión en el lado derecho de (5.6) se aproxima a $2(b-a)$.

Por la Fórmula de Simetría de Selberg, para toda constante $c > 0$ tenemos

$$\sum_{n \leq cx} \Lambda_2(n) = 2cx \log x + o(x \log x),$$

que implica que

$$\frac{1}{x \log x} \sum_{ax < n \leq bx} \Lambda_2(n) = 2(b-a) + o(1) \tag{5.7}$$

conforme x tiende a infinito. Sea $\varepsilon > 0$, supongamos que x es lo suficientemente grande de tal manera que $ax < n \leq bx$ implica

$$1 - \varepsilon \leq \frac{\log x}{\log n} \leq 1 + \varepsilon.$$

Multiplicando por $\frac{\Lambda_2(n)}{\log x}$ y sumando todas las desigualdades para $n \in (ax, bx]$ obtenemos

$$1 - \varepsilon \leq \frac{\frac{1}{x}\sum_{ax < n \leq bx} \frac{\Lambda_2(n)}{\log n}}{\frac{1}{x \log x}\sum_{ax < n \leq bx} \Lambda_2(n)} \leq 1 + \varepsilon.$$

Con esto demostramos que

$$\frac{\frac{1}{x}\sum_{ax < n \leq bx} \frac{\Lambda_2(n)}{\log n}}{\frac{1}{x \log x}\sum_{ax < n \leq bx} \Lambda_2(n)} \to 1 \text{ cuando } x \to \infty.$$

Utilizando la Ecuación (5.7), lo anterior implica que

$$\frac{1}{x}\sum_{ax<n\leq bx}\frac{\Lambda_2(n)}{\log n}=2(b-a)+o(1), \tag{5.8}$$

pero el lado izquierdo de esta expresión es igual a

$$\sum_{ax<n\leq bx} e^{\log n-\log x}\frac{\Lambda_2(n)}{n\log n}$$

y haciendo $h=\log x$, esto se convierte en el lado derecho de la Ecuación (5.6). □

Teorema 5.2.2. *Para toda sucesión (h_n) de números reales, tales que $h_n\to\infty$ cuando $n\to\infty$, existe una subsucesión (h_{n_k}) y una función Borel medible $D\colon\mathbb{R}\to[0,2]$ tal que para cualquier $f\in C_c(\mathbb{R})$ se tiene*

$$\lim_{k\to\infty}\int f\,d\tau_{h_{n_k}}\nu_\Lambda=\int fD\,d\mu_\ell.$$

Demostración. Sea l un entero positivo fijo. Tomamos una sucesión $\tau_{h_n}\nu_\Lambda$ restringida a $[-l,l]$, de nuevo por lo discutido al inicio del capítulo, podemos asociar a cada elemento en esta sucesión con un funcional $\varphi_n\in C[-l,l]^*$. El primer paso es demostrar que estos funcionales son acotados (respecto a la norma de operador). Consideremos la función

$$f_l(x)=\begin{cases}1 & \text{si } |x|\leq l,\\ l+1-|x| & \text{si } |x|\in[l,l+1],\\ 0, & \text{si } |x|\geq l+1.\end{cases}$$

Por la Proposición 5.2.1, conforme $h\to\infty$ sabemos que

$$\tau_h\nu_{sym}([-l,l])\leq 2\int f_l\,d\tau_h\nu_{sym}\to\int f_l(x)dx=2(2l+1). \tag{5.9}$$

Además, para todo $B\in\mathcal{B}$, integrando χ_B y usando la Ecuacion (5.4) implica que $\tau_h\nu_\Lambda(B)\leq\tau_h\nu_{sym}(B)$. Entonces (5.9) y la desigualdad del triangulo para integrales nos garantizan que los funcionales φ_n son acotados. Consideramos entonces la bola más pequeña que contiene a toda la sucesión:

$$B_l:=B_{R_l}^{C[-l,l]^*}$$

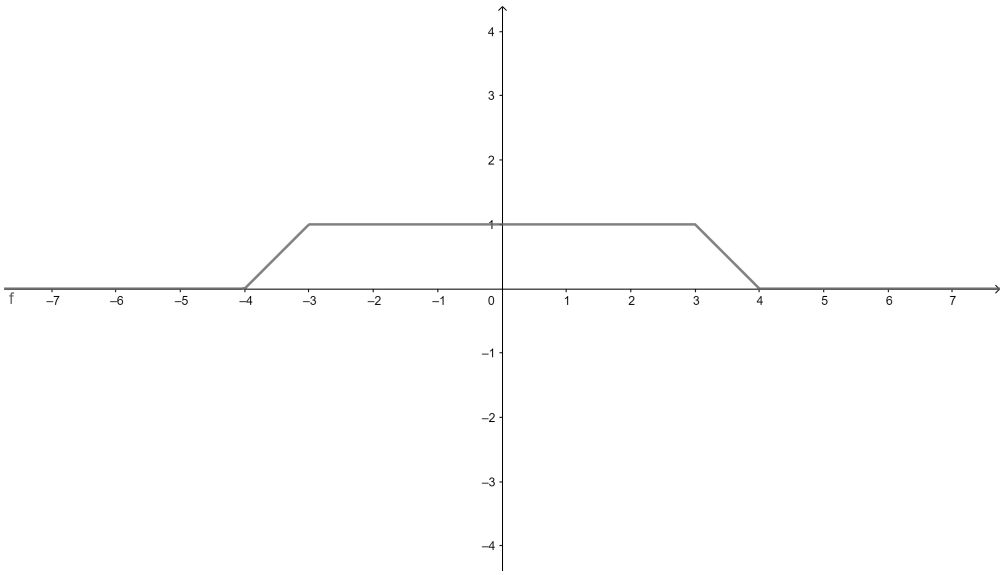

Figura 5.1. f_l con $l = 3$.
Fuente: Elaborado con GeoGebra.

donde $R_l := \sup_{n\geq 1} \tau_{h_n}\nu_\Lambda([-l,l])$. Por el Teorema de Banach-Alouglu-Bourbaki, esta bola es compacta y metrizable en la topología débil*. Entonces existe una subsucesión (φ_{n_k}) y un funcional φ en $C[-l,l]$ tal que $\varphi_{n_k} \xrightarrow{*} \varphi$. En otras palabras, $\varphi_{n_k}(f) \to \varphi(f)$ para todo $f \in C[-l,l]$.

Puesto que l era arbitrario, claramente podemos repetir el argumento con cualquier entero y considerar

$$X = \prod_{l\geq 1} B_l.$$

El espacio X también es compacto y metrizable en la topología producto por el Teorema 1.3.9. Por lo que existe una subsucesión $(\tau_{n_k}\nu_\Lambda)$ tal que $\lim_{k\to\infty} \int f\, d\tau_{n_k}\nu_\Lambda$ converge para todo $f \in C_c(\mathbb{R})$. Como los operadores límite e integral son lineales, entonces esto nos proporciona un funcional lineal en $C_c(\mathbb{R})$. Más aún, este funcional es claramente no negativo si $f \in C_c(\mathbb{R})$ lo es, así que por el Teorema de representación de Riesz 3.3.3 existe una medida de Radon μ en $\mathbb{R}$ tal que

$$\lim_{k\to\infty} \int f\, d\tau_{n_k}\nu_\Lambda = \int f\, d\mu,$$

para todo $f \in C_c(\mathbb{R})$. Además, si $f \geq 0$,

$$\int f\, d\mu = \lim_{k\to\infty} \int f\, d\tau_{n_k}\nu_\Lambda \leq \lim_{k\to\infty} \int f\, d\tau_{n_k}\nu_{sym} = 2\int f\, d\mu_\ell.$$

Por la densidad de $C_c(\mathbb{R})$ en $L_{1,\mu+\mu_\ell}(\mathbb{R})$ (Teorema 3.3.1) para todo conjunto B borel medible acotado, se puede aproximar χ_B utilizando funciones no negativas de $C_c(\mathbb{R})$ tanto para μ como para μ_ℓ, lo cual implica $\mu(B) \leq 2\mu_\ell(B)$ y esto a su vez que $\mu \ll \mu_\ell$. Así que por el Teorema 1.7.1 la derivada de Radon-Nikodym $D = \frac{d\mu}{d\mu_\ell}$ existe y

$$\mu(B) = \int_B D d\mu_\ell \leq 2 \int_B d\mu_\ell$$

por lo que D toma valores en $[0, 2]$ μ_ℓ-casi en todas partes. □

6. Demostración del Teorema

6.1. Primera Reformulación del Teorema

Como se mencionó en la sección 2, la motivación de introducir la función Λ es trabajar con otro conjunto más simple de manipular pero que nos da información útil para TNP. El siguiente teorema exhibe que en efecto es posible hacer una conexión entre Λ y π que puede ser utilizada para demostrar TNP.

Teorema 6.1.1 (Primera Reformulación)**.** *Si*

$$\sum_{1\leq n\leq x} \Lambda(n) \sim x \tag{6.1}$$

entonces TNP se sigue.

Antes de la demostración presentamos unas observaciones que inspiran las ideas claves de la misma. Buscamos que $\lim_{x\to\infty} \frac{\pi(x)\log x}{x} = 1$, esto lo podemos interpretar intuitivamente como

$$\pi(x) \leq \frac{x}{\log x} + E_1(x)$$

$$\pi(x) \geq \frac{x}{\log x} + E_2(x)$$

donde E_1 y E_2 son un «error» que tiende a 0 conforme x crece. Como $\pi(x) \leq x$ para todo x, y $\log x \leq x^\delta$ para todo $\delta > 0$ y x suficientemente grande, podriamos aproximarnos al valor de $\pi(x)$ via potencias de x. Es decir, tomar $\delta > 0$ pequeño y considerar

$$\pi(x^{1-\delta}) \leq x^{1-\delta} \leq \frac{x}{\log x} \qquad \text{para } x \text{ grande.}$$

Entonces el análisis se resume, en términos generales, a estimar $\pi(x) - \pi(x^{1-\delta})$ que esperamos que sea pequeño si δ también lo es y x es grande. Resulta, que este mismo argumento puede utilizarse para establecer la segunda desigualdad.

Demostración. Con lo anterior en mente, comencemos escribiendo

$$\sum_{x^{1-\delta}\leq n\leq x}\Lambda(n)=\sum_{x^{1-\delta}\leq p\leq x}\log p+\sum_{\substack{x^{1-\delta}\leq p^k\leq x\\ k>1}}\log p. \tag{6.2}$$

El primer término en esta ecuación es el de particular interés pues involucra exclusivamente a los primos. De hecho, podemos establecer fácilmente una relación entre este y la cantidad $\pi(x)-\pi(x^{1-\delta})$:

$$(1-\delta)(\pi(x)-\pi(x^{1-\delta}))\log x\leq\sum_{x^{1-\delta}\leq p\leq x}\log p\leq(\pi(x)-\pi(x^{1-\delta}))\log x. \tag{6.3}$$

El segundo término de (6.3) surge como el excedente de considerar un conjunto más grande que los primos. Sin embargo, como comentamos al inicio de esta discusión, esperamos que su aporte asintótico sea negligible.

Para ver esto, sea $\delta>0$ pequeño. Por la hipótesis se tiene que

$$\sum_{1\leq n\leq x^{1-\delta}}\Lambda(n)=x^{1-\delta}+o(x^{1-\delta}),$$
$$\sum_{1\leq n\leq x}\Lambda(n)=x+o(x).$$

Como $x^{1-\delta}=o(x)$ y $o(x)-o(x^{1-\delta})=o(x)$, al restar las dos ecuaciones anteriores obtenemos que

$$\sum_{x^{1-\delta}\leq n\leq x}\Lambda(n)=x+o_\delta(x) \tag{6.4}$$

donde $o_\delta(x)$ representa una función de menor orden que x asociada al δ escogido. Ahora, si $x^{1-\delta}\leq p^k\leq x$ entonces $x^{\frac{1-\delta}{k}}\leq p\leq x^{1/k}$, reordenando la segunda suma de la derecha usando a k como índice

$$\sum_{k}\sum_{x^{\frac{1-\delta}{k}}\leq p\leq x^{1/k}}\log p.$$

Pero como $p=2$ es el primo más pequeño, podemos tomar $x^{1/k}\geq 2$ que implica que $k\leq\frac{\log x}{\log 2}$ y en vista de $k>1$, $\log p\leq\log x^{1/2}$ y

$$\sum_{k=2}^{\lfloor\frac{\log x}{\log 2}\rfloor}\sum_{x^{\frac{1-\delta}{k}}\leq p\leq x^{1/k}}\log p\leq\sum_{k=2}^{\lfloor\frac{\log x}{\log 2}\rfloor}\sum_{x^{\frac{1-\delta}{k}}\leq p\leq x^{1/k}}\log\sqrt{x}$$

$$\leq \log\sqrt{x} \cdot \sum_{k=2}^{\frac{\log x}{\log 2}} \pi(x^{1/k})$$
$$= \log\sqrt{x} \cdot O(x^{1/2}\log x)$$
$$= O(x^{1/2}\log^2 x) = o(x).$$

Entonces por (6.2), (6.3) y (6.4) llegamos a

$$\frac{x}{\log x} + o_\delta\left(\frac{x}{\log x}\right) \leq \pi(x) - \pi(x^{1-\delta}) \leq \frac{x}{(1-\delta)\log x} + o_\delta\left(\frac{x}{\log x}\right).$$

Finalmente, por la discusión previa a la demostración, $\pi(x^{1-\delta}) = o(\frac{x}{\log x})$ y al ser δ arbitrario se obtiene el resultado. □

6.2. Segunda Reformulación

A partir del Teorema 6.1.1, Terence Tao halló una reformulación de TNP en el lenguaje del análisis funcional. Para ello, definió $||\cdot||_\Lambda : C_c(\mathbb{R}) \to \mathbb{R}$ como

$$||f||_\Lambda := \limsup_{h\to\infty} \left| \sum_{n\geq 1} \frac{\Lambda(n)}{n} f(\log n - h) - \int_{\mathbb{R}} f(t)\, dt \right| \tag{6.5}$$

la cual resulta ser una seminorma en $C_c(\mathbb{R})$. Notemos que el primer término en la resta de la definición de $||\cdot||_\Lambda$ está relacionado a la Ecuación (5.1), mientras que el segundo está relacionado a la medida de Lebesgue μ_ℓ porque f es continua (1.1). Además, la Proposición 5.2.1 y el Teorema 5.2.2 sugieren que $||\cdot||_\Lambda$ no diverge.

Teorema 6.2.1. *La función definida en (6.5) es una seminorma en $C_c(\mathbb{R})$, que cumple*

$$||f||_\Lambda \leq ||f||_1, \quad \textit{para } f \in C_c(\mathbb{R}) \tag{6.6}$$

y

$$||f_1 * f_2||_\Lambda \leq ||f_1||_\Lambda ||f_2||_\Lambda, \ \ \textit{para } f_1, f_2 \in C_c(\mathbb{R}). \tag{6.7}$$

Demostración. Gracias a la linealidad de los operadores suma, integral y lím sup, se cumple la propiedad de homogeneidad de los escalares, así como también la desigualdad triangular. La función $||\cdot||_\Lambda$ es claramente no negativa, por lo que lo único a verificar es que es menor que ∞ para todo $f \in C_c(\mathbb{R})$. Para ello, demostraremos la desigualdad (6.6).

Sea $f \in C_c(\mathbb{R})$, tomemos una sucesión $h_n \to \infty$ conforme $n \to \infty$. Entonces

$$||f||_\Lambda = \lim_{n\to\infty} \left| \sum_{n\geq 1} \frac{\Lambda(n)}{n} f(\log n - h_n) - \int_{\mathbb{R}} f \, d\mu_\ell \right|.$$

Aplicando el Teorema 5.2.2, podemos hallar una subsucesión $\tau_{h_{n_k}} \nu_\Lambda$ tal que

$$\lim_{k\to\infty} \int f \, d\tau_{h_{n_k}} \nu_\Lambda = \int fD \, d\mu_\ell,$$

donde D toma valores en $[0, 2]$. Por lo tanto,

$$||f||_\Lambda = \left| \int f(D-1) \, d\mu_\ell \right| \leq \int |f| d\mu_\ell = ||f||_1.$$

Puesto que $C_c(\mathbb{R}) \subset L_1$, esto demuestra que $|| \cdot ||_\Lambda$ en efecto es una seminorma.

Falta demostrar la desigualdad de álgebra Banach

$$\left| \sum_{n\geq 1} \frac{\Lambda(n)}{n} (f_1 * f_2)(\log n - h) - \int_{\mathbb{R}} f \, d\mu_\ell \right| \leq ||f_1||_\Lambda ||f_2||_\Lambda + o_{f_1,f_2}(1) \tag{6.8}$$

conforme $h \to \infty$. Recordemos de (5.1) que la suma de la izquierda en (6.8) representa la integral de $f_1 * f_2$ respecto de la medida $\tau_h \nu_\Lambda$, entonces todo el lado izquierdo puede ser interpretado como la integral de $f_1 * f_2$ respecto de la medida $\tau_h \nu_\Lambda - \mu_\ell$.

Definimos $\mu_{\ell+} := \mu_\ell|_{[0,\infty)}$ y notemos que $\tau_h \mu_{\ell+} \to \mu_\ell$ si $h \to \infty$. Por esta razón podemos ver la integral respecto a $\tau_h \nu_\Lambda - \mu_\ell$ como una integral respecto a τ_h aplicado a $\nu_\Lambda - \mu_{\ell+}$. Además de esto, $\mu_{\ell+}$ cumple la propiedad

$$(\mu_{\ell+} * \mu_{\ell+})(B) = \int_0^\infty t \chi_B(t) dt. \tag{6.9}$$

A continuación se presenta un bosquejo de los pasos a seguir a partir de aquí para demostrar la desigualdad deseada.

Primero, se utiliza la Proposición 5.2.1 para reescribir $\nu_\Lambda - \mu_{\ell+}$ como suma de tres medidas con signo

$$\nu_\Lambda - \mu_{\ell+} = \rho_1 + \rho_2 + \rho_3 \tag{6.10}$$

No presentaremos una definición formal del concepto de medida con signo. Sin embargo, para fines prácticos el lector puede considerarlo como una "medida" capaz de

tomar valores negativos. En particular,

$$\begin{aligned}\rho_1 &= \nu_{sym} - 2\mu_{\ell+}\\ \rho_2 &= -\frac{1}{t}d\big((\nu_\Lambda - \mu_{\ell+}) * (\nu_\Lambda - \mu_{\ell+})\big)(t)\\ \rho_3 &= 2\mu_{\ell+} - 2\frac{1}{t}d(\nu_\Lambda * \mu_{\ell+})(t).\end{aligned}$$

De nuevo por la Proposición 5.2.1 se sigue que

$$\tau_h\rho_1 = \tau_h(\nu_{sym} - 2\mu_{\ell+}) \xrightarrow{*} 2\mu_\ell - 2\mu_\ell = 0. \tag{6.11}$$

conforme $h \to \infty$. Luego, al examinar el término

$$\frac{1}{t}d(\nu_\Lambda * \mu_{\ell+})(t)$$

en ρ_3 notamos que corresponde la medida cuya derivada de Radon-Nikodym respecto a μ_ℓ es

$$\frac{1}{t}\sum_{n\leq e^t}\frac{\Lambda(n)}{n}.$$

Por el Teorema de Merten, (Teorema 6.2.4 demostrado adelante) esta función es $1 + o(1)$ conforme $t \to \infty$. Entonces

$$\tau_h\left(\frac{1}{t}d(\nu_\Lambda * \mu_{\ell+})\right) \xrightarrow{*} \mu_\ell \tag{6.12}$$

y por consiguiente

$$\tau_h\rho_3 \xrightarrow{*} 2\mu_\ell - 2\mu_\ell = 0 \tag{6.13}$$

conforme $h \to \infty$.

Gracias al comportamiento asintótico de $\tau_h\rho_1$ y $\tau_h\rho_3$ descrito en (6.11) y (6.13) para obtener la desigualdad (6.8) basta mostrar que

$$\left|\int f_1 * f_2 d\tau_h\rho_2\right| \leq ||f_1||_\Lambda||f_2||_\Lambda + o(1) \tag{6.14}$$

Para esto, se utiliza de nuevo la Proposición 5.2.1, Fubini extendido a medidas con signo, argumentos de compacidad y (6.12). El uso explícito de estos elementos puede hallarse a detalle en la sección 14.2.6 de [8]. □

Teorema 6.2.2 (Segunda Reformulación)**.** *Si* $||f||_\Lambda = 0$ *para todo* $f \in C_c(\mathbb{R})$ *entonces TNP es cierto.*

Demostración. Si la seminorma es cero y sustituimos $h = \log x$ en la Ecuación (6.5) entonces

$$\sum_{n\geq 1} \frac{\Lambda(n)}{n} f\left(\log \frac{n}{x}\right) = \int_{\mathbb{R}} f(t)\, dt + o(1)$$

conforme $x \to \infty$ para todo $f \in C_c(\mathbb{R})$. Si escogemos $g \in C_c((0,\infty))$ y definimos $f(t) := e^t g(e^t)$ entonces $f \in C_c(\mathbb{R})$ y la ecuación de arriba se convierte en

$$\sum_{n\geq 1} \frac{\Lambda(n)}{n} \frac{n}{x} g\left(\frac{n}{x}\right) = \int_{\mathbb{R}} g(e^t) e^t\, dt + o(1)$$

y si $u = e^t$ en el lado derecho,

$$\sum_{n\geq 1} \Lambda(n) g\left(\frac{n}{x}\right) = x \int_0^\infty g(u)\, du + o(x).$$

Sea $\delta > 0$, es posible elegir dos funciones de soporte compacto g_+ y g_- que cumplan $0 \leq g_- \leq \chi_{[\frac{1}{2},1]} \leq g_+ \leq 1$ y $\int (g_+ - g_-)dx < \delta$ como podemos apreciar de la figura 6.1.

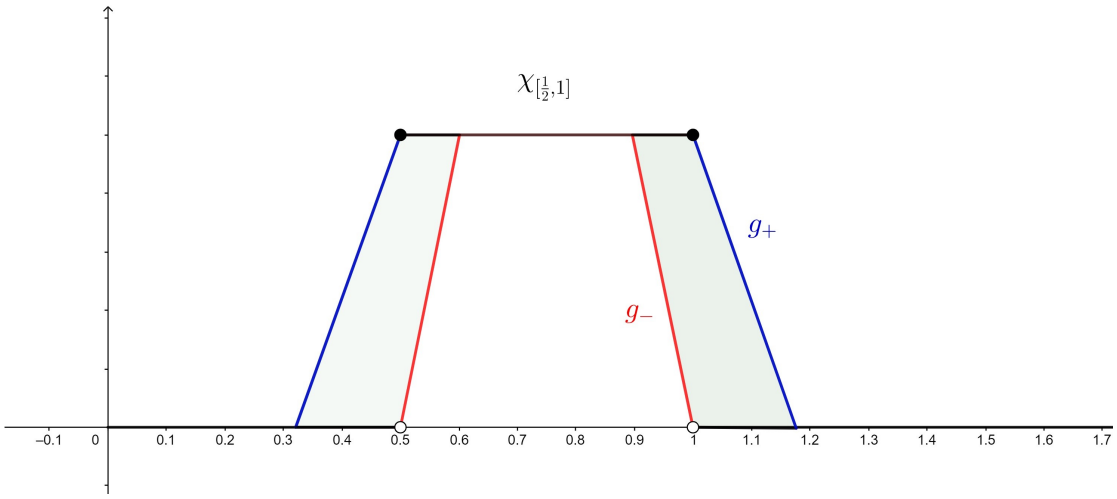

Figura 6.1. Funciones g_+ y g_-
Fuente: Elaborado con GeoGebra.

Aplicando el argumento anterior a g_+ obtenemos

$$\sum_{n\geq 1} \Lambda(n) g_+\left(\frac{n}{x}\right) = x \int_0^\infty g_+(u)\, du + o_\delta(x)$$

$$\begin{aligned}
&\leq x\left(\delta + \int_0^\infty g_-(u)du\right) + o_\delta(x)\\
&\leq x\left(\delta + \int_0^\infty \chi_{[\frac{1}{2},1]}\, du\right) + o_\delta(x)\\
&= x\left(\delta + \frac{1}{2}\right) + o_\delta(x).
\end{aligned}$$

Además,

$$\sum_{n\geq 1}\Lambda(n)\chi_{[\frac{1}{2},1]}\left(\frac{n}{x}\right) \leq \sum_{n\geq 1}\Lambda(n)g_+\left(\frac{n}{x}\right)$$

que implica

$$\sum_{\frac{x}{2}\leq n\leq x}\Lambda(n) \leq \sum_{n\geq 1}\Lambda(n)g_+\left(\frac{n}{x}\right) \leq x\left(\delta + \frac{1}{2}\right) + o_\delta(x).$$

Análogamente

$$x\left(\frac{1}{2} - \delta\right) + o_\delta(x) \leq \sum_{\frac{x}{2}\leq n\leq x}\Lambda(n),$$

pero al ser δ arbitrario se llega a

$$\sum_{\frac{x}{2}\leq n\leq x}\Lambda(n) = \frac{x}{2} + o(x). \tag{6.15}$$

A continuación, separaremos la suma $\sum_{n\leq x}\Lambda(n)$ por intervalos cuyos extremos son de la forma $\frac{x}{2^j}$ y cada una de estas subsumas la estimaremos usando (6.15).

$$\begin{aligned}
\sum_{n\leq x}\Lambda(n) &= \sum_{\frac{x}{2}<n\leq x}\Lambda(n) + \sum_{\frac{x}{2^2}<n\leq \frac{x}{2}}\Lambda(n) + \cdots + \sum_{\frac{x}{2^{l+1}}<n\leq \frac{x}{2^l}}\Lambda(n) + \sum_{1\leq n\leq \frac{x}{2^{l+1}}}\Lambda(n)\\
&= \sum_{j=1}^{l+1}\frac{x}{2^j} + \sum_{j=0}^{l} o\left(\frac{x}{2^j}\right) + \sum_{1\leq n\leq \frac{x}{2^{l+1}}}\Lambda(n).
\end{aligned}$$

Para estimar la segunda y tercera suma, fijemos $\varepsilon > 0$, para este, existe una constante $M_\varepsilon \geq 1$ tal que el término de error en $\sum_{\frac{y}{2}\leq n\leq y}\Lambda(n)$ es menor en valor absoluto que εy para todo $y \geq M_\varepsilon$. Si escogemos l lo mayor posible tal que $\frac{x}{2^l} \geq M_\varepsilon$, entonces

$$\sum_{n\leq x}\Lambda(n) \leq x\left(1 - \frac{1}{2^{l+1}}\right) + \sum_{j=0}^{l} o\left(\frac{x}{2^j}\right) + \sum_{1\leq n\leq M_\varepsilon}\Lambda(n)$$

$$\leq x + M + 2\varepsilon x$$
$$= x + 2\varepsilon x + o_\varepsilon(x).$$

Al ser ε arbitrario, se llega a la Ecuación (6.1). □

A partir de aquí, suponemos que $||\cdot||_\Lambda$ es una seminorma no trivial. Esto nos permitirá construir caracteres con propiedades específicas, que inevitablemente implicarán una contradicción. Para comenzar, notemos que si la seminorma es no trivial, podemos considerar el cociente de $C_c(\mathbb{R})$ por el kernel de $||\cdot||_\Lambda$ para obtener un espacio normado. Luego, podemos completar este para obtener el espacio de Banach $\mathcal{A}_\Lambda$. Además, la continuidad de la convolución como multiplicación en el álgebra normada permite que se extienda a $\mathcal{A}_\Lambda$. Finalmente, si consideramos el mapeo canónico de $C_c(\mathbb{R})$ a $\mathcal{A}_\Lambda$, por la desigualdad del Teorema 6.2.1, este se extiende a un homorfismo $\Phi\colon L_1(\mathbb{R}) \to \mathcal{A}_\Lambda$.

Usando Φ y la transformada de Fourier $f \mapsto \widehat{f}$ construiremos un caracter no trivial de $C_c(\mathbb{R})$ que nos garantizará una condición de igualdad sobre $||\cdot||_\Lambda$ y la norma usual en $L_1(\mathbb{R})$.

Teorema 6.2.3 (Teorema del espectro no trivial)**.** *Supongamos que $||\cdot||_\Lambda$ es no trivial en $C_c(\mathbb{R})$, de modo que la completación $\mathcal{A}_\Lambda$, asociada a este espacio, es no trivial. Entonces existe un $\xi \in \mathbb{R}$, tal que el mapeo $f \mapsto \widehat{f}(\xi)$ con $f \in C_c(\mathbb{R})$, es continuo respecto a $||\cdot||_\Lambda$. Además, dado tal ξ, si $f(t)e^{-2\pi it\xi} \geq 0$ para todo $t \in \mathbb{R}$, entonces $||f||_\Lambda = ||f||_1$.*

Demostración. Notemos que a partir de todo caracter $\psi\colon \mathcal{A}_\Lambda \to \mathbb{C}$, se puede construir el caracter $\psi \circ \Phi\colon L_1(\mathbb{R}) \to \mathbb{C}$. Si ψ es no trivial, entonces $\psi \circ \Phi$ tampoco lo es. Luego por el Teorema 4.2.4 existe $\xi \in \mathbb{R}$ tal que $\psi \circ \Phi$ toma la forma $f \mapsto \widehat{f}(\xi)$, obviamente continuo respecto a la norma. Por lo que basta con demostrar que existe un caracter no trivial en $\mathcal{A}_\Lambda$.

Por la fórmula del radio espectral (Teorema 3.4.3) y la fórmula de Gelfand (Corolario 3.5.9), la existencia de un caracter así se sigue si existe un elemento $f \in \mathcal{A}_\Lambda$ cuyo radio espectral es no nulo.

Tomemos $g_0 \in C_c(\mathbb{R})$ que cumpla $||g_0||_\Lambda > 0$. Por la densidad de $\mathscr{S}(\mathbb{R})$ en $L_1(\mathbb{R})$ en la norma $||\cdot||_1$ podemos hallar un elemento $g_1 \in \mathscr{S}(\mathbb{R})$ tal que

$$||g_0 - g_1||_1 < ||g_0||_\Lambda. \tag{6.16}$$

Como $||f||_\Lambda \leq ||f||_1$ para todo $f \in C_c(\mathbb{R})$, obtenemos

$$||g_0||_\Lambda \leq ||g_0 - g_1||_\Lambda + ||g_1||_\Lambda \leq ||g_0 - g_1||_1 + ||g_1||_\Lambda$$

y de esto

$$0 < ||g_0||_\Lambda - ||g_0 - g_1||_1 \leq ||g_1||_\Lambda.$$

De hecho, como veremos a continuación, podemos suponer $\widehat{g_1} \in C_c^\infty(\mathbb{R})$. Recordemos del Teorema 4.1.4 que la transformada inversa de Fourier definida en (4.2) es un mapeo continuo. Como $C_c^\infty(\mathbb{R})$ es denso en $L_1(\mathbb{R})$, el Corolario 1.3.6 implica que $\widetilde{C_c^\infty(\mathbb{R})}$ es denso en $\widetilde{L_1(\mathbb{R})}$. Por el Teorema 4.1.4 sabemos que $\widetilde{\mathscr{S}(\mathbb{R})} = \mathscr{S}(\mathbb{R})$ y por lo tanto $\mathscr{S}(\mathbb{R}) \subset \widetilde{L_1(\mathbb{R})}$. Así que $\widetilde{C_c^\infty(\mathbb{R})}$ es denso en $\mathscr{S}(\mathbb{R})$ y por transitividad también es denso en $L_1(\mathbb{R})$. Es decir, podemos escoger a $g_1 \in \widetilde{C_c^\infty(\mathbb{R})}$ que cumpla (6.16) y por el Teorema 4.1.2 en efecto $\widehat{g_1} \in C_c^\infty(\mathbb{R})$.

Ahora, como $\text{supp}(\widehat{g_1})$ es compacto existe un intervalo I tal que $\text{supp}(\widehat{g_1}) \subset I$. Luego, podemos hallar una función $h \in \mathscr{S}(\mathbb{R})$ tal que $h \equiv 1$ en $\text{supp}(\widehat{g_1})$ utilizando el argumento en [10] y del Teorema 4.1.4 sabemos que $\check{h} \in \mathscr{S}(\mathbb{R})$. Llamemos $f = \check{h}$, de modo que por propiedades de la transformada de Fourier (Teorema 4.1.1) tenemos

$$\widehat{g_1 * f^{*n}} = \widehat{g_1}\widehat{f^n} = \widehat{g_1}$$

y por el Teorema de inversión de Fourier (Teorema 4.1.2) $g_1 * f^{*n} = g_1$. Esto indica que f se comporta como el elemento identidad para g_1. Aplicando Φ,

$$0 < ||g_1||_\Lambda = ||g_1 * f^{*n}||_\Lambda \leq ||g_1||_\Lambda ||f^{*n}||_\Lambda$$

que implica $||f^{*n}||_\Lambda \geq 1$ para todo $n \geq 1$, así que f tiene un espectro no trivial, y por lo tanto, existe un caracter ψ no trivial.

Finalmente, supongamos que para el ξ asociado a ψ, existe una función $f \in C_c(\mathbb{R})$ para la cual $f(t)e^{-2\pi i t\xi}$ es no negativa. Entonces,

$$||f||_1 = \int |f(t)|dt = \int f(t)e^{-2\pi i t\xi} = |\widehat{f}(\xi)| \leq ||\psi||\, ||f||_\Lambda \leq ||f||_\Lambda. \qquad \square$$

La parte más esencial del teorema anterior para la prueba de TNP es la desigualdad $||f||_1 \leq ||f||_\Lambda$. Esto porque veremos que la desigualdad estricta $||f||_1 > ||f||_\Lambda$ también debe cumplirse y esto genera la contradicción necesaria. Implica que la seminorma $||\cdot||_\Lambda$ es cero para todo $f \in C_c(\mathbb{R})$ y por la reformulación de Tao, se

obtiene TNP. Antes de esto, enunciamos y demostramos el Teorema de Merten.

Teorema 6.2.4 (Teorema de Merten)**.** *Para $x \geq 1$,*

$$\sum_{n \leq x} \frac{\Lambda(n)}{n} = \log x + O(1).$$

Demostración. Gracias a la ecuación (6.6) del Teorema 6.2.1 tenemos $||f||_\Lambda \leq ||f||_1$, que implica

$$\limsup_{x \to \infty} \left| \sum_{n \geq 1} \frac{\Lambda(n)}{n} f\Big(\log \frac{n}{x}\Big) \right| \leq 2||f||_1,$$

para todo $f \in C_c\big((0,\infty)\big)$. Si utilizamos esta desigualdad para una función $f(t) = e^t g_+(e^t)$, donde $g_+ \in C_c(\mathbb{R})$ es un función no negativa con $\chi_{[\frac{1}{2},1]} \leq g_+$, obtenemos, por argumentos similares a los empleados para demostrar el Teorema 6.2.2, que

$$\limsup_{x \to \infty} \frac{1}{x} \sum_{\frac{x}{2} < n \leq x} \Lambda(n) \leq 2 \int_0^\infty g_+(t) dt,$$

con lo cual se llega a

$$\sum_{\frac{x}{2} < n \leq x} \Lambda(n) \leq Cx$$

para algún $C \geq 1$ y todo $x \geq 1$. Esto permite estimar la suma por pedazos,

$$\begin{aligned} \sum_{n \leq x} \Lambda(n) &= \sum_{\frac{x}{2} < n \leq x} \Lambda(n) + \sum_{\frac{x}{4} < n \leq \frac{x}{2}} \Lambda(n) + \cdots + \sum_{\frac{x}{2^{k+1}} < n \leq \frac{x}{2^k}} \Lambda(n) \\ &\leq Cx + \frac{C}{2}x + \cdots + \frac{C}{2^k}x \leq 2Cx, \end{aligned}$$

donde $k = \lfloor \log_2 x \rfloor$, con lo que concluimos que

$$\sum_{n \leq x} \Lambda(n) = O(x). \tag{6.17}$$

Recordemos que $\log n = \sum_{d|n} \Lambda(d)$; a partir de esto

$$\sum_{n \leq x} \log n = \sum_{n \leq x} \sum_{d|n} \Lambda(d) = \sum_{d \leq x} \Lambda(d) \Big\lfloor \frac{x}{d} \Big\rfloor.$$

Por un lado, $\lfloor x \rfloor = x + O(1)$, así que (6.17) implica

$$\sum_{n \leq x} \log n = x \sum_{d \leq x} \frac{\Lambda(d)}{d} + O(x). \tag{6.18}$$

Por otro, aplicando la desigualdad 2.5 a $f(t) = \log t$

$$\sum_{n \leq x} \log n = \int_1^x \log t \, dt + O\big(\log(1+x)\big) = x \log x + O(x). \tag{6.19}$$

Al igualar las expresiones de (6.17) y (6.18) y dividir por x se obtiene el resultado. □

Teorema 6.2.5 (Teorema del Espectro Trivial)**.** *Para todo $\xi \in \mathbb{R}$, existe $f \in C_c(\mathbb{R})$ tal que $f(t)e^{-2\pi i t\xi} \geq 0$ para todo $t \in \mathbb{R}$ y $||f||_\Lambda < ||f||_1$.*

Demostración. Caso 1 ($\xi \neq 0$): Consideremos la función (ver figura 6.1)

$$g(t) = \begin{cases} 1 & \text{si } |t| \leq 1, \\ 2 - |t| & \text{si } |t| \in [1,2], \\ 0 & \text{en cualquier otro caso.} \end{cases}$$

Para un $\xi \neq 0$ fijo, definimos $f(t) = g(t)e^{2\pi i t\xi}$ que también está en $C_c(\mathbb{R})$. De la definición de $||f||_\Lambda$, podemos tomar una sucesión de números reales (h_n) tal que $h_n \to \infty$ y por lo tanto

$$||f||_\Lambda = \lim_{n\to\infty} \left| \int f \, d\tau_{h_n}\nu_\Lambda - \int f \, d\mu_\ell \right|.$$

Por el Teorema 5.2.2, podemos hallar una función Borel medible $D : \mathbb{R} \to [0,2]$ que cumple

$$||f||_\Lambda = \left| \int (D-1) f \, d\mu_\ell \right|.$$

Sea $\theta \in \mathbb{R}$ tal que

$$||f||_\Lambda = e^{i\theta} \int (D-1) f \, d\mu_\ell = \int_{-2}^{2} \mathrm{Re}(e^{i\theta} e^{2\pi i t\xi}) g(t)(D(t)-1) dt.$$

Notemos que $|D(t) - 1| \leq 1$ y $\mathrm{Re}(e^{i\theta}e^{2\pi i t\xi}) = \cos(\theta + 2\pi t\xi) < 1$ casi en todas partes.

De modo que

$$||f||_\Lambda < \int_{-2}^{2} g(t)dt = ||g||_1 = ||f||_1.$$

Caso 2 ($\xi = 0$): Consideremos la función

$$f(t) = \begin{cases} 1, & \text{si } t \in [-N, 0], \\ 1 - t, & \text{si } t \in [0, 1], \\ N + 1 + t, & \text{si } t \in [-(N+1), -N], \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

donde N es un entero a determinar luego (comparar con figura 5.1). Por el Teorema de Merten, para $h \geq N + 1$ se tiene

$$\begin{aligned} \sum_{n\geq 1} \frac{\Lambda(n)}{n} f(\log n - h) &\leq \sum_{e^{h-(N+1)} \leq n \leq e^{h+1}} \frac{\Lambda(n)}{n} \\ &= \log e^{h+1} - \log e^{h-(N+1)} + O(1) \\ &= N + O(1). \end{aligned}$$

De manera similar, para $h \geq N$ se tiene

$$\begin{aligned} \sum_{n\geq 1} \frac{\Lambda(n)}{n} f(\log n - h) &\geq \sum_{e^{h-N} \leq n \leq e^{h}} \frac{\Lambda(n)}{n} \\ &= \log e^{h} - \log e^{h-N} + O(1) \\ &= N + O(1). \end{aligned}$$

Al tomar N suficientemente grande y usar el hecho $\int f(t)dt = N + 1$ hallamos

$$\left| \sum_{n\geq 1} \frac{\Lambda(n)}{n} f\left(\log \frac{n}{h} \right) - \int f(t)dt \right| < N < ||f||_1$$

para todo h suficientemente grande. Entonces $||f||_\Lambda < ||f||_1$. □

CONCLUSIONES

1. Las funciones $\pi(x)$ y $\frac{x}{\log x}$ son asintóticamente equivalentes.

2. La contradicción generada por los Teoremas 6.2.3 y 6.2.5 es el argumento fundamental en esta demostración del Teorema de los Números Primos y surge como consecuencia de que en el espacio $C_c(\mathbb{R})$ abundan elementos que se comportan localmente como la unidad.

3. Las herramientas desarrolladas de análisis funcional permitieron varios argumentos de «análisis suave», que hicieron de las demostraciones de resultados como el Teorema 6.2.3 y el Teorema de Merten, directas y de pocos cálculos. En el caso del Teorema de Merten, comparar esta demostración con alguna de sus demostraciones «elementales» (ver teorema 1.10 de [1]) ejemplifica lo conveniente que puede ser emplear herramientas de este tipo.

RECOMENDACIONES

1. Para un examen detallado de las demostraciones más técnicas de este trabajo es conveniente un estudio más a fondo de los elementos básicos y teoremas fundamentales de Teoría de la Medida presentados en el capítulo 1.

2. El blog de Terence Tao [14] posee una gran variedad de entradas, algunas de las cuales ilustran resultados que pueden ser explorados como parte de una tesis de licenciatura.

3. En el presente documento se utilizaron algunos resultados usados por Selberg en su demostración elemental del Teorema de los Números Primos. Muchos de estos podrían ser agregados al contenido del curso de Teoría de Números de la licenciatura en matemática aplicada de la Universidad de San Carlos.

BIBLIOGRAFÍA

[1] E. Antley. *Towards the Prime Number Theorem.* s.e. 2015.

[2] S. Axler. *Measure, Integration & Real Analysis.* Springer, 2020.

[3] R. Bartle. *The Elements of Integration and Lebesgue Measure.* Wiley, Nueva York, 1995.

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

[5] J. Cerda. *Linear functional analysis.* American Mathematical Society; Real Sociedad Matemática Española, Providence: Madrid, 2010.

[6] K. Chung y K. Zhong. *A course in probability theory.* Academic press, Londres, 2001.

[7] M. Curtins y G. Anand. The importance of numbers and the need to study primes: The prime questions. *Journal of Mathematics and Statistics*, **7**(4):262–269, 2011.

[8] M. Einsiedler y T. Ward. *Functional analysis, spectral theory and applications.* Springer, 2017.

[9] J. Escamilla-Castillo. *Topología.* 2.ª ed. s.e., Guatemala, 1992.

[10] D. Giraudo. Compactly supported infinitely differentiable function constant on an interval. Consultado en junio 2021 en `https://math.stackexchange.com/questions/1088674`

[11] A. Hildebrand. *Short Course on Asymptotics.* s.e.

[12] T. Koshy. *Elementary Number Theroy with applications.* Academic Press, 2002.

[13] E. Kreyszig. *Introductory functional analysis with applications, (volumen 1)* Wiley, New York, 1978.

[14] T. Tao. A Banach algebra proof of the prime number theorem. Consultado en julio 2020 en `https://terrytao.wordpress.com/2014/10/25/a-banach-algebra-proof-of-the-prime-number-theorem/`.