UNA DEMOSTRACIÓN ELEMENTAL DEL TEOREMA DE LOS NÚMEROS PRIMOS

Larry Antonio Flores Luna

UNIVERSIDAD DE CÓRDOBA
FACULTAD DE CIENCIAS BÁSICAS
DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA
MONTERÍA
2022

UNA DEMOSTRACIÓN ELEMENTAL DEL TEOREMA DE LOS NÚMEROS PRIMOS

Larry Antonio Flores Luna

Trabajo presentado como requisito parcial para optar al título de

Matemático

Asesor:

Ph. D. Jerson Manuel Borja Soto

UNIVERSIDAD DE CÓRDOBA
FACULTAD DE CIENCIAS BÁSICAS
DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA
MONTERÍA
2022

UNIVERSIDAD DE CÓRDOBA
FACULTAD DE CIENCIAS BÁSICAS
DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA

Los jurados abajo firmantes certifican que han leído y que aprueban el trabajo de grado titulado: **UNA DEMOSTRACIÓN ELEMENTAL DEL TEOREMA DE LOS NÚMEROS PRIMOS**, el cual es presentado por el estudiante **Larry Antonio Flores Luna**.

Fecha: Febrero de 2022

Asesor: _____
Ph. D. Jerson Manuel Borja Soto

Jurado: _____
M.Sc. Sergio Miguel Avilez Ortiz

Jurado: _____ 25.03.22.
Ph. D. Luis Enrique Benítez Babilonia

Roberto Flores
Y Cledys Luna

Resumen

Dado un número real positivo x, la cantidad de números primos menores o iguales que x se denota $\pi(x)$. En este trabajo presentaremos una demostración elemental del conocido *teorema de los números primos*, que establece que la cantidad $\pi(x)$ es asintóticamente igual al cociente $x/\ln x$ cuando $x \to \infty$. Para lograr tal demostración, haremos uso de técnicas elementales de la teoría analítica de números para probar la llamada *identidad de Selberg*, a partir de la cual se derivará dicha demostración elemental.

Abstract

Given a real positive number x, the quantity of prime numbers less than or equal to x is denoted by $\pi(x)$. In this work, we will present an elementary proof of famous *prime number theorem*, which asserts that the quantity $\pi(x)$ is asymptotically equivalent to the quotient $x/\ln x$ as $x \to \infty$. To do this demonstration, we will use elementary techniques of analytic number theory to demonstrate *Selberg's asymptotic formula*, from which we will derive the elementary proof of the prime number theorem.

Agradecimientos

Quiero agradecer y rendir un homenaje a mi madre Cledys del Rosario Luna González, que aunque no esté a mi lado ha sido mi apoyo espiritual, me ha llenado de fuerza en los momentos más difíciles, por siempre aconsejarme estudiar una carrera profesional y por su infinito amor. Seguidamente, aunque no de forma menos importante, le agradezco a mi padre Roberto Antonio Flores Rada, por el apoyo incondicional que me ha dado a lo largo de la carrera y a lo largo de mi vida, es la persona que más admiro y por la cual me he esforzado para ver siempre una sonrisa en su rostro. No puede faltar mi hermana, Sindy Flores, por sus consejos y por ayudarme a entender cuando hago las cosas mal, a mi novia Maria Florez, por apoyarme moralmente y ayudarme a levantar cada vez que caigo. A la Universidad de Córdoba, por brindarme una educación de calidad y en particular al Departamento de Matemáticas y Estadística, y sus profesores, por haberme impartido tan valiosos conocimientos, los cuales han ayudado a formarme en esta maravillosa área del saber, como son las matemáticas. Finalmente quiero agradecer a mi asesor de tesis Jerson Borja, pues me ha enseñado mucho a lo largo de este trabajo y de la carrera, por su paciencia y dedicación en este trabajo y por ser un ejemplo a seguir como matemático; no pueden faltar mis compañeros de estudio, amigos como Mario Viloria, Jean Carlos Sanches, Jean Carlos Otero, Duban Benítez y Jesús Javier Soto, quienes han significado mucho para mí y a mis demás compañeros.

Montería, Colombia — Larry Antonio Flores Luna

Febrero de 2022

Índice general

Resumen ***iv***

Abstract ***v***

Introducción **1**

1. Algunas funciones aritméticas y el producto de Dirichlet **4**

1.1. La función μ de Möbius 4

1.2. Producto de Dirichlet 6

1.3. Inversa de Dirichlet y la fórmula de inversión de Möbius 8

1.4. La Función de Von Mangoldt 10

2. Sumas parciales de funciones aritméticas **12**

2.1. La notación O grande 13

2.2. Fórmula de sumación de Euler 16

2.3. Sumas parciales de un producto de Dirichlet 18

2.4. Sumas parciales que involucran a la función Λ; la función ψ de Chebyshev 21

3. El teorema de los números primos **25**

3.1. El teorema de los números primos y una forma equivalente con la función ψ 25

3.2. La fórmula asintótica de Selberg 29

3.3. Estrategia de la demostración 32

3.4. Demostración del teorema de los números primos 33

3.4.1. La función $S(y)$ 33

3.4.2. Fórmulas asintóticas y desigualdades que involucran a la función $S(y)$ 36
3.4.3. Reescalamiento logarítmico 43
3.4.4. Demostración final 48

Bibliografía **53**

Introducción

Uno de los problemas fundamentales en la teoría de números es comprender cómo están distribuidos los números primos en el conjunto de los números naturales. Los números primos son aquellos enteros positivos que solo son divisibles por 1 y por sí mismos (como lo son $2, 3, 5, 7$, etc), y estos, aparentemente, no presentan una regla que permita predecir, dado un primo p_n, el siguiente primo p_{n+1}. No obstante, en el año 1737 el gran matemático *Leonhard Euler* descubrió una maravillosa identidad capaz de relacionar los números naturales con los números primos mediante la siguiente fórmula:

$$\sum_{n=1}^{\infty} n^{-s} = \prod_{p} (1 - 1/p^s)^{-1},$$

para todo $s > 1$, donde el producto es tomado sobre todos los primos p. Luego, a finales del siglo XVIII, *Carl Friedrich Gauss* y *Adrien-Marie Legendre* conjeturaron que dado un número $x > 0$, la cantidad de números primos menores o iguales a x, debía aproximarse asintóticamente a $x/\ln x$. Cabe resaltar que *Gauss* se dió cuenta de este comportamiento utilizando una inmensa tabla de números primos que él mismo había construido, con una precisión que resulta desconcertante para una época previa a los ordenadores.

En la teoría de números, se denota la cantidad de números primos menores o iguales a x por $\pi(x)$. Así, en lenguaje moderno, la conjetura de Gauss y Legendre establece que

$$\lim_{x\to\infty} \frac{\pi(x)\ln x}{x} = 1.$$

La demostración formal de esta conjetura la hicieron por primera vez, y de forma independiente, tanto *Jacques Hadamard* en su artículo de 1896, *Sur la distribution des*

zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques [2], como *Charles-Jean de la Vallée Poussin* en un artículo de ese mismo año, titulado *Recherches analytiques sur la théorie des nombres premiers* [6]. Este resultado es hoy conocido como el *Teorema de los números primos*.

Las demostraciones de Hadamard y de la Vallée Poussin no son en absoluto pruebas elementales. Ambas utilizan la teoría de funciones de variable compleja, siguiendo el camino trazado por Bernhard Riemann, quien descubrió que el teorema de los números primos está íntimamente relacionado con el comportamiento de la función zeta en el plano complejo, y encontró muchas propiedades de esta función que hoy lleva su nombre, y hace parte del que tal vez sea el problema sin resolver más importante de las matemáticas: *la hipótesis de Riemann*. Las ideas de Riemann fueron explotadas e implementadas en las pruebas de Hadamard y de la Vallée Poussin.

Eventualmente, muchos matemáticos trataron de encontrar una prueba elemental del teorema de los números primos; (*elemental en el sentido de no tener que usar la teoría de funciones complejas*), pero no tuvieron éxito, hasta que, en 1949, Atle Selberg encontró tal demostración elemental, siguiendo el camino trazado por el ruso Pafnuty Lvóvich Chebyshev, quien había estudiado el problema a través de las propiedades del logaritmo, en lugar de funciones de variable compleja. Selberg publicó su demostración en un artículo de 1949 titulado *An elementary proof of the prime-number theorem*, en la prestigiosa *Annals of Mathematics*. Esta demostración no hace uso de la teoría de funciones complejas, y consiste en probar una forma equivalente al teorema de los números primos:

$$\lim_{x\to\infty} \frac{\psi(x)}{x} = 1,$$

donde $\psi(x)$ es la función de *Chebyshev* dada por

$$\psi(x) = \sum_{n\leq x} \Lambda(n),$$

para $x > 0$, y n recorre todos los enteros positivos menores o iguales que x y $\Lambda(n)$

representa la función de Von Mangoldt definida por

$$\Lambda(n) = \begin{cases} \ln p, & \text{si } n = p^k, \text{ donde } p \text{ es primo y } k \geq 1; \\ 0, & \text{en otro caso.} \end{cases}$$

En esta demostración, *Selberg* introduce la fórmula asintótica

$$\psi(x) \ln x + \sum_{n \leq x} \Lambda(n) \psi\left(\frac{x}{n}\right) = 2x \ln x + O(x),$$

la cual fue probada por él mismo y es el punto de partida para hallar una demostración elemental del teorema de los números primos. Esta ecuación es hoy conocida como *la identidad de Selberg*.

En este trabajo deseamos presentar con suficientes detalles, una demostración elemental del teorema de los números primos. Seguiremos la ideas de la demostración de Selberg expuestas por Levinson (1969). Básicamente nuestro trabajo tiene dos partes importantes: la primera es construir la identidad de Selberg, y la segunda es derivar la demostración elemental del teorema de los números primos a partir de la identidad de Selberg. En el primer capítulo estudiaremos las funciones aritméticas que necesitaremos en nuestro trabajo; el segundo capítulo se encarga de las herramientas del análisis necesarias para estudiar el comportamiento asintótico de las sumas parciales de funciones aritméticas; y en el tercer capítulo presentamos una prueba de la identidad de Selberg, y mostraremos cómo se deriva a partir de esta, la prueba elemental del teorema de los números primos.

Capítulo 1

Algunas funciones aritméticas y el producto de Dirichlet

El objeto central de estudio de este capítulo son las *funciones aritméticas*. Una función aritmética es una función f definida en los enteros positivos, que toma valores reales o complejos. Trataremos con algunas funciones aritméticas esenciales para el desarrollo de nuestro trabajo, como la función de Möbius y la función de von Mangoldt. Además, definiremos el producto de Dirichlet entre funciones aritméticas y presentaremos la importante *fórmula de inversión de Möbius*.

1.1. La función μ de Möbius

Cuando se estudia una función aritmética f, interesa determinar, para cada entero positivo n, las sumas de la forma $\sum_{d|n} f(d)$, donde la suma se toma sobre todos los divisores positivos d de n. A este tipo de sumas las llamamos *sumas de divisores* de la función f. En esta sección presentamos la definición de la función μ de Möbius y calcularemos sus sumas de divisores.

Definición 1.1.1. Sea n un entero positivo. Definimos $\mu(n)$ como sigue:

$$\mu(n) = \begin{cases} 1, & \text{si } n = 1; \\ (-1)^k, & \text{si } n = p_1 p_2 \cdots p_k \ (p_i \text{ primos distintos}); \\ 0, & \text{en otro caso.} \end{cases}$$

Esta función aritmética es conocida como la función de Möbius.

Un entero $n > 1$ es llamado *libre de cuadrado* si n no tiene ningún factor de la forma m^2 con $m > 1$. De esta manera tenemos que $\mu(n) = 0$ si y solamente si $n > 1$ y n no es libre de cuadrado. En caso de que $n > 1$ es libre de cuadrado, $\mu(n) = 1$ si y solo si el número de primos distintos que divide a n es par, y $\mu(n) = -1$ si y solo si el número de primos distintos que divide a n es impar. En la siguiente tabla observamos los valores $\mu(n)$ para $1 \leq n \leq 15$:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\mu(n)$	1	-1	-1	0	-1	1	-1	0	0	1	-1	0	-1	1	1

Para $x \in \mathbb{R}$, $\lfloor x \rfloor$ representa el mayor entero menor o igual que x. Así, $\lfloor x \rfloor$ es el único entero que satisface las desigualdades

$$\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1.$$

Definición 1.1.2. La función aritmética I se define como sigue:

$$\mathrm{I}(n) = \left\lfloor \frac{1}{n} \right\rfloor, \text{para todo } n.$$

Claramente tenemos la siguiente igualdad

$$\mathrm{I}(n) = \begin{cases} 1, & \text{si } n = 1; \\ 0, & \text{si } n > 1. \end{cases} \tag{1.1}$$

En realidad podemos definir la función I usando (1.1), sin necesidad de recurrir a $\left\lfloor \frac{1}{n} \right\rfloor$, pero esto último nos permite tener fórmulas compactas, y no por tramos.

En el siguiente teorema mostramos el cálculo de las sumas de divisores de la función de Möbius. Este resultado será de gran importancia más adelante cuando demostremos la fórmula de inversión de Möbius.

Teorema 1.1.1. Para todo entero $n \geq 1$,

$$\sum_{d|n} \mu(d) = \left\lfloor \frac{1}{n} \right\rfloor.$$

Demostración. Para $n = 1$ el resultado es claro. Supongamos que $n > 1$ y sea $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ la representación canónica de n[1]. En la suma $\sum_{d|n} \mu(d)$ los únicos términos distintos de cero provienen de $d = 1$ y de aquellos divisores de n que son productos de primos distintos. Por lo tanto

$$\begin{aligned}
\sum_{d|n} \mu(d) &= \mu(1) + \mu(p_1) + \cdots + \mu(p_k) + \mu(p_1 p_2) + \cdots + \mu(p_{k-1} p_k) + \cdots + \mu(p_1 \cdots p_k) \\
&= 1 + [(-1) + \cdots + (-1)] + \left[(-1)^2 + \cdots + (-1)^2\right] + \cdots + (-1)^k \\
&= 1 + \binom{k}{1}(-1) + \binom{k}{2}(-1)^2 + \cdots + \binom{k}{k}(-1)^k \\
&= (1-1)^k = 0.
\end{aligned}$$

Esto termina la demostración del teorema. □

En la siguiente sección definiremos el producto de Dirichlet y estudiaremos algunas propiedades algebraicas de este mismo.

1.2. Producto de Dirichlet

Antes de ver otras funciones aritméticas, definimos el producto de Dirichlet. Esta operación entre funciones aritméticas le confiere a la colección de todas las funciones aritméticas una estructura algebraica lo suficientemente rica como para deducir resultados importantes sobre las mismas. Prueba de esto será la demostración de la fórmula de inversión de Möbius, que está ligada al estudio de la existencia de inversas de funciones aritméticas, con respecto al producto de Dirichlet.

[1] Todo entero positivo $n > 1$ se puede representar de manera única como un producto de potencias de primos $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ donde $p_1 < p_2 < \cdots < p_k$ son primos y a_i es un entero positivo para $i = 1, 2, \ldots, k$. Esta representación es llamada *representación canónica* o *forma estándar* de n.

Definición 1.2.1. Sean f y g funciones aritméticas. Se define el *producto de Dirichlet de f y g*, denotado por $f * g$, como la función aritmética dada por

$$(f * g)(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right),$$

para todo n.

Para cada divisor positivo d de n tenemos $n = d(n/d)$. Haciendo $a = d$ y $b = n/d$ notamos que tener un divisor positivo d de n equivale a tener una representación de n como producto $n = ab$. De esta forma observamos que

$$(f * g)(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right) = \sum_{n=ab} f(a) g(b)$$

En el siguiente teorema mostramos las propiedades algebraicas generales del producto de Dirichlet.

Teorema 1.2.1. Sean f, g y h funciones aritméticas. Entonces:

1. $f * g = g * f$ (propiedad conmutativa),
2. $(f * g) * h = f * (g * h)$ (propiedad asociativa),
3. $\mathrm{I} * f = f * \mathrm{I} = f$ (identidad).

Demostración. Sabemos que para todo $n \geq 1$, $(f * g)(n)$ se puede escribir de la siguiente manera

$$(f * g)(n) = \sum_{n=ab} f(a) g(b) = \sum_{n=ba} g(b) f(a) = (g * f)(n).$$

Esto demuestra la propiedad conmutativa. Para la asociatividad tenemos

$$\begin{aligned}
((f * g) * h)(n) &= \sum_{n=ab} (f * g)(a) h(b) \\
&= \sum_{n=ab} \left(\sum_{a=cd} f(c) g(d) \right) h(b) \\
&= \sum_{n=cdb} f(c) g(d) h(b) \\
&= \sum_{n=ca} f(c) \left(\sum_{a=db} g(d) h(b) \right) \\
&= \sum_{n=ca} f(c) (g * h)(a) = (f * (g * h))(n).
\end{aligned}$$

Ahora, por conmutatividad, solo necesitamos probar que $f * \mathrm{I} = f$. En efecto, para todo entero positivo n

$$(f * \mathrm{I})(n) = \sum_{d|n} f(d)\mathrm{I}\left(\frac{n}{d}\right) = \sum_{d|n} f(d)\left\lfloor \frac{d}{n} \right\rfloor = f(n),$$

puesto que $\lfloor d/n \rfloor = 0$ si $d < n$. $\square$

1.3. Inversa de Dirichlet y la fórmula de inversión de Möbius

Para una función aritmética f, una inversa (*de Dirichlet*) para f es una función aritmética g tal que $f * g = g * f = \mathrm{I}$. Usando las propiedades de asociatividad e identidad podemos probar que si una función aritmética f tiene una inversa, entonces tal inversa es única. En efecto, si g y h son inversas de f, entonces

$$g = g * \mathrm{I} = g * (f * h) = (g * f) * h = \mathrm{I} * h = h.$$

Cuando la inversa de f existe, la denotaremos por f^{-1}.

No toda función aritmética f tienen inversa. De hecho, si $f(1) = 0$, entonces f no tiene inversa. En efecto, suponiendo que $f(1) = 0$ y f tiene una inversa g, entonces se obtiene que $f * g = \mathrm{I}$ y así $(f * g)(1) = \mathrm{I}(1)$, pero esto es imposible, puesto que $\sum_{d|1} f(d)g(d) = f(1)g(1) = 0$ e $\mathrm{I}(1) = 1$. En el siguiente teorema mostramos que esta condición es suficiente y necesaria para que f tenga inversa, y hallamos una manera recursiva de determinar la inversa de f.

Teorema 1.3.1. Sea f una función aritmética. Entonces f^{-1} existe si, y solamente si, $f(1) \neq 0$. Además, cuando $f(1) \neq 0$, f^{-1} viene dada por la fórmula recursiva

$$f^{-1}(1) = \frac{1}{f(1)}, \quad f^{-1}(n) = -\frac{1}{f(1)} \sum_{\substack{d|n \\ d<n}} f^{-1}(d) f\left(\frac{n}{d}\right), \quad \textit{para } n > 1. \tag{1.2}$$

Demostración. Si suponemos que f^{-1} existe y $f(1) = 0$, como vimos anteriormente, llegamos a una contradicción. Por lo tanto, si f^{-1} existe, entonces $f(1) \neq 0$.

Debemos probar ahora que si $f(1) \neq 0$, entonces f^{-1} existe y obedece a la fórmula recursiva indicada. Suponga entonces que $f(1) \neq 0$, esta condición nos permite, de manera recursiva, definir una función aritmética g como sigue:

$$g(1) = \frac{1}{f(1)}, \quad g(n) = -\frac{1}{f(1)} \sum_{\substack{d|n \\ d<n}} g(d) f\left(\frac{n}{d}\right), \quad \textit{para } n > 1. \tag{1.3}$$

Ahora notemos que

$$(g * f)(1) = \sum_{d|1} g(d) f\left(\frac{1}{d}\right) = g(1)f(1) = \frac{f(1)}{f(1)} = 1,$$

de igual forma podemos mostrar que $(f * g)(1) = 1$. Ahora, para todo $n > 1$

$$\begin{aligned}(g * f)(n) &= \sum_{d|n} g(d) f\left(\frac{n}{d}\right) = \sum_{\substack{d|n \\ d<n}} g(d) f\left(\frac{n}{d}\right) + f(1)g(n) \\ &= \sum_{\substack{d|n \\ d<n}} g(d) f\left(\frac{n}{d}\right) + f(1)\left(-\frac{1}{f(1)} \sum_{\substack{d|n \\ d<n}} g(d) f\left(\frac{n}{d}\right)\right) = 0.\end{aligned}$$

Análogamente también mostramos que $(f * g)(n) = 0$. De esta forma, lo que hemos probado es que

$$f * g = g * f = \mathrm{I}.$$

Esto muestra que g es la inversa de Dirichlet de f. Por lo tanto f^{-1} existe y coincide con g. Finalmente, reemplazando g por f^{-1} en (1.3) encontramos que f^{-1} viene dada por la fórmula recursiva (1.2). $\square$

A continuación determinamos la inversa de la función de Möbius. Aplicando el Teorema 1.1.1 sabemos que para todo $n \geq 1$, $\sum_{d|n} \mu(d) = \mathrm{I}(n)$. Si definimos una función u mediante $u(n) = 1$ para todo n, entonces podemos reescribir esta igualdad como $\sum_{d|n} \mu(d) u\left(\frac{n}{d}\right) = \mathrm{I}(n)$. Así, $(\mu * u)(n) = I(n)$, para todo n, es decir, $\mu * u = \mathrm{I}$. Por lo tanto, μ y u son inversas entre sí: $u = \mu^{-1}$ y $\mu = u^{-1}$.

Ahora estamos en posición de demostrar la fórmula de inversión de Möbius.

Teorema 1.3.2 (Fórmula de inversión de Möbius)**.** Sean f y g funciones aritméticas. Para $n \geq 1$, se tiene

$$f(n) = \sum_{d|n} g(d) \tag{1.4}$$

si, y solo si

$$g(n) = \sum_{d|n} f(d)\mu\left(\frac{n}{d}\right). \tag{1.5}$$

Demostración. La ecuación (1.4) es equivalente a $f = g*u$, mientras que (1.5) equivale a $g = f * \mu$. Así, debemos probar que $f = g * u$ si y solo si $g = f * \mu$. En efecto, si $f = g * u$, entonces

$$f * \mu = (g * u) * \mu = g * (u * \mu) = g * \mathrm{I} = g.$$

Recíprocamente, si $g = f * \mu$, entonces

$$g * u = (f * \mu) * u = f * (\mu * u) = f * \mathrm{I} = f. \qquad \square$$

1.4. La Función de Von Mangoldt

En esta sección presentamos la función Λ de Mangoldt, que juega un papel central en el estudio de la distribución de los números primos.

Definición 1.4.1. Para todo entero $n \geq 1$ se define

$$\Lambda(n) = \begin{cases} \ln p, & \text{si } n = p^k, \text{ donde } p \text{ es primo y } k \geq 1; \\ 0, & \text{en otro caso.} \end{cases}$$

En la siguiente tabla mostramos los primeros 10 valores de $\Lambda(n)$.

n	1	2	3	4	5	6	7	8	9	10
$\Lambda(n)$	0	$\ln 2$	$\ln 3$	$\ln 2$	$\ln 5$	0	$\ln 7$	$\ln 2$	$\ln 3$	0

La función de Mangoldt no tiene inversa puesto que $\Lambda(1) = 0$. En el siguiente teorema determinamos las sumas de divisores de Λ.

Teorema 1.4.1 (Sumas de divisores de $\Lambda(n)$)**.** Para todo entero $n \geq 1$,

$$\sum_{d|n} \Lambda(d) = \ln n. \tag{1.6}$$

Demostración. Para $n = 1$, el resultado es claro. Supongamos que $n > 1$ y sea $n = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$ la descomposición canónica de n. Por las propiedades del logaritmo tenemos

$$\ln n = \sum_{i=1}^{r} a_i \ln p_i.$$

En la función de Mangoldt los únicos términos distintos de cero provienen de potencias de primos, por lo que

$$\begin{aligned}
\sum_{d|n} \Lambda(d) &= \Lambda(p_1) + \Lambda(p_1^2) + \cdots + \Lambda(p_1^{a_1}) + \cdots + \Lambda(p_r) + \Lambda(p_r^2) + \cdots + \Lambda(p_r^{a_r}) \\
&= \underbrace{\ln p_1 + \ln p_1 + \cdots + \ln p_1}_{a_1-veces} + \cdots + \underbrace{\ln p_r + \ln p_r + \cdots + \ln p_r}_{a_r-veces} \\
&= a_1 \ln p_1 + \cdots + a_r \ln p_r \\
&= \sum_{i=1}^{r} a_i \ln p_i = \ln n.
\end{aligned}$$

□

Aplicando la fórmula de inversión de Möbius a (1.6) podemos obtener una fórmula alternativa para $\Lambda(n)$ en términos de la función logaritmo y μ, como lo mostramos en el siguiente teorema.

Teorema 1.4.2. Si $n \geq 1$, entonces

$$\Lambda(n) = -\sum_{d|n} \mu(d) \ln d.$$

Demostración. Por la fórmula de inversión de Möbius aplicada a (1.6) obtenemos

$$\begin{aligned}
\Lambda(n) = \sum_{d|n} \mu(d) \ln\left(\frac{n}{d}\right) &= \sum_{d|n} \mu(d) \left(\ln n - \ln d\right) \\
&= \sum_{d|n} \mu(d) \ln n - \sum_{d|n} \mu(d) \ln d \\
&= \ln(n) \sum_{d|n} \mu(d) - \sum_{d|n} \mu(d) \ln d.
\end{aligned}$$

Notemos que $\ln n \sum_{d|n} \mu(d) = 0$, pues para $n > 1$, se tiene que $\sum_{d|n} \mu(d) = \lfloor \frac{1}{n} \rfloor = 0$, y para $n = 1$, $\ln n = 0$. Así

$$\Lambda(n) = -\sum_{d|n} \mu(d) \ln d.$$

□

Capítulo 2

Sumas parciales de funciones aritméticas

Muchas veces ocurre que entender el comportamiento asintótico de una función aritmética dada f es un problema difícil. En su lugar, puede resultar provechoso entender el comportamiento asintótico del promedio $\frac{1}{n}\sum_{k=1}^{n} f(k)$, lo que hace que, naturalmente, dirijamos nuestra atención al estudio de las sumas $\sum_{k=1}^{n} f(k)$. Es conveniente reemplazar el índice superior n por un número real positivo arbitrario x y considerar en su lugar las *sumas parciales* de la forma

$$\sum_{k\leq x} f(k),$$

donde k recorre todos los enteros de 1 a $\lfloor x \rfloor$. Si n es un entero positivo, el valor de la suma $\sum_{k\leq x} f(k)$ es constante para $n \leq x < n+1$, y posiblemente tenga una salto en $x = n+1$, dependiendo del valor de $f(n+1)$. Si $0 < x < 1$, el valor de la suma es, por convención, 0. Al considerar estas sumas parciales, podemos involucrar herramientas del análisis matemático en el estudio de funciones aritméticas. Nuestro objetivo en este capítulo es desarrollar herramientas que nos permitan entender el comportamiento asintótico de las sumas parciales de funciones aritméticas.

2.1. La notación O grande

Una de las principales herramientas para el estudio del comportamiento asintótico de funciones es la notación O ("O grande"), que definimos a continuación. Aunque la definición puede darse en un contexto más general, la aplicaremos solo para funciones reales definidas en los reales positivos.

Definición 2.1.1. Sean $f, g : (0, \infty) \to \mathbb{R}$ funciones, y suponga que $g(x) > 0$ siempre que $x \geq a$, para algún $a \in \mathbb{R}$. Decimos que $f(x)$ es "O grande" de $g(x)$ y escribimos

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

si existe una constante $M > 0$ tal que

$$|f(x)| \leq Mg(x)$$

para todo $x \geq a$.

De manera equivalente, tener $f(x) = O(g(x))$ significa que el cociente $f(x)/g(x)$ está acotado para todo $x \geq a$.

Una ecuación de la forma $f = h + O(g)$ significa que $f - h = O(g)$. Mostramos a continuación algunas propiedades de la notación "O grande".

Teorema 2.1.1. Sean f, g, k y h funciones reales. Entonces los siguientes enunciados son ciertos:

1. Si $f(x) = O(h(x))$ y $h(x) = O(g(x))$, entonces $f(x) = O(g(x))$.

2. Si $f(x) = O(g(x))$ y $k(x) = O(h(x))$, entonces $f(x) + k(x) = O(g(x) + h(x))$.

3. Si $f(x) = O(h(x))$ y $k(x) = O(g(x))$, entonces $f(x)k(x) = O(h(x)g(x))$.

4. $k(x)O(g(x)) = O(|k(x)|g(x))$.

5. $O(f(x)) + O(g(x)) = O(f(x) + g(x))$.

6. $O\left(O(g(x))\right) = O(g(x))$.

7. Si $h(x) = O(g(x))$, entonces $O(h(x)) + O(g(x)) = O(g(x))$.

8. Sea $r \in \mathbb{R}$ con $r \neq 0$. Entonces $f(x) = O(g(x))$ si, y solo si, $rf(x) = O(g(x))$.

9. Si $f(t) = O(g(t))$, y f y g son integrables en $[x_0, \infty)$ para algún $x_0 \in \mathbb{R}$, entonces existe un $a \in \mathbb{R}$ tal que $\int_a^x f(t)dt = O\left(\int_a^x g(t)dt\right)$.

Las propiedades 4, 5, 6 y 7 en realidad pueden ser vistas como igualdades entre conjuntos, y de hecho, la demostración de estas propiedades estarán enfocadas en este sentido.

Demostración. 1. Por definición existe una constante M_1 tal que $|f(x)| \leq M_1 h(x)$ siempre que $x \geq a_1$ para algún $a_1 \in \mathbb{R}$ e igualmente existe una constante M_2 y un real a_2 tal que $|h(x)| \leq M_2 g(x)$ para todo $x \geq a_2$. Tomando $a = \text{máx}\{a_1, a_2\}$ encontramos que $|f(x)| \leq M_1 h(x) = M_1 |h(x)| \leq M_1 M_2 g(x)$ siempre que $x \geq a$. Así, $f(x) = O(g(x))$.

2. Existen constantes M_1, M_2 y reales a_1, a_2 tales que $|f(x)| \leq M_1 g(x)$ siempre que $x \geq a_1$ y $|k(x)| \leq M_2 h(x)$ siempre que $x \geq a_2$. Tomando $a = \text{máx}\{a_1, a_2\}$ y $M = \text{máx}\{M_1, M_2\}$, tenemos

$$|f(x) + k(x)| \leq |f(x)| + |k(x)| \leq M_1 g(x) + M_2 h(x) \leq M(g(x) + h(x)),$$

para todo $x \geq a$. En consecuencia $f(x) + k(x) = O(g(x) + h(x))$.

3. Como $f(x) = O(h(x))$ y $k(x) = O(g(x))$. Entonces, por definición, existen constantes M_1, M_2 y reales a_1, a_2 tales que $|f(x)| \leq M_1 h(x)$ siempre que $x \geq a_1$ e igualmente se obtiene $|k(x)| \leq M_2 g(x)$ siempre que $x \geq a_2$. Tomando $a = \text{máx}\{a_1, a_2\}$ encontramos que $|f(x)k(x)| \leq M_1 M_2 h(x) g(x)$ siempre que $x \geq a$. En consecuencia, $f(x)k(x) = O(h(x)g(x))$.

4. Sea σ una función real y suponga que $\sigma(x) = O(g(x))$. Dada la definición de O grande, tener $h(x) = O(|k(x)|g(x))$ implica que $|k(x)|g(x) > 0$, para x suficientemente grande, de donde $k(x) \neq 0$ para todo x con x suficientemente grande. Luego, $k(x) = O(|k(x)|)$. Aplicando la propiedad 3 para $k(x) = O(|k(x)|)$ y $\sigma(x) = O(g(x))$, encontramos que $k(x)\sigma(x) = O(|k(x)|g(x))$ por lo tanto $k(x)O(g(x)) = O(|k(x)|g(x))$.

5. Para mostrar que $O(f(x)) + O(g(x)) = O(f(x) + g(x))$, tomemos funciones arbitrarias σ_1, σ_2 tales que $\sigma_1(x) = O(f(x))$ y $\sigma_2(x) = O(g(x))$. Ahora, usando la propiedad 3, $\sigma_1(x)+\sigma_2(x) = O(f(x)+g(x))$. Así, $O(f(x))+O(g(x)) = O(f(x)+g(x))$.

6. Vamos a mostrar que $O\left(O(g(x))\right) = O(g(x))$. Tomemos funciones arbitrarias f, h tales que $f(x) = O(g(x))$ y $h(x) = O\left(f(x)\right)$, entonces existen constantes M_1, M_2 y reales a_1, a_2 tales que $|f(x)| \leq M_1 g(x)$ siempre que $x \geq a_1$, y $|h(x)| \leq M_2 f(x)$ siempre que $x \geq a_2$. Luego, para $a = \text{máx}\{a_1, a_2\}$ tenemos que

$$|h(x)| \leq M_2 f(x) = M_2 |f(x)| \leq M_1 M_2 g(x),$$

para todo $x \geq a$. En consecuencia $h(x) = O(g(x))$, así $O\left(O(g(x))\right) = O(g(x))$.

7. Veamos que $O(h(x)) + O(g(x)) = O(g(x))$ si $h(x) = O(g(x))$. Considerar funciones arbitrarias σ_1 y σ_2 tales que $\sigma_1(x) = O(h(x))$ y $\sigma_2(x) = O(g(x))$. Por hipótesis tenemos que $h(x) = O(g(x))$ así, de la propiedad 1, se obtiene $\sigma_1(x) = O(g(x))$. Luego, de la propiedad 2, aplicada a $\sigma_1(x) = O(g(x))$ y $\sigma_2(x) = O(g(x))$, se tiene que $\sigma_1(x) + \sigma_2(x) = O(2g(x))$. Esto claramente implica que $\sigma_1(x) + \sigma_2(x) = O(g(x))$ y por lo tanto $O(h(x)) + O(g(x)) = O(g(x))$.

8. Sea $r \neq 0$. Mostremos primero que $f(x) = O(g(x))$ implica $rf(x) = O(g(x))$. Por definición, sabemos que existe una constante M y un real a de tal forma que $|f(x)| \leq Mg(x)$ siempre que $x \geq a$, luego, $|rf(x)| \leq |r|Mg(x)$ siempre que $x \geq a$. De esta manera tenemos que $rf(x) = O(g(x))$. Recíprocamente, si $rf(x) = O(g(x))$ existe una constante M y un real a tales que $|rf(x)| \leq Mg(x)$ siempre que $x \geq a$. Como $r \neq 0$, podemos dividir por $|r|$ para obtener que $|f(x)| \leq (M/|r|)g(x)$ para todo $x \geq a$, o equivalentemente $f(x) = O(g(x))$.

9. Como $f(t) = O(g(t))$, entonces existe una constante M y un real a_0 tales que $|f(t)| \leq Mg(t)$ siempre que $x \geq a_0$. Tomando $a = \text{máx}\{a_0, x_0\}$ encontramos que $|\int_a^x f(t)dt| \leq \int_a^x |f(t)|dt \leq M \int_a^x g(t)dt$ para todo $x \geq a$. En consecuencia se obtiene que $\int_a^x f(t)dt = O\left(\int_a^x g(t)dt\right)$. $\square$

2.2. Fórmula de sumación de Euler

La fórmula de sumación de Euler es una herramienta importante en la estimación asintótica de sumas de la forma $\sum_{y<n\leq x} f(n)$. En esta sección daremos la prueba de la fórmula, y la aplicaremos para obtener algunas fórmulas asintóticas especiales.

Teorema 2.2.1 (Fórmula de sumación de Euler)**.** Sea f una función diferenciable en el intervalo $[y, x]$, donde $0 < y < x$. Entonces

$$\sum_{y<n\leq x} f(n) = \int_y^x f(t)dt + \int_y^x (t - \lfloor t \rfloor) f'(t)dt + (\lfloor x \rfloor - x)f(x) - (\lfloor y \rfloor - y)f(y).$$

Demostración. Sean $m = \lfloor y \rfloor$ y $k = \lfloor x \rfloor$. Para enteros consecutivos n y $n-1$, tenemos

$$\begin{aligned}\int_{n-1}^n \lfloor t \rfloor f'(t)dt &= \int_{n-1}^n (n-1)f'(t)dt = (n-1)[f(n) - f(n-1)] \\ &= [nf(n) - (n-1)f(n-1)] - f(n).\end{aligned}$$

Sumando desde $n = m+1$ a $n = k$, se obtiene que

$$\begin{aligned}\int_m^k \lfloor t \rfloor f'(t)dt &= \sum_{n=m+1}^k \int_{n-1}^n \lfloor t \rfloor f'(t)dt \\ &= \sum_{n=m+1}^k \{[nf(n) - (n-1)f(n-1)] - f(n)\} \\ &= \sum_{n=m+1}^k [nf(n) - (n-1)f(n-1)] - \sum_{n=m+1}^k f(n) \\ &= kf(k) - mf(m) - \sum_{y<n\leq x} f(n).\end{aligned}$$

En consecuencia

$$\sum_{y<n\leq x} f(n) = -\int_m^k \lfloor t \rfloor f'(t)dt + kf(k) - mf(m). \tag{2.1}$$

Ahora, como $m \leq y \leq k \leq x$, se sigue que

$$\int_m^k \lfloor t \rfloor f'(t)dt + \int_k^x \lfloor t \rfloor f'(t)dt = \int_m^y \lfloor t \rfloor f'(t)dt + \int_y^x \lfloor t \rfloor f'(t)dt. \tag{2.2}$$

Notando que $\lfloor t \rfloor = \lfloor x \rfloor$ para todo t con $\lfloor x \rfloor \leq t \leq x$, obtenemos lo siguiente

$$\int_{\lfloor x \rfloor}^x \lfloor t \rfloor f'(t)dt = \lfloor x \rfloor \int_{\lfloor x \rfloor}^x f'(t)dt = \lfloor x \rfloor f(x) - \lfloor x \rfloor f(\lfloor x \rfloor),$$

como $k = \lfloor x \rfloor$ la anterior igualdad implica que $\int_k^x \lfloor t \rfloor f'(t)dt = kf(x) - kf(k)$. De forma análoga podemos mostramos que $\int_m^y \lfloor t \rfloor f'(t)dt = mf(y) - mf(m)$, llevando esto a (2.2) y reordenando términos se obtiene la siguiente igualdad

$$-\int_m^k \lfloor t \rfloor f'(t)dt + kf(k) - mf(m) = -\int_y^x \lfloor t \rfloor f'(t)dt + kf(x) - mf(y). \tag{2.3}$$

Recordando que $m = \lfloor y \rfloor$ y $k = \lfloor x \rfloor$, se sigue de (2.1) y (2.3) que

$$\sum_{y<n\leq x} f(n) = -\int_y^x \lfloor t \rfloor f'(t)dt + \lfloor x \rfloor f(x) - \lfloor y \rfloor f(y). \tag{2.4}$$

La integración por partes aplicada a $\int_y^x f(t)dt$ nos da lo siguiente:

$$\begin{aligned}
\int_y^x f(t)dt &= xf(x) - yf(y) - \int_y^x tf'(t)dt \\
&= xf(x) - yf(y) - \int_y^x (t - \lfloor t \rfloor + \lfloor t \rfloor) f'(t)dt \\
&= xf(x) - yf(y) - \int_y^x (t - \lfloor t \rfloor) f'(t)dt - \int_y^x \lfloor t \rfloor f'(t)dt,
\end{aligned}$$

así que

$$-\int_y^x \lfloor t \rfloor f'(t)dt = \int_y^x f(t)dt + \int_y^x (t - \lfloor t \rfloor) f'(t)dt - xf(x) + yf(y). \tag{2.5}$$

Finalmente, reemplazando (2.5) en (2.4) se obtiene el resultado deseado. □

Aplicando la fórmula de sumación de Euler a la función $f(x) = 1/x$ obtenemos

$$\begin{aligned}
\sum_{n\leq x} \frac{1}{n} = 1 + \sum_{1<n\leq x} \frac{1}{n} &= 1 + \int_1^x \frac{1}{t}dt - \int_1^x \frac{t - \lfloor t \rfloor}{t^2}dt + \frac{\lfloor x \rfloor - x}{x} \\
&= 1 + \ln x - \int_1^x \frac{t - \lfloor t \rfloor}{t^2}dt + \frac{\lfloor x \rfloor - x}{x}.
\end{aligned}$$

Luego

$$\sum_{n\leq x} \frac{1}{n} - \ln x = 1 - \int_1^x \frac{t - \lfloor t \rfloor}{t^2}dt + \frac{\lfloor x \rfloor - x}{x}.$$

Cuando $x \to \infty$, en la ecuación anterior resulta

$$\lim_{x\to\infty} \left(\sum_{n\leq x} \frac{1}{n} - \ln x \right) = 1 - \int_1^\infty \frac{t - \lfloor t \rfloor}{t^2}dt. \tag{2.6}$$

Como $0 \leq (t - \lfloor t \rfloor)/t^2 < 1/t^2$, se sigue que la integral impropia $\displaystyle\int_1^\infty \frac{t - \lfloor t \rfloor}{t^2} dt$ converge. Más aún,

$$0 \leq \int_1^\infty \frac{t - \lfloor t \rfloor}{t^2} dt < 1.$$

De esta forma, el límite de la izquierda en (2.6) existe y es positivo.

Definición 2.2.1. La constante γ de Euler-Mascheroni se define como

$$\gamma = \lim_{x \to \infty} \left(\sum_{n \leq x} \frac{1}{n} - \ln x \right).$$

El siguiente teorema nos da una fórmula asintótica de suma importancia en este trabajo la cual es consecuencia de la fórmula de sumación de Euler.

Teorema 2.2.2. Para $x \geq 1$

$$\sum_{n \leq x} \frac{1}{n} = \ln x + \gamma + O\left(\frac{1}{x}\right). \tag{2.7}$$

Demostración. Aplicando la fórmula de sumación de Euler con $f(t) = 1/t$ obtenemos.

$$\begin{aligned}
\sum_{n \leq x} \frac{1}{n} &= 1 + \ln x - \int_1^x \frac{t - \lfloor t \rfloor}{t^2} dt + \frac{\lfloor x \rfloor - x}{x} \\
&= \ln x + \left(1 - \int_1^\infty \frac{t - \lfloor t \rfloor}{t^2} dt\right) + \int_x^\infty \frac{t - \lfloor t \rfloor}{t^2} dt + O\left(\frac{1}{x}\right) \\
&= \ln x + \gamma + O\left(\int_x^\infty \frac{1}{t^2} dt\right) + O\left(\frac{1}{x}\right) \\
&= \ln x + \gamma + O\left(\frac{1}{x}\right) + O\left(\frac{1}{x}\right) \\
&= \ln x + \gamma + O\left(\frac{1}{x}\right).
\end{aligned}$$

Por lo tanto se cumple (2.7). □

2.3. Sumas parciales de un producto de Dirichlet

Generalizamos el producto de Dirichlet de tal forma que para una función aritmética dada α y una función F definida para todo real positivo, con $F(x) = 0$ para $0 < x < 1$, obtengamos una función $\alpha \circ F$ definida para todo real positivo.

Definición 2.3.1 (Producto generalizado de Dirichlet)**.** Sean α una función aritmética y F una función real definida en $(0,\infty)$ tal que $F(x)=0$ para todo $0<x<1$. Se define la función $\alpha\circ F:(0,\infty)\to\mathbb{R}$ por

$$(\alpha\circ F)(x)=\sum_{n\leq x}\alpha(n)F\left(\frac{x}{n}\right),$$

para todo $x>0$.

Observemos que la función $\alpha\circ F$ también se anula para todo $0<x<1$.

Si tuviéramos que $F(x)=0$ para todo x que no sea entero y restringimos F al conjunto de los enteros positivos, encontramos que

$$(\alpha\circ F)(n)=(\alpha*F)(n),$$

para todo entero $n\geq 1$, por lo que esta operación, de hecho, puede considerarse como una generalización del producto de Dirichlet. A continuación probamos la propiedad asociativa que es satisfecha por la operación $\circ$.

Teorema 2.3.1. Sean α y β funciones aritméticas. Si F es una función real definida en $(0,\infty)$ tal que $F(x)=0$ para todo $0<x<1$, entonces

$$\alpha\circ(\beta\circ F)=(\alpha*\beta)\circ F.$$

Demostración. Si $x>0$, tenemos

$$\begin{aligned}(\alpha\circ(\beta\circ F))(x)&=\sum_{n\leq x}\alpha(n)(\beta\circ F)\left(\frac{x}{n}\right)\\&=\sum_{n\leq x}\alpha(n)\sum_{m\leq x/n}\beta(m)F\left(\frac{x}{mn}\right)\\&=\sum_{mn\leq x}\alpha(n)\beta(m)F\left(\frac{x}{mn}\right).\end{aligned}$$

Haciendo $k=mn$ notamos que tener un par (m,n) tal que $mn\leq x$ es equivalente a tener un par (k,n) tal que $k\leq x$ y $n\mid k$. Así,

$$\begin{aligned}\sum_{mn\leq x}\alpha(n)\beta(m)F\left(\frac{x}{mn}\right)&=\sum_{k\leq x}\left(\sum_{n\mid k}\alpha(n)\beta\left(\frac{k}{n}\right)\right)F\left(\frac{x}{k}\right)\\&=\sum_{k\leq x}(\alpha*\beta)(k)F\left(\frac{x}{k}\right)\\&=((\alpha*\beta)\circ F)(x).\end{aligned}$$

Esto muestra que $\alpha \circ (\beta \circ F) = (\alpha * \beta) \circ F$. $\square$

Definimos una función $U : (0, \infty) \to \mathbb{R}$ como sigue:

$$U(x) = \begin{cases} 0, & \text{si } 0 < x < 1; \\ 1, & \text{si } x \geq 1. \end{cases}$$

Si f es una función aritmética, entonces tenemos que

$$(f \circ U)(x) = \sum_{n \leq x} f(n) U\left(\frac{x}{n}\right) = \sum_{n \leq x} f(n),$$

es decir, la suma parcial $\sum_{n \leq x} f(n)$ es igual a $(f \circ U)(x)$. En el siguiente teorema mostramos cómo se relacionan las sumas parciales de un producto de Dirichlet $f * g$, con las sumas parciales de f y g.

Teorema 2.3.2 (Sumas parciales de un producto de Dirichlet)**.** Sean f, g funciones aritméticas y $h = f * g$. Defina

$$H(x) = \sum_{n \leq x} h(n), \quad F(x) = \sum_{n \leq x} f(n), \quad y \quad G(x) = \sum_{n \leq x} g(n).$$

Entonces

$$H(x) = \sum_{n \leq x} f(n) G\left(\frac{x}{n}\right) = \sum_{n \leq x} g(n) F\left(\frac{x}{n}\right). \tag{2.8}$$

Demostración. Por las observaciones hechas arriba tenemos que $F = f \circ U$, $G = g \circ U$ y $H = h \circ U$. Luego, del Teorema 2.3.1 se sigue que

$$f \circ G = f \circ (g \circ U) = (f * g) \circ U = h \circ U = H,$$
$$g \circ F = g \circ (f \circ U) = (g * f) \circ U = h \circ U = H.$$

Finalmente, aplicado la Definición 2.3.1 a las igualdades

$$H(x) = (f \circ G)(x) = (g \circ F)(x),$$

se obtiene (2.8). $\square$

El siguiente resultado es una consecuencia directa del Teorema 2.3.2 y será de gran importancia para mostrar importantes propiedades sobre las función Λ.

Corolario 2.3.3. Sea f una función aritmética. Si $F(x) = \sum_{n \leq x} f(n)$, entonces

$$\sum_{n \leq x} \sum_{d|n} f(d) = \sum_{n \leq x} f(n) \left\lfloor \frac{x}{n} \right\rfloor = \sum_{n \leq x} F\left(\frac{x}{n}\right). \tag{2.9}$$

Demostración. La prueba se obtiene aplicando el Teorema 2.3.2 con $g(n) = 1$ para todo n, teniendo en cuenta que $\sum_{m \leq x/n} 1 = \left\lfloor \frac{x}{n} \right\rfloor$. □

2.4. Sumas parciales que involucran a la función Λ; la función ψ de Chebyshev

En esta sección definiremos la función ψ de Chebyshev la cual, como veremos en el siguiente capítulo, nos permite tener una forma equivalente al teorema de los números primos. Además, estableceremos una fórmula asintótica para las sumas parciales de $\Lambda(n)/n$. Comenzamos con una fórmula asintótica para las sumas parciales de $\ln n$.

Teorema 2.4.1. Si $x \geq 2$, entonces

$$\sum_{n \leq x} \ln n = x \ln x - x + O(\ln x). \tag{2.10}$$

Demostración. Aplicando la fórmula de sumación de Euler con $f(t) = \ln t$ se tiene

$$\begin{aligned}
\sum_{n \leq x} \ln n = \sum_{1 < n \leq x} \ln n &= \int_1^x \ln t dt + \int_1^x \frac{t - \lfloor t \rfloor}{t} dt - (x - \lfloor x \rfloor) \ln x \\
&= (x \ln x - x + 1) + O\left(\int_1^x \frac{1}{t} dt\right) + O(\ln x) \\
&= x \ln x - x + O(\ln x),
\end{aligned}$$

lo que termina la prueba del teorema. □

El Corolario 2.3.3 nos permite tener una forma alternativa de ver el anterior resultado, ya que aplicado a $f(n) = \Lambda(n)$ nos da

$$\sum_{n \leq x} \Lambda(n) \left\lfloor \frac{x}{n} \right\rfloor = \sum_{n \leq x} \sum_{d|n} \Lambda(d) = \sum_{n \leq x} \ln n,$$

y de esta manera obtenemos

$$\sum_{n \leq x} \Lambda(n) \left\lfloor \frac{x}{n} \right\rfloor = x \ln x - x + O(\ln x). \tag{2.11}$$

Definición 2.4.1. Definimos la función ψ de Chebyshev por la fórmula

$$\psi(x) = \sum_{n \leq x} \Lambda(n),$$

para todo $x > 0$.

Cabe resaltar que la función $\Lambda(n)$ aporta valores a la suma si, y solamente si, n es una potencia de primo, por lo que $\psi(x)$ se mantiene constante para valores de x entre dos potencias sucesivas de primos y da un salto en cada potencia de primo; así $\psi(x)$ es discontinua en $x = p^m$, donde p es un número primo y m cualquier entero mayor o igual a uno.

Teorema 2.4.2. Para x suficientemente grande se tiene que $\psi(x) < (3/2)x$. En particular, $\psi(x) = O(x)$.

Demostración. Teniendo en cuanta que para todo real y se tiene $\lfloor y \rfloor \geq 2\lfloor y/2 \rfloor$, tenemos que para $x \geq 2$

$$\begin{aligned}
\sum_{n \leq x} \Lambda(n) \left\lfloor \frac{x}{n} \right\rfloor - 2 \sum_{n \leq x/2} \Lambda(n) \left\lfloor \frac{x}{2n} \right\rfloor &= \sum_{n \leq x/2} \Lambda(n) \left(\left\lfloor \frac{x}{n} \right\rfloor - 2 \left\lfloor \frac{x}{2n} \right\rfloor \right) + \sum_{x/2 < n \leq x} \Lambda(n) \left\lfloor \frac{x}{n} \right\rfloor \\
&\geq \sum_{x/2 < n \leq x} \Lambda(n) \left\lfloor \frac{x}{n} \right\rfloor \\
&= \sum_{x/2 < n \leq x} \Lambda(n) \\
&= \psi(x) - \psi\left(\frac{x}{2}\right).
\end{aligned}$$

Por otro lado, se sigue de la fórmula (2.11) que

$$\begin{aligned}
\sum_{n \leq x} \Lambda(n) \left\lfloor \frac{x}{n} \right\rfloor - 2 \sum_{n \leq x/2} \Lambda(n) \left\lfloor \frac{x}{2n} \right\rfloor &= x \ln x - x + O(\ln x) \\
&\quad - 2\left(\frac{x}{2} \ln \frac{x}{2} - \frac{x}{2} + O\left(\ln \frac{x}{2}\right)\right) \\
&= x \ln x - x + O(\ln x) - x(\ln x - \ln 2) \\
&\quad - x + O\left(\ln \frac{x}{2}\right) = x \ln 2 + O(\ln x).
\end{aligned}$$

Hemos mostrado así que $\psi(x) - \psi(x/2) = x \ln 2 + O(\ln x)$. Esto quiere decir que existe una constante M tal que para x suficientemente grande se tiene

$$\psi(x) - \psi\left(\frac{x}{2}\right) \leq x \ln 2 + M \ln x. \tag{2.12}$$

Teniendo en cuenta que $\ln(x/2^n) \leq \ln x$ para todo entero positivo n, si reemplazamos x sucesivamente por $x, x/2, x/4, \ldots$ en la desigualdad (2.12) podemos construir las siguientes desigualdades:

$$\begin{aligned}
\psi(x) - \psi\left(\frac{x}{2}\right) &\leq x\ln 2 + M\ln x,\\
\psi\left(\frac{x}{2}\right) - \psi\left(\frac{x}{4}\right) &\leq \frac{x}{2}\ln 2 + M\ln x,\\
\psi\left(\frac{x}{4}\right) - \psi\left(\frac{x}{8}\right) &\leq \frac{x}{4}\ln 2 + M\ln x,\\
&\vdots\\
\psi\left(\frac{x}{2^n}\right) - \psi\left(\frac{x}{2^{n+1}}\right) &\leq \frac{x}{2^n}\ln 2 + M\ln x.
\end{aligned}$$

Ahora notemos que $\psi(x/2^n) = 0$ cuando $2^n > x$, o equivalentemente cuando $n < \ln x/\ln 2$, así que solo estamos considerando a lo sumo $\ln x/\ln 2$ desigualdades. Sumando tales desigualdades obtenemos

$$\psi(x) \leq x\ln 2\left(1 + \frac{1}{2} + \frac{1}{4} + \cdots\right) + \frac{\ln x}{\ln 2}(M\ln x) = 2x\ln 2 + \frac{M}{\ln 2}\ln^2 x. \tag{2.13}$$

Usando la regla de L'Hôpital obtenemos que $\lim\limits_{x\to\infty} \dfrac{M\ln^2 x}{x\ln 2} = 0$. En consecuencia, para x suficientemente grande se tiene que

$$2\ln 2 + \frac{M\ln^2 x}{x\ln 2} \leq 2\ln 2 + 0.1,$$

pero $2\ln 2 + 0.1 < 3/2$, lo que implica que

$$2x\ln 2 + \frac{M\ln x\ln x}{\ln 2} < \frac{3}{2}x,$$

para x suficientemente grande. Esto combinado con (2.13) termina la prueba. $\square$

Finalizamos la sección con la fórmula prometida para las sumas parciales de $\Lambda(n)/n$. Este resultado juega un papel crucial a la hora de demostrar la fórmula asintótica de Selberg.

Teorema 2.4.3. Para $x \geq 1$

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

Demostración. Usando que $\lfloor x \rfloor = x + O(1)$ y el Teorema 2.4.2 se sigue que

$$\begin{aligned}
\sum_{n \leq x} \Lambda(n) \left\lfloor \frac{x}{n} \right\rfloor &= \sum_{n \leq x} \Lambda(n) \left(\frac{x}{n} + O(1) \right) \\
&= x \sum_{n \leq x} \frac{\Lambda(n)}{n} + O\left(\sum_{n \leq x} \Lambda(n) \right) \\
&= x \sum_{n \leq x} \frac{\Lambda(n)}{n} + O(\psi(x)) \\
&= x \sum_{n \leq x} \frac{\Lambda(n)}{n} + O(x).
\end{aligned}$$

En consecuencia

$$\sum_{n \leq x} \frac{\Lambda(n)}{n} = \frac{1}{x} \sum_{n \leq x} \Lambda(n) \left\lfloor \frac{x}{n} \right\rfloor + O(1),$$

pero de (2.11) se sigue que $\sum_{n \leq x} \Lambda(n) \left\lfloor \frac{x}{n} \right\rfloor = x \ln x + O(x)$, por lo que

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

Así termina la prueba. $\square$

Capítulo 3

El teorema de los números primos

En este capítulo presentamos una demostración elemental del teorema de los números primos. Las ideas principales son las desarrolladas por Selberg [4], pero seguimos la exposición de Levinson [3]. La idea consiste en demostrar una formulación equivalente al teorema de los números primos, y esto se logrará a través de la fórmula de Selberg.

3.1. El teorema de los números primos y una forma equivalente con la función ψ

Para cada $x > 0$, $\pi(x)$ representa la cantidad de números primos que son menores o iguales que x. Simbólicamente podemos escribir

$$\pi(x) = \sum_{p \leq x} 1,$$

donde p recorre todos los primos menores o iguales que x. El teorema de los números primos afirma que

$$\lim_{x \to \infty} \frac{\pi(x) \ln x}{x} = 1. \tag{3.1}$$

Esto significa que la cantidad $\pi(x)$ se aproxima a $x/\ln x$ cuando x es muy grande. En la Figura 3.1 comparamos las gráficas de $\pi(x)$ y $x/\ln x$.

Ahora procedemos a probar la forma equivalente del teorema de los números primos que involucra la función ψ de Chebyshev.

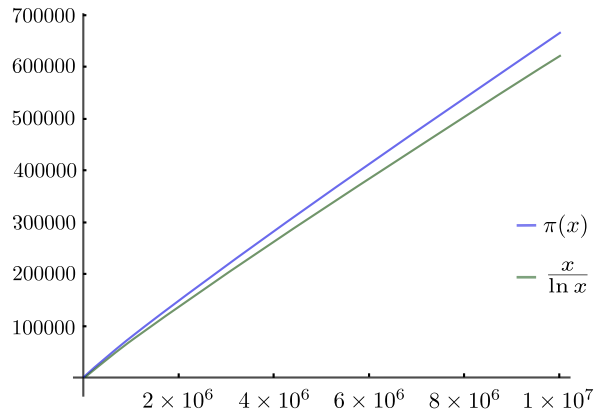

Figura 3.1: $\pi(x)$ versus $x/\ln x$

Recordemos que se define ψ como las sumas parciales de la función Λ de Mangoldt:

$$\psi(x) = \sum_{n \leq x} \Lambda(n).$$

Dado que $\Lambda(n) = 0$ a menos que n sea una potencia de primo, podemos extender $\psi(x)$ como sumas sobre potencias de primos, de la siguiente manera:

$$\begin{aligned}\psi(x) &= \sum_{p \leq x} \Lambda(p) + \sum_{p^2 \leq x} \Lambda(p^2) + \sum_{p^3 \leq x} \Lambda(p^3) + \cdots \\ &= \sum_{m=1}^{\infty} \sum_{p^m \leq x} \Lambda(p^m) = \sum_{m=1}^{\infty} \sum_{p \leq x^{1/m}} \ln p,\end{aligned}$$

Observemos que la suma sobre p es cero si $x^{1/m} < 2$, o equivalentemente, si $m > \log_2 x$. Por lo tanto

$$\begin{aligned}\psi(x) &= \sum_{m=1}^{\lfloor \log_2 x \rfloor} \sum_{p \leq x^{1/m}} \ln p = \sum_{p \leq x} \ln p + \sum_{m=2}^{\lfloor \log_2 x \rfloor} \sum_{p \leq x^{1/m}} \ln p \\ &\leq \sum_{p \leq x} \ln x + \sum_{m=2}^{\lfloor \log_2 x \rfloor} \sum_{p \leq x^{1/2}} \ln p \leq \ln x \sum_{p \leq x} 1 + \sum_{m=2}^{\lfloor \log_2 x \rfloor} \ln(x^{1/2}) \sum_{p \leq x^{1/2}} 1 \\ &= \pi(x) \ln x + \frac{1}{2} \pi(x^{1/2}) \ln x \sum_{m=2}^{\lfloor \log_2 x \rfloor} 1 \leq \pi(x) \ln x + \frac{1}{2} x^{1/2} \ln x \log_2 x \\ &= \pi(x) \ln x + \frac{x^{1/2} \ln^2 x}{2 \ln 2}\end{aligned}$$

donde se ha usado que $\pi(y) \leq y$ para todo $y \geq 2$ (*esto es claro, pues para cualquier real $y \geq 2$ es posible encontrar un entero menor o igual que $\lfloor y \rfloor$ y no es primo*). Los

cálculos anteriores muestran que

$$\psi(x) \leq \pi(x)\ln x + \frac{x^{1/2}\ln^2 x}{2\ln 2}. \tag{3.2}$$

Por otro lado, se sigue de los cálculos anteriores que

$$\psi(x) = \sum_{p\leq x}\Lambda(p) + \sum_{p^2\leq x}\Lambda(p^2) + \sum_{p^3\leq x}\Lambda(p^3) + \cdots \geq \sum_{p\leq x}\Lambda(p) = \sum_{p\leq x}\ln p.$$

Ahora, para x suficientemente grande

$$\begin{aligned}\psi(x) \geq \sum_{p\leq x}\ln p \geq \sum_{x/\ln^2 x\leq p\leq x}\ln p &\geq \ln\left(\frac{x}{\ln^2 x}\right)\sum_{x/\ln^2 x\leq p\leq x}1 \\ &\geq \ln\left(\frac{x}{\ln^2 x}\right)\left(\pi(x) - \pi\left(\frac{x}{\ln^2 x}\right)\right).\end{aligned}$$

Utilizando las propiedades del logaritmo, obtenemos

$$\ln\left(\frac{x}{\ln^2 x}\right)\left(\pi(x) - \pi\left(\frac{x}{\ln^2 x}\right)\right) \geq (\ln x - 2\ln(\ln x))\left(\pi(x) - \frac{x}{\ln^2 x}\right),$$

así que;

$$\psi(x) \geq (\ln x - 2\ln(\ln x))\left(\pi(x) - \frac{x}{\ln^2 x}\right).$$

Como $\ln x - 2\ln(\ln x) > 0$, para todo $x > 1$; podemos dividir la desigualdad anterior por $\ln x - 2\ln(\ln x)$ y multiplicarla por $\ln x$ para obtener

$$\frac{\psi(x)\ln x}{\ln x - 2\ln(\ln x)} \geq \pi(x)\ln x - \frac{x}{\ln x}. \tag{3.3}$$

Pero notemos que

$$\begin{aligned}\frac{\psi(x)\ln x}{\ln x - 2\ln(\ln x)} &= \frac{\psi(x)\ln x - 2\psi(x)\ln(\ln x) + 2\psi(x)\ln(\ln x)}{\ln x - 2\ln(\ln x)} \\ &= \frac{\psi(x)(\ln x - 2\ln(\ln x)) + 2\psi(x)\ln(\ln x)}{\ln x - 2\ln(\ln x)} \\ &= \psi(x) + \psi(x)\frac{2\ln(\ln x)}{\ln x - 2\ln(\ln x)},\end{aligned}$$

por lo tanto la desigualdad (3.3) puede ser reescrita como sigue:

$$\pi(x)\ln x \leq \psi(x) + \psi(x)\frac{2\ln(\ln x)}{\ln x - 2\ln(\ln x)} + \frac{x}{\ln x}. \tag{3.4}$$

Por otro lado, usando la regla de L'Hôpital, se muestra fácilmente que

$$\lim_{x\to\infty} \frac{2\ln(\ln x)}{(1/4)\ln x} = 0.$$

De esta manera podemos afirmar que, para x suficientemente grande, $2\ln(\ln x) \leq (1/4)\ln x$. Luego

$$\ln x - 2\ln(\ln x) \geq \ln x - (1/4)\ln x = (3/4)\ln x > 0.$$

Tomando inverso multiplicativo y multiplicando por $2\ln(\ln x)$ se obtiene la siguiente desigualdad

$$\frac{2\ln(\ln x)}{\ln x - 2\ln(\ln x)} \leq \frac{8}{3}\frac{\ln(\ln x)}{\ln x}.$$

Además, aplicando el Teorema 2.4.2 tenemos que para x suficientemente grande $\psi(x) < (3/2)x$. Usando este hecho y la desigualdad de arriba,

$$\psi(x)\frac{2\ln(\ln x)}{\ln x - 2\ln(\ln x)} \leq \left(\frac{3}{2}x\right)\left(\frac{8}{3}\frac{\ln(\ln x)}{\ln x}\right) = \frac{4x\ln(\ln x)}{\ln x}. \tag{3.5}$$

Finalmente, de las desigualdades (3.4) y (3.5) obtenemos que para x suficientemente grande

$$\pi(x)\ln x \leq \psi(x) + \frac{4x\ln(\ln x)}{\ln x} + \frac{x}{\ln x}. \tag{3.6}$$

Ya estamos listos para nuestra forma equivalente al teorema de los números primos en términos de $\psi(x)$.

Teorema 3.1.1. El teorema de los números primos es equivalente a

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

Demostración. El teorema estará probado si podemos mostrar que

$$\lim_{x\to\infty} \frac{\psi(x)}{x} = \lim_{x\to\infty} \frac{\pi(x)\ln x}{x}. \tag{3.7}$$

De las desigualdades (3.2) y (3.6) encontramos que, para x suficientemente grande

$$-\frac{4\ln(\ln x)+1}{\ln x} \leq \frac{\psi(x)}{x} - \frac{\pi(x)\ln x}{x} \leq \frac{\ln^2 x}{(2\ln 2)\sqrt{x}}.$$

Si hacemos que $x \to \infty$ en la desigualdad anterior y aplicamos la regla de L'Hôpital encontramos que $(4\ln(\ln x)+1)/(\ln x)$ y $(\ln^2 x)/(2\ln 2\sqrt{x})$ tienden a cero por lo que se cumple la igualdad (3.7). $\square$

3.2. La fórmula asintótica de Selberg

La fórmula de Selberg establece que

$$\psi(x)\ln x + \sum_{n\leq x}\psi\left(\frac{x}{n}\right)\Lambda(n) = 2x\ln x + O(x).$$

Esta sección está dedicada a la demostración de esta identidad. Para este propósito seguiremos las ideas desarrolladas por Tatuzawa e Iseki [5]. La fórmula asintótica de Selberg constituye el paso crucial de la demostración elemental del teorema de los números primos. Primero, probamos el siguiente resultado, que puede entenderse como una fórmula de inversión para funciones definidas en $(0,\infty)$.

Teorema 3.2.1. Sea $F:(0,\infty)\to\mathbb{R}$ una función. Si

$$G(x) = \ln x\sum_{n\leq x}F\left(\frac{x}{n}\right),$$

entonces

$$F(x)\ln x + \sum_{n\leq x}F\left(\frac{x}{n}\right)\Lambda(n) = \sum_{d\leq x}\mu(d)G\left(\frac{x}{d}\right).$$

Demostración. Dado que $\left\lfloor\frac{1}{n}\right\rfloor = 0$ para todo $n\geq 2$, $\displaystyle\sum_{n\leq x}\left\lfloor\frac{1}{n}\right\rfloor F\left(\frac{x}{n}\right)\ln\frac{x}{n} = F(x)\ln x$. Además, del Teorema 1.1.1, $\displaystyle\left\lfloor\frac{1}{n}\right\rfloor = \sum_{d|n}\mu(d)$, luego tenemos que;

$$F(x)\ln x = \sum_{n\leq x}\left\lfloor\frac{1}{n}\right\rfloor F\left(\frac{x}{n}\right)\ln\left(\frac{x}{n}\right) = \sum_{n\leq x}F\left(\frac{x}{n}\right)\sum_{d|n}\mu(d)\ln\left(\frac{x}{n}\right).$$

Por el Teorema 1.4.1 sabemos que $\sum_{d|n}\Lambda(d) = \ln n$. Aplicando la fórmula de inversión de Möbius a esta igualdad, obtenemos

$$\Lambda(n) = \sum_{d|n}\mu(d)\ln\left(\frac{n}{d}\right),$$

de donde

$$\sum_{n\leq x}F\left(\frac{x}{n}\right)\Lambda(n) = \sum_{n\leq x}F\left(\frac{x}{n}\right)\sum_{d|n}\mu(d)\ln\left(\frac{n}{d}\right).$$

Ahora notemos que

$$\begin{aligned}
F(x)\ln x + \sum_{n\le x} F\left(\frac{x}{n}\right)\Lambda(n) &= \sum_{n\le x} F\left(\frac{x}{n}\right)\ln\left(\frac{x}{n}\right)\sum_{d|n}\mu(d) + \sum_{n\le x} F\left(\frac{x}{n}\right)\sum_{d|n}\mu(d)\ln\left(\frac{n}{d}\right)\\
&= \sum_{n\le x} F\left(\frac{x}{n}\right)\sum_{d|n}\mu(d)\left\{\ln\left(\frac{x}{n}\right)+\ln\left(\frac{n}{d}\right)\right\}\\
&= \sum_{n\le x}\sum_{d|n} F\left(\frac{x}{n}\right)\mu(d)\ln\left(\frac{x}{d}\right).
\end{aligned}$$

Escribiendo $n = qd$ notamos que tener $n \le x$ y $d \mid n$ es equivalente a tener $d \le x$ y $q \le x/d$. Así,

$$\sum_{n\le x}\sum_{d|n} F\left(\frac{x}{n}\right)\mu(d)\ln\left(\frac{x}{d}\right) = \sum_{d\le x}\mu(d)\ln\left(\frac{x}{d}\right)\sum_{q\le x/d} F\left(\frac{x}{qd}\right) = \sum_{d\le x}\mu(d)G\left(\frac{x}{d}\right).$$

Esto termina la prueba del teorema. □

Teorema 3.2.2 (Fórmula asintótica de Selberg)**.** Si $x > 0$, entonces

$$\psi(x)\ln x + \sum_{n\le x}\psi\left(\frac{x}{n}\right)\Lambda(n) = 2x\ln x + O(x).$$

Demostración. Inicialmente notemos que tener un para (n, d) tal que $n \le x$ y $d \le x/n$ es equivalente a tener un par (n, d) tal que $n \le x$ y $d \mid n$. Aplicando esto junto con el Teorema 1.4.1 encontramos que

$$\ln x\sum_{n\le x}\psi\left(\frac{x}{n}\right) = \ln x\sum_{n\le x}\sum_{d\le x/n}\Lambda(d) = \ln x\sum_{n\le x}\sum_{d|n}\Lambda(d) = \ln x\sum_{n\le x}\ln n,$$

pero, del Teorema 2.4.1 se sigue que, $\ln x\sum_{n\le x}\ln n = x\ln^2 x - x\ln x + O(\ln^2 x)$. Luego,

$$\ln x\sum_{n\le x}\psi\left(\frac{x}{n}\right) = x\ln^2 x - x\ln x + O(\ln^2 x). \tag{3.8}$$

Por otro lado, aplicando el Teorema 2.2.2 tenemos

$$\begin{aligned}
\ln x\sum_{n\le x}\left(\frac{x}{n}-\gamma-1\right) &= x\ln x\sum_{n\le x}\frac{1}{n} - (\gamma+1)\ln x\sum_{n\le x}1\\
&= x\ln x\left(\ln x+\gamma+O\left(\frac{1}{x}\right)\right) - (\gamma+1)\ln x\,\lfloor x\rfloor\\
&= x\ln^2 x + \gamma x\ln x + O(\ln x) - (\gamma+1)\ln x\,(x+O(1))\\
&= x\ln^2 x + \gamma x\ln x + O(\ln x) - \gamma x\ln x - x\ln x + O(\ln x)\\
&= x\ln^2 x - x\ln x + O(\ln x),
\end{aligned}$$

es decir,

$$\ln x \sum_{n\leq x}\left(\frac{x}{n}-\gamma-1\right)=x\ln^2 x-x\ln x+O(\ln x). \tag{3.9}$$

Comparando las estimaciones (3.8) y (3.9), se obtiene que

$$\ln x \sum_{n\leq x}\left\{\psi\left(\frac{x}{n}\right)-\frac{x}{n}+\gamma+1\right\}=O\left(\ln^2 x\right). \tag{3.10}$$

Tomando $G(x)=\ln x\sum_{n\leq x}\{\psi(x/n)-x/n+\gamma+1\}$ y notando que $\ln^2 x=O(\sqrt{x})$. La ecuación (3.10) implica que

$$G(x)=O\left(\sqrt{x}\right). \tag{3.11}$$

Aplicando el Teorema 3.2.1 con $F(x)=\psi(x)-x+\gamma+1$ tenemos

$$\{\psi(x)-x+\gamma+1\}\ln x+\sum_{n\leq x}\left\{\psi\left(\frac{x}{n}\right)-\frac{x}{n}+\gamma+1)\right\}\Lambda(n)=\sum_{d\leq x}\mu(d)G\left(\frac{x}{d}\right),$$

pero de (3.11) se sigue que

$$\sum_{d\leq x}\mu(d)G\left(\frac{x}{d}\right)=O\left(\sqrt{x}\sum_{d\leq x}\frac{1}{\sqrt{d}}\right)=O\left(\sqrt{x}\int_0^x\frac{1}{\sqrt{t}}dt\right)=O(x),$$

por lo tanto

$$\{\psi(x)-x+\gamma+1\}\ln x+\sum_{n\leq x}\left\{\psi\left(\frac{x}{n}\right)-\frac{x}{n}+\gamma+1)\right\}\Lambda(n)=O(x),$$

ordenando términos y aplicando el Teorema 2.4.3 resulta

$$\begin{aligned}\psi(x)\ln x+\sum_{n\leq x}\psi\left(\frac{x}{n}\right)\Lambda(n)&=(x-\gamma-1)\ln x+\sum_{n\leq x}\left(\frac{x}{n}-\gamma-1\right)\Lambda(n)+O(x)\\&=x\ln x+x\sum_{n\leq x}\frac{\Lambda(n)}{n}-(\gamma+1)\sum_{n\leq x}\Lambda(n)+O(x)\\&=x\ln x+x(\ln x+O(1))+O(x)\\&=2x\ln x+O(x).\end{aligned}$$

Esto termina la prueba. □

3.3. Estrategia de la demostración

Esta sección está dedicada a describir los pasos principales que haremos, a partir de la fórmula de Selberg, para realizar nuestra demostración elemental del teorema de los números primos. Los detalles completos los veremos en secciones posteriores.

En la sección anterior demostramos la fórmula de Selberg

$$\psi(x)\ln x + \sum_{n\leq x}\psi\left(\frac{x}{n}\right)\Lambda(n) = 2x\ln x + O(x),$$

que es nuestro punto de partida. Lo primero que haremos será definir una función S mediante

$$S(y) = \int_2^y \frac{\psi(t)-t}{t}dt,$$

para todo $y \geq 2$, y $S(y) = 0$, para $y < 2$. Resulta que si mostramos que

$$\lim_{y\to\infty}\frac{S(y)}{y} = 0, \tag{3.12}$$

entonces se sigue el teorema de los números primos. Así, nuestro objetivo se convierte en mostrar (3.12).

A partir de la fórmula de Selberg, usando algunas propiedades de S, derivaremos la siguinete versión de la fórmula de Selberg para la función S:

$$S(y)\ln y + \sum_{n\leq y}\Lambda(n)S\left(\frac{y}{n}\right) = O(y).$$

Luego, con un poco de esfuerzo y la ayuda de la fórmula de sumación de Abel probaremos la desigualdad siguiente:

$$|S(y)|\ln^2 y \leq 2\sum_{m\leq y}|S(y/m)|\ln m + K_1 y\ln y,$$

de la cual derivaremos una "versión integral" de la misma:

$$|S(y)|\ln^2 y \leq 2\int_2^y |S(y/u)|\ln u du + K_2 y\ln y.$$

El paso siguiente es definir $W(x) = e^{-x}S(e^x)$, y la desigualdad anterior lleva a la siguiente:

$$|W(x)| \leq \frac{2}{x^2}\int_0^x\left(\int_0^u |W(v)|dv\right)du + \frac{K_2}{x}.$$

Además, por la definición de $W(x)$ notamos que basta mostrar que $W(x) \to 0$, cuando $x \to \infty$; para concluir que $S(y)/y \to 0$, cuando $y \to \infty$. Así, nuestro objetivo se transforma en probar que $\lim_{x\to\infty} W(x) = 0$, y para lograr esto, lo que demostramos como paso final es que

$$\alpha := \limsup_{x\to\infty} |W(x)| = 0,$$

utilizando propiedades especiales de la función $|W(x)|$ que probaremos en detalle más adelante.

3.4. Demostración del teorema de los números primos

En esta sección completamos los detalles de los pasos descritos en la sección anterior. Dado lo extenso que pueden ser tales detalles, dividimos esta sección en varias partes correspondientes a los diferentes pasos descritos en la sección anterior.

3.4.1. La función $S(y)$

Definición 3.4.1. Definimos la función $S : (0, \infty) \to \mathbb{R}$ como sigue:

$$S(y) = \int_2^y \frac{\psi(t) - t}{t} dt$$

siempre que $y \geq 2$, y para $y < 2$ definimos $S(y) = 0$.

El siguiente teorema establece que, (3.12) implica el teorema de los números primos.

Teorema 3.4.1. Si $\lim\limits_{y\to\infty} \dfrac{S(y)}{y} = 0$, entonces $\lim\limits_{y\to\infty} \dfrac{\psi(y)}{y} = 1$.

Demostración. Supongamos que (3.12) se cumple y sea $\varepsilon > 0$ dado. Para y suficientemente grande tenemos

$$|S(y)| \leq \frac{\varepsilon^2}{2(1+\varepsilon)} y.$$

Ahora, notemos que

$$S(y(1+\varepsilon)) - S(y) \leq \frac{\varepsilon^2}{2}y + \frac{\varepsilon^2}{2(1+\varepsilon)}y \leq \varepsilon^2 y,$$

que por la definición de S, vemos que es equivalente a

$$\int_y^{y(1+\varepsilon)} \frac{\psi(t)-t}{t}dt \leq \varepsilon^2 y. \tag{3.13}$$

Usando el hecho de que la función ψ es creciente, para $y \leq t \leq y(1+\varepsilon)$, podemos construir la siguiente desigualdad

$$\frac{\psi(y)}{y(1+\varepsilon)} \leq \frac{\psi(t)}{t}.$$

Esto junto con (3.13) muestran que

$$\frac{\psi(y)}{y(1+\varepsilon)}\int_y^{y(1+\varepsilon)} dt - \int_y^{y(1+\varepsilon)} dt \leq \varepsilon^2 y,$$

resolviendo estas integrales se obtiene que $\dfrac{\psi(y)}{y(1+\varepsilon)}y\varepsilon - y\varepsilon \leq \varepsilon^2 y$, o equivalentemente,

$$\frac{\psi(y)}{y(1+\varepsilon)}y\varepsilon \leq y\varepsilon(1+\varepsilon),$$

de donde $\psi(y)/y \leq (1+\varepsilon)^2$.

Un razonamiento análogo al que acabamos de hacer muestra que para y suficientemente grande se tiene $S(y) - S(y(1-\varepsilon)) \geq -\varepsilon^2 y$, lo cual conduce a $\psi(y)/y \geq (1-\varepsilon)^2$. Como ε fue tomado de forma arbitraria, concluimos que $\lim_{y\to\infty} \psi(y)/y = 1$. □

Dada la definición, tenemos claramente que S es una función continua. El siguiente resultado muestra, además, que esta continuidad es uniforme.

Teorema 3.4.2. Existe una constante c tal que $|S(y_2) - S(y_1)| \leq c|y_2 - y_1|$, para todo $y_1, y_2 > 0$.

Demostración. Del Teorema 2.4.2 se sigue que $-x < \psi(x) - x < \frac{1}{2}x$ para x suficientemente grande. En consecuencia

$$\limsup_{x\to\infty} \frac{|\psi(x) - x|}{x} \leq 1. \tag{3.14}$$

Fijemos un real $\lambda_1 > 1$. Entonces, existe un x_1 tal que $|\psi(x) - x| \leq \lambda_1 x$ para todo $x \geq x_1$. Además, es claro que para $0 < x < x_1$ existe una constante λ_2 tal que $|\psi(x) - x| \leq \lambda_2 x$. Tomando $c = \text{máx}\{\lambda_1, \lambda_2\}$ encontramos que para todo $x > 0$,

$$|\psi(x) - x| \leq cx. \tag{3.15}$$

Por otro lado, notemos que $S'(y) = (\psi(y) - y)/y$, excepto cuando y es una potencia de primo, puesto que allí $\psi(y)$ es discontinua. Así, la desigualdad (3.15) implica que

$$|S'(y)| \leq c, \tag{3.16}$$

siempre que y no sea una potencia de primo.

Ahora procedemos a demostrar la desigualdad requerida. Sean $y_1, y_2 > 0$. Si tuviéramos que $y_1 = y_2$, el resultado es claro. Supongamos que $y_1 \neq y_2$, y sin perdida de generalidad, asumamos que $y_1 < y_2$.

En el caso en que el intervalo abierto (y_1, y_2) no contenga ninguna potencia de primo, S es diferenciable en (y_1, y_2), y, por T.V.M existe $y \in (y_1, y_2)$ tal que

$$|S(y_2) - S(y_1)| = |S'(y)||y_2 - y_1| \leq c|y_2 - y_1|.$$

De esta forma, en todo intervalo cerrado $[y_1, y_2]$ que no contenga potencias de primos en su interior, vale $|S(y_2) - S(y_1)| \leq c|y_2 - y_1|$.

Ahora, para el caso general, consideramos una partición $x_0, x_1, \ldots, x_n$ del intervalo $[y_1, y_2]$ de tal forma que los x_i con $0 < i < n$ sean precisamente las potencias de primos que están en el interior de $[y_1, y_2]$ (si es que las hay). De esta forma, para todo $i = 1, 2, \ldots, n$, los intervalos abiertos (x_{i-1}, x_i) no contienen potencias de primos. Luego

$$|S(y_2) - S(y_1)| \leq \sum_{i=1}^{n} |S(x_i) - S(x_{i-1})| \leq \sum_{i=1}^{n} c|x_i - x_{i-1}| = c|y_2 - y_1|.$$

De esta manera queda probado el teorema. □

Corolario 3.4.3. Existe una constante c tal que $|S(y)| \leq cy$, para todo $y > 0$. En particular, $S(y) = O(y)$.

Demostración. Por el Teorema 3.4.2, existe una constante c tal que

$$|S(y_2) - S(y_1)| \leq c|y_2 - y_1|,$$

para todo $y_1, y_2 > 0$. Supongamos entonces que $y > 2$. Tomando $y_1 = 2$ e $y_2 = y$, $|S(y)| = |S(y) - S(2)| \leq c|y - 2| = c(y - 2) \leq cy$. Para $y \leq 2$ la desigualdad se cumple trivialmente, pues $S(y) = 0$. □

3.4.2. Fórmulas asintóticas y desigualdades que involucran a la función $S(y)$

En esta sección presentamos algunas fórmulas asintóticas para la función $S(y)$ derivadas a partir de la identidad de Selberg. Lo primero que probaremos es la fórmula asintótica

$$(\psi(x) - x)\ln x + \sum_{n \leq x} \Lambda(n)\left(\psi\left(\frac{x}{n}\right) - \frac{x}{n}\right) = O(x). \tag{3.17}$$

Para probarla, comenzamos observando que

$$\begin{aligned}
\sum_{n \leq x} \Lambda(n)\left(\psi\left(\frac{x}{n}\right) - \frac{x}{n}\right) &= \sum_{n \leq x} \Lambda(n)\psi\left(\frac{x}{n}\right) - x\sum_{n \leq x} \frac{\Lambda(n)}{n} \\
&= \sum_{n \leq x} \Lambda(n)\psi\left(\frac{x}{n}\right) - x(\ln x + O(1)) \\
&= \sum_{n \leq x} \Lambda(n)\psi\left(\frac{x}{n}\right) - x\ln x + O(x),
\end{aligned}$$

por lo que

$$\sum_{n \leq x} \Lambda(n)\psi\left(\frac{x}{n}\right) = \sum_{n \leq x} \Lambda(n)\left(\psi\left(\frac{x}{n}\right) - \frac{x}{n}\right) + x\ln x + O(x).$$

Sustituyendo esta última igualdad en la fórmula de Selberg, y agrupando términos, nos queda la identidad deseada

$$(\psi(x) - x)\ln x + \sum_{n \leq x} \Lambda(n)\left(\psi\left(\frac{x}{n}\right) - \frac{x}{n}\right) = O(x).$$

Por otro lado, la integración por partes aplicada a la integral $\int_2^y \frac{S(x)}{x}dx$ conduce a

$$\int_2^y \frac{\psi(x) - x}{x}\ln x\ dx = S(y)\ln y - S(2)\ln 2 - \int_2^y \frac{S(x)}{x}dx,$$

como $S(2) = 0$ y $S(x) = O(x)$, se sigue la identidad

$$\int_2^y \frac{\psi(x) - x}{x} \ln x \; dx = S(y) \ln y + O(y). \tag{3.18}$$

Utilizando el cambio de variable $z = x/n$, obtenemos

$$\int_2^y \left(\psi\left(\frac{x}{n}\right) - \frac{x}{n} \right) \frac{dx}{x} = \int_2^{y/n} \frac{\psi(z) - z}{z} dz = S(y/n). \tag{3.19}$$

Si ahora dividimos ambos lados de (3.17) por x e integramos desde 2 hasta y con respecto a x obtenemos

$$\int_2^y \frac{\psi(x) - x}{x} \ln x dx + \sum_{n \leq x} \Lambda(n) \int_2^y \left(\psi\left(\frac{x}{n}\right) - \frac{x}{n} \right) \frac{dx}{x} = O(y),$$

combinando esto con (3.18) y (3.19), se sigue la siguiente fórmula asintótica para $S(y)$:

$$S(y) \ln y + \sum_{n \leq y} \Lambda(n) S\left(\frac{y}{n}\right) = O(y). \tag{3.20}$$

Para el siguiente paso, necesitaremos la *identidad de Abel*.

Teorema 3.4.4 (Identidad de Abel)**.** Para cualquier función aritmética $a(n)$, sea

$$A(x) = \sum_{n \leq x} a(n),$$

donde $A(x) = 0$, para $0 < x < 1$. Asuma que f tiene derivada continua en el intervalo $[y, x]$, donde $0 < y < x$. Entonces

$$\sum_{y < n \leq x} a(n) f(n) = A(x) f(x) - A(y) f(y) - \int_y^x A(t) f'(t) \; dt. \tag{3.21}$$

Demostración. Sean $k = \lfloor x \rfloor$ y $m = \lfloor y \rfloor$, así que $A(x) = A(k)$ y $A(y) = A(m)$.

Entonces

$$\begin{aligned}
\sum_{y<n\leq x} a(n)f(n) &= \sum_{n=m+1}^{k} a(n)f(n) = \sum_{n=m+1}^{k} [A(n)-A(n-1)]f(n) \\
&= \sum_{n=m+1}^{k} A(n)f(n) - \sum_{n=m}^{k-1} A(n)f(n+1) \\
&= \sum_{n=m+1}^{k-1} A(n)[f(n)-f(n+1)] + A(k)f(k) - A(m)f(m+1) \\
&= -\sum_{n=m+1}^{k-1} A(n)\int_n^{n+1} f'(t)\ dt + A(k)f(k) - A(m)f(m+1) \\
&= -\sum_{n=m+1}^{k-1} \int_n^{n+1} A(t)f'(t)\ dt + A(k)f(k) - A(m)f(m+1) \\
&= -\int_{m+1}^{k} A(t)f'(t)\ dt + A(x)f(x) - \int_k^x A(t)f'(t)\ dt \\
&\quad - A(y)f(y) - \int_y^{m+1} A(t)f'(t)\ dt \\
&= A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)\ dt.
\end{aligned}$$

De esta forma se termina la prueba del teorema. □

Teorema 3.4.5. Si $C_n = \Lambda(n)\ln n + \sum_{mk=n}\Lambda(m)\Lambda(k)$, entonces

$$\sum_{n\leq x}(C_n - 2\ln n) = O(x).$$

Demostración. Tomando $a(n) = \Lambda(n)$ y $f(x) = \ln x$ en la identidad de Abel (con $y < 1$), se tiene $A(x) = \psi(x)$ y obtenemos

$$\sum_{n\leq x}\Lambda(n)\ln n = \psi(x)\ln x - \int_1^x \frac{\psi(t)}{t}dt.$$

Reordenando términos y recordando que $\psi(t) = O(t)$ por el Teorema 2.4.2, obtenemos

$$\psi(x)\ln x = \sum_{n\leq x}\Lambda(n)\ln n + \int_1^x \frac{\psi(t)}{t}dt = \sum_{n\leq x}\Lambda(n)\ln n + O(x). \tag{3.22}$$

Por otro lado, notemos que

$$\sum_{m\leq x}\Lambda(m)\psi\left(\frac{x}{m}\right) = \sum_{m\leq x}\Lambda(m)\sum_{k\leq x/m}\Lambda(k) = \sum_{n\leq x}\sum_{mk=n}\Lambda(m)\Lambda(k). \tag{3.23}$$

Reemplazamos (3.22) y (3.23) en la fórmula de Selberg y agrupando términos, obtenemos

$$\sum_{n\leq x}\left\{\Lambda(n)\ln n + \sum_{mk=n}\Lambda(m)\Lambda(k)\right\} = 2x\ln x + O(x),$$

por lo tanto, $\sum_{n\leq x} C_n = 2x\ln x + O(x)$. Además, del Teorema 2.4.1 se sabe que $\sum_{n\leq x}\ln n = x\ln x + O(x)$. Luego

$$\sum_{n\leq x}(C_n - 2\ln n) = \sum_{n\leq x} C_n - 2\sum_{n\leq x}\ln n = 2x\ln x + O(x) - 2(x\ln x + O(x)) = O(x),$$

esto termina la prueba. □

Antes de nuestro siguiente resultado, analizamos un poco la función

$$T(x) = \sum_{n\leq x}(C_n - 2\ln n).$$

Reemplacemos x por un entero $m \geq 2$ y observemos que $T(m) - T(m-1) = C_m - 2\ln m$. Usando esto encontramos que

$$\begin{aligned}
\left|\sum_{m\leq x}(C_m - 2\ln m)\left|S\left(\frac{x}{m}\right)\right|\right| &= \left|\sum_{m\leq x}(T(m) - T(m-1))\left|S\left(\frac{x}{m}\right)\right|\right| \\
&= \left|\sum_{m\leq x} T(m)\left|S\left(\frac{x}{m}\right)\right| - \sum_{m\leq x} T(m-1)\left|S\left(\frac{x}{m}\right)\right|\right| \\
&= \left|\sum_{m\leq x} T(m)\left|S\left(\frac{x}{m}\right)\right| - \sum_{m\leq x} T(m)\left|S\left(\frac{x}{m+1}\right)\right|\right| \\
&= \left|\sum_{m\leq x} T(m)\left(\left|S\left(\frac{x}{m}\right)\right| - \left|S\left(\frac{x}{m+1}\right)\right|\right)\right| \\
&\leq \sum_{m\leq x}|T(m)|\left|\left|S\left(\frac{x}{m}\right)\right| - \left|S\left(\frac{x}{m+1}\right)\right|\right| \\
&\leq \sum_{2\leq m\leq x}|T(m)|\left|S\left(\frac{x}{m}\right) - S\left(\frac{x}{m+1}\right)\right|,
\end{aligned}$$

donde se ha tomado $2 \leq m \leq x$ debido a que $S(x) = 0$ para todo $x < 2$ (en el último paso hemos usado la desigualdad $||a| - |b|| \leq |a - b|$, para números reales a y b). Aplicando los Teoremas 3.4.2 y 3.4.5 podemos elegir una constante K de tal forma que

$$|T(m)|\left|S\left(\frac{x}{m}\right) - S\left(\frac{x}{m+1}\right)\right| \leq Km\left(\frac{x}{m} - \frac{x}{m+1}\right) = \frac{Kx}{m+1}.$$

Retomando las cálculos anteriores podemos ver que

$$\left|\sum_{m\leq x}(C_m - 2\ln m)\left|S\left(\frac{x}{m}\right)\right|\right| \leq Kx\sum_{2\leq m\leq x}\left(\frac{1}{m+1}\right) \leq Kx\int_1^x \frac{du}{u} \leq Kx\ln x.$$

Hemos probado así la fórmula asintótica siguiente:

$$\sum_{m\leq x}(C_m - 2\ln m)\left|S\left(\frac{x}{m}\right)\right| = O(x\ln x). \tag{3.24}$$

Teorema 3.4.6. Existe una constante K_1 tal que para y suficientemente grande

$$|S(y)|\ln^2 y \leq 2\sum_{m\leq y}|S(y/m)|\ln m + K_1 y\ln y. \tag{3.25}$$

Demostración. Por (3.20), tenemos $S(y)\ln y + \sum_{n\leq y}\Lambda(n)S(y/n) = O(y)$. En esta fórmula reemplazamos y por y/k, y luego multiplicamos por $\Lambda(k)$ para obtener

$$\Lambda(k)S\left(\frac{y}{k}\right)\ln\left(\frac{y}{k}\right) + \sum_{n\leq y/k}\Lambda(k)\Lambda(n)S\left(\frac{y}{nk}\right) = \Lambda(k)O\left(\frac{y}{k}\right).$$

Las propiedades del logaritmo y la notación O nos conducen a

$$\Lambda(k)S\left(\frac{y}{k}\right)\ln y - \Lambda(k)S\left(\frac{y}{k}\right)\ln k + \sum_{n\leq y/k}\Lambda(k)\Lambda(n)S\left(\frac{y}{nk}\right) = O\left(y\frac{\Lambda(k)}{k}\right).$$

Tomando sumas parciales en la fórmula anterior y observando que del Teorema 2.4.3, $O\left(y\sum_{k\leq y}\frac{\Lambda(k)}{k}\right) = O\left(y(\ln y + O(1))\right) = O\left(yO(\ln y)\right) = O\left(O(y\ln y)\right) = O(y\ln y)$, obtenemos que

$$\ln y\sum_{k\leq y}\Lambda(k)S\left(\frac{y}{k}\right) - \sum_{k\leq y}\Lambda(k)S\left(\frac{y}{k}\right)\ln k + \sum_{k\leq y}\sum_{n\leq y/k}\Lambda(k)\Lambda(n)S\left(\frac{y}{nk}\right) = O(y\ln y).$$

Haciendo $nk = m$, notamos que tener un par (n,k) tal que $k\leq y$ y $n\leq y/k$ es equivalente a tener un par (m,k) tal que $m\leq y$ y $nk = m$. Así, la fórmula anterior se convierte en

$$\ln y\sum_{m\leq y}\Lambda(m)S\left(\frac{y}{m}\right) - \sum_{m\leq y}\Lambda(m)S\left(\frac{y}{m}\right)\ln m + \sum_{m\leq y}\sum_{nk=m}\Lambda(k)\Lambda(n)S\left(\frac{y}{m}\right) = O(y\ln y),$$

donde se ha cambiado el subíndice k por m en las primeras sumas parciales. De la fórmula (3.20) se sigue que $\sum_{m\leq y}\Lambda(m)S\left(y/m\right) = -S(y)\ln y + O(y)$; reemplazando esto en la anterior desigualdad y agrupado términos obtenemos la siguiente fórmula

$$S(y)\ln^2 y + \sum_{m\leq y}S\left(\frac{y}{m}\right)\left\{\Lambda(m)\ln m - \sum_{nk=m}\Lambda(n)\Lambda(k)\right\} = O(y\ln y).$$

Tomando $A_m = \Lambda(m)\ln m - \sum_{nk=m}\Lambda(n)\Lambda(k)$ y usando que $||a|-|a|| \leq |a+b|$ para todo $a, b \in \mathbb{R}$, se sigue de la definición de O grande que existe una constante K tal que para y suficientemente grande

$$\left|\left|S(y)\ln^2 y\right| - \left|\sum_{m\leq y} S(y/m)A_m\right|\right| \leq \left|S(y)\ln^2 y + \sum_{m\leq y} S(y/m)A_m\right| \leq Ky\ln y.$$

En consecuencia

$$|S(y)|\ln^2 y - \left|\sum_{m\leq y} S(y/m)A_m\right| \leq Ky\ln y. \tag{3.26}$$

Ahora, dado que los términos $\Lambda(m)\ln m$ y $\sum_{nk=m}\Lambda(n)\Lambda(k)$ son positivos tenemos que $C_m = \Lambda(m)\ln m + \sum_{nk=m}\Lambda(n)\Lambda(k)$ es positivo, y usando nuevamente la desigualdad $||a|-|b|| \leq |a+b|$, resulta que $|A_m| \leq C_m$. Luego

$$\left|\sum_{m\leq y} A_m S(y/m)\right| \leq \sum_{m\leq y} |A_m S(y/m)| \leq \sum_{m\leq y} C_m |S(y/m)|,$$

por lo que $|S(y)|\ln^2 y - |\sum_{m\leq y} S(y/m)\Lambda_m| \geq |S(y)|\ln^2 y - \sum_{m\leq y} C_m|S(y/m)|$, esto combinado con (3.26) muestran que

$$|S(y)|\ln^2 y \leq \sum_{m\leq y} C_m|S(y/m)| + Ky\ln y. \tag{3.27}$$

Por otro lado, de (3.24) sabemos que existe una constante K' tal que

$$\sum_{m\leq y} C_m|S(y/m)| \leq 2\sum_{m\leq y}|S(y/m)|\ln m + K'y\ln y$$

para y suficientemente grande. Usando este hecho junto con (3.27), terminamos la prueba tomando $K_1 = K + K'$. □

Del teorema anterior deducimos ahora la siguiente desigualdad que involucra a $S(y)$.

Teorema 3.4.7. Existe una constante K_2 tal que para y suficientemente grande,

$$|S(y)|\ln^2 y \leq 2\int_2^y |S(y/u)|\ln u du + K_2 y\ln y. \tag{3.28}$$

Demostración. Comenzamos con la Figura 3.2.

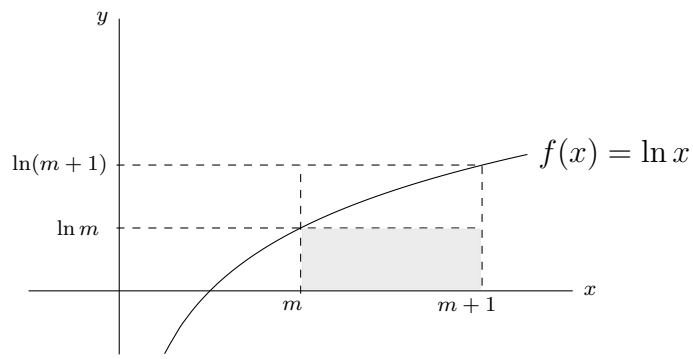

Figura 3.2: Área bajo el logaritmo

El área de la región sombreada (rectángulo) es menor que el área bajo la curva $f(x) = \ln x$ que está entre $x = m$ y $x = m+1$. Esto significa que $\ln m \leq \int_m^{m+1} \ln u \; du$. Luego

$$|S(y/m)| \ln m \leq \int_m^{m+1} |S(y/m)| \ln u \; du.$$

Usando la desigualdad triangular, $|S(y/m)| \leq |S(y/u)| + |S(y/m) - S(y/u)|$, así que

$$|S(y/m)| \ln m \leq \int_m^{m+1} |S(y/u)| \ln u \; du + J_m, \tag{3.29}$$

donde

$$J_m = \int_m^{m+1} |S(y/m) - S(y/u)| \ln u \; du.$$

De la Figura 3.2 también podemos ver que el área de la región rectangular con altura $\ln(m+1)$ que se encuentra entre $x = m$ y $x = m+1$ es mayor que el área bajo la curva $f(x) = \ln x$ que está entre $x = m$ y $x = m + 1$, esto es, $\int_m^{m+1} \ln u \; du \leq \ln(m + 1)$. Además, del Teorema 3.4.2 y el hecho de que $\ln(m + 1) \leq m$ podemos elegir una constante c tal que

$$J_m \leq c \left(\frac{y}{m} - \frac{y}{m+1} \right) \int_m^{m+1} \ln u \; du \leq \frac{cy \ln(m+1)}{m(m+1)} \leq \frac{cy}{m+1}.$$

Usando este hecho en (3.29) tenemos

$$|S(y/m)| \ln m \leq \int_m^{m+1} |S(y/m)| \ln u \; du + \frac{cy}{m+1}.$$

Aplicando esta desigualdad en (3.25) nos da lo siguiente:

$$\begin{aligned}
|S(y)|\ln^2 y &\leq 2\sum_{m\leq y}\left(\int_m^{m+1}|S(y/u)|\ln u\ du+\frac{cy}{m+1}\right)+K_1 y\ln y\\
&= 2\sum_{m\leq y}\int_m^{m+1}|S(y/u)|\ln u\ du+2cy\sum_{m\leq y}\frac{1}{m+1}+K_1 y\ln y\\
&\leq 2\int_2^y|S(y/u)|\ln u\ du+2cy\int_1^y\frac{du}{u}+K_1 y\ln y\\
&\leq 2\int_2^y|S(y/u)|\ln u\ du+2cy\ln y+K_1 y\ln y\\
&= 2\int_2^y|S(y/u)|\ln u\ du+K_2 y\ln y,
\end{aligned}$$

donde se ha tomado $K_2 = K_1 + 2c$. Esto termina la prueba del teorema. □

3.4.3. Reescalamiento logarítmico

En esta sección, continuando desde el Teorema 3.4.7, encontraremos nuevas fórmulas asintóticas que nos llevarán más cerca de nuestra prueba para el teorema de los números primos. Definamos $W(x) = e^{-x}S(e^x)$, para todo número real $x > 0$. Notemos que probar que $\lim_{y\to\infty} S(y)/y = 0$, lo que tenemos que probar es que $\lim_{x\to\infty} W(x) = 0$; así, nuestra prueba del teorema de los números primos estará hecha cuando probemos que $\limsup_{x\to\infty}|W(x)| = 0$.

Haciendo el cambio de variables $v = \ln(y/u)$, y tomando $x = \ln y$ en la desigualdad (3.28), obtenemos

$$x^2|S(e^x)| \leq 2\int_0^{x-\ln 2}|S(e^v)|(x-v)e^{x-v}dv + K_2 x e^x.$$

Dividiendo por x^2e^x y sustituyendo $e^{-x}S(e^x)$ por $W(x)$, llegamos a lo siguiente:

$$|W(x)| \leq \frac{2}{x^2}\int_0^{x-\ln 2}(x-v)|W(v)|dv + \frac{K_2}{x} \leq \frac{2}{x^2}\int_0^x\int_v^x|W(v)|dudv + \frac{K_2}{x}.$$

Cambiamos el orden de integración:

$$\int_0^x\int_v^x|W(v)|dudv = \int_0^u\int_0^x|W(v)|dudv = \int_0^x\left(\int_0^u|W(v)|dv\right)du,$$

para obtener las *desigualdad integral* siguiente

$$|W(x)| \leq \frac{2}{x^2}\int_0^x\left(\int_0^u|W(v)|dv\right)du + \frac{K_2}{x}. \tag{3.30}$$

Teorema 3.4.8. Sean

$$\alpha := \limsup_{x\to\infty} |W(x)|, \quad \beta := \limsup_{x\to\infty} \frac{1}{x}\int_0^x |W(v)|dv.$$

Entonces $\alpha \leq 1$ y $\alpha \leq \beta$.

Demostración. Primero veamos que $\alpha \leq 1$. De (3.14) sabemos que:

$$\limsup_{x\to\infty} \frac{|\psi(x)-x|}{x} \leq 1.$$

Sea $r > 0$. Entonces $\limsup_{x\to\infty} \dfrac{|\psi(x)-x|}{x} < 1+r$, y por lo tanto existe un x_r tal que $\dfrac{|\psi(x)-x|}{x} \leq 1+r$ siempre que $x \geq x_r$. En consecuencia $\displaystyle\int_{x_r}^{y} \frac{|\psi(x)-x|}{x}dx \leq (1+r)y$ para todo $y \geq x_r$. Usando este hecho, para y suficientemente grande, se obtiene lo siguiente:

$$\begin{aligned}|S(y)| \leq \int_2^y \frac{|\psi(x)-x|}{x}dx &= \int_2^{x_r} \frac{|\psi(x)-x|}{x}dx + \int_{x_r}^{y} \frac{|\psi(x)-x|}{x}dx \\ &\leq \int_2^{x_r} \frac{|\psi(x)-x|}{x}dx + (1+r)y.\end{aligned}$$

Dividiendo por y se sigue que

$$\frac{|S(y)|}{y} \leq \frac{1}{y}\int_2^{x_r} \frac{|\psi(x)-x|}{x}dx + 1 + r.$$

Haciendo que $y \to \infty$ encontramos que $\limsup_{y\to\infty} \frac{|S(y)|}{y} \leq 1+r$ y dado que r se tomó de forma arbitraria tenemos que $\limsup_{y\to\infty} \frac{|S(y)|}{y} \leq 1$. Como $|W(x)| = e^{-x}|S(e^x)|$, entonces, tomando $y = e^x$ tenemos que $|W(x)| = \frac{|S(y)|}{y}$. Pero $y = e^x$ tiende a infinito si, y solo si, x tiende a infinito. En consecuencia, tener $\limsup_{y\to\infty} \frac{|S(y)|}{y} \leq 1$ es equivalente a tener que $\alpha = \limsup_{x\to\infty} |W(x)| \leq 1$.

Ahora veamos que $\alpha \leq \beta$. Para esto, tomemos $\varepsilon > 0$ fijo, pero arbitrario. Se sigue de la definición de β que, para u suficientemente grande,

$$\int_0^u |W(v)|dv \leq (\beta+\varepsilon)u.$$

Usando esto en (3.30) encontramos que, para x suficientemente grande

$$|W(x)| \leq \frac{2}{x^2}\int_0^x (\beta+\varepsilon)udu + \varepsilon = \frac{2(\beta+\varepsilon)}{x^2}\int_0^x udu + \varepsilon = (\beta+\varepsilon)+\varepsilon = \beta+2\varepsilon.$$

Concluimos que $\alpha \leq \beta + 2\varepsilon$. Como ε fue tomado de forma arbitraria se concluye que $\alpha \leq \beta$. $\square$

Como lo hemos mencionado, lo que tenemos que hacer para lograr nuestra prueba del teorema de los números primos, es probar que $\alpha = 0$; pero antes de eso, necesitamos tres resultados más acerca de la función $W(x)$.

Teorema 3.4.9. Existe una constante N tal que $|W(x_2) - W(x_1)| \leq N|x_2 - x_1|$, para todo $x_1, x_2 > 0$.

Demostración. Como $W(x) = e^{-x}S(e^x)$ y $S(y)$ es diferenciable para todo y que no sea potencia de primo, encontramos que

$$|W'(x)| \leq e^{-x}|S(e^x)| + |S'(e^x)|,$$

siempre que e^x no sea una potencia de primo. Por lo tanto, por el Corolario 3.4.3 y la desigualdad (3.16) podemos elegir una constante N tal que $|W'(x)| \leq N$, siempre que e^x no sea potencia de primo. De forma similar a como se demostró el Teorema 3.4.2, obtenemos que $|W(x_2) - W(x_1)| \leq N|x_2 - x_1|$ para todo $x_1, x_2 > 0$. □

Teorema 3.4.10. Existe una constante M tal que

$$\left|\int_{x_1}^{x_2} W(x)dx\right| \leq M,$$

para todo $x_1, x_2 > 0$. Además, si $W(x) \neq 0$ para todo x con $x_1 < x < x_2$, entonces $\displaystyle\int_{x_1}^{x_2} |W(x)|dx \leq M$.

Demostración. Usando el Teorema 2.4.2 y tomando $a(n) = \Lambda(n)$ y $f(t) = 1/t$ en la identidad de Abel, obtenemos que $\displaystyle\int_2^y \frac{\psi(t)}{t^2}dt = \ln y + O(1)$, de lo cual se sigue que

$$\int_2^y \frac{\psi(t) - t}{t^2}dt = \int_2^y \frac{\psi(t)}{t^2}dt - \int_2^y \frac{dt}{t} = \ln y + O(1) - \ln y + \ln 2 = O(1).$$

De (3.15) resulta que $(t - \psi(t))/t = O(1)$, y se sigue que

$$\frac{1}{y}\int_2^y \frac{t - \psi(t)}{t}dt = O\left(\frac{1}{y}\int_2^y dt\right) = O(1).$$

Usando todo esto obtenemos

$$\begin{aligned}\int_2^y \frac{S(z)}{z^2}dz &= \int_2^y \frac{1}{z^2}\int_2^z \frac{\psi(t) - t}{t}dtdz = \int_2^y \frac{\psi(t) - t}{t}\left(\int_t^y \frac{dz}{z^2}\right)dt \\ &= \int_2^y \frac{\psi(t) - t}{t}\left(\frac{1}{t} - \frac{1}{y}\right)dt \\ &= \int_2^y \frac{\psi(t) - t}{t^2}dt + \frac{1}{y}\int_2^y \frac{t - \psi(t)}{t}dt = O(1).\end{aligned}$$

Si tomamos $y = e^x$ y hacemos el cambio de variable $z = e^u$ se obtiene que

$$\int_2^y \frac{S(z)}{z^2}dz = \int_{\ln 2}^{\ln y} \frac{S(e^u)}{e^u}du = \int_{\ln 2}^{x} W(u)du,$$

en consecuencia $\displaystyle\int_{\ln 2}^{x} W(u)du = O(1)$. Ahora, para x_1 y x_2 cualesquiera con $x_1 < x_2$ tenemos

$$\int_{x_1}^{x_2} W(u)du = \int_{x_1}^{\ln 2} W(u)du + \int_{\ln 2}^{x_2} W(u)du = O(1).$$

Por lo tanto, existe una constante M_1 y un real a tales que $\displaystyle\left|\int_{x_1}^{x_2} W(u)du\right| \leq M_1$ para todo $x_1, x_2 \geq a$. Notando que para $0 < x_1, x_2 < a$, el Corolario 3.4.3 implica que $\displaystyle\left|\int_{x_1}^{x_2} W(u)du\right|$, como función de x_1 y x_2, es acotada. Existe una constante M_2 tal que $\displaystyle\left|\int_{x_1}^{x_2} W(u)du\right| \leq M_2$. Si tomamos $M = \text{máx}\{M_1, M_2\}$, entonces $\displaystyle\left|\int_{x_1}^{x_2} W(u)du\right| \leq M$ para todo $x_1, x_2 > 0$.

Finalmente mostremos que si $W(u) \neq 0$, para todo $x_1 < u < x_2$, la anterior desigualdad conduce a $\displaystyle\int_{x_1}^{x_2} |W(u)|du \leq M$. Para demostrar esto vamos a tener en cuenta que $W(x)$ es continua. Tener $W(u) \neq 0$ para todo $x_1 < u < x_2$ implica que $W(u) < 0$ para todo $x_1 < u < x_2$ o, por el contrario, $W(u) > 0$ para todo $x_1 < u < x_2$. En el primer caso tenemos que

$$\left|\int_{x_1}^{x_2} W(u)du\right| = -\int_{x_1}^{x_2} W(u)du = \int_{x_1}^{x_2} -W(u)du = \int_{x_1}^{x_2} |W(u)|du.$$

Similarmente, si $W(u) > 0$ para todo $x_1 < u < x_2$, $\displaystyle\left|\int_{x_1}^{x_2} W(u)du\right| = \int_{x_1}^{x_2} |W(u)|du$. En consecuencia $\displaystyle\int_{x_1}^{x_2} |W(u)|du \leq M$ siempre que $W(u) \neq 0$ en (x_1, x_2). $\square$

Para mostrar el siguiente resultado, vamos a tener en cuenta que si f es una función continua en un intervalo finito $[a, b]$ y $\displaystyle\int_a^b f(x)dx = 0$, entonces $f(x) = 0$ para algún $x \in (a, b)$. Además recuerde que $\psi(x)$ se mantiene constante en todo intervalo finito $[a, b]$ que no contenga potencias de primo.

Teorema 3.4.11. Los ceros de $W(x)$ no se acumulan en el intervalo $[1, +\infty)$.

Demostración. Sea $x_0 \geq 1$ fijo, pero arbitrario. Inicialmente probaremos que los ceros de $W(x)$ no se acumulan a la izquierda de x_0. Para esto vamos a suponer lo contrario

y tratar de llegar a una contradicción. En otras palabras, vamos a suponer que para todo $\varepsilon > 0$ existe un $\delta \in (x_0 - \varepsilon, x_0)$ tal que $W(\delta) = 0$.

Consideremos $k \neq e^{x_0}$, la potencia de primo más grande a la izquierda de e^{x_0} (*este k siempre existe ya que estamos tomando $x_0 \geq 1$ por lo que $e^{x_0} \geq e > 2$*). Así, el intervalo abierto (k, e^{x_0}) no contiene potencias de primo. Ahora, como la función $g : (0, \infty) \to (1, \infty)$ dada por $g(x) = e^x$ es una función sobreyectiva, estrictamente creciente y $k \in (1, \infty)$, podemos garantizar la existencia de un x_1 tal que $e^{x_1} = k$ y $0 < x_1 < x_0$.

Tomando $\varepsilon = x_0 - x_1$, existe un $\delta_1 \in (x_1, x_0)$ tal que $W(\delta_1) = 0$, si ahora tomamos $\varepsilon = x_0 - \delta_1$ existe un $\delta_2 \in (\delta_1, x_0)$ tal que $W(\delta_2) = 0$, y finalmente, tomando $\varepsilon = x_0 - \delta_2$ podemos garantizar la existencia de un $\delta_3 \in (\delta_2, x_0)$ tal que $W(\delta_3) = 0$. Todo esto muestra que existen ceros δ_1, δ_2 y δ_3 de $W(x)$ con $x_1 < \delta_1 < \delta_2 < \delta_3 < x_0$.

Además, $e^{x_1} < e^{\delta_1} < e^{\delta_2} < e^{\delta_3} < e^{x_0}$ implica que $[e^{\delta_1}, e^{\delta_3}]$ no contiene potencias de primo, pues el intervalo abierto (e^{x_1}, e^{x_0}) no contiene potencias de primo.

Por otro lado, como δ_1, δ_2 y δ_3 son ceros de $W(x)$, se sigue de la definición de $W(x)$ que

$$\int_2^{e^{\delta_1}} \frac{\psi(x) - x}{x} dx = \int_2^{e^{\delta_2}} \frac{\psi(x) - x}{x} dx = \int_2^{e^{\delta_3}} \frac{\psi(x) - x}{x} dx = 0,$$

luego

$$\int_{e^{\delta_1}}^{e^{\delta_2}} \frac{\psi(x) - x}{x} dx = \int_2^{e^{\delta_2}} \frac{\psi(x) - x}{x} dx - \int_2^{e^{\delta_1}} \frac{\psi(x) - x}{x} dx = 0,$$

de forma similar tenemos que $\displaystyle\int_{e^{\delta_2}}^{e^{\delta_3}} \frac{\psi(x) - x}{x} dx = 0$. Ahora, como el intervalo $[e^{\delta_1}, e^{\delta_3}]$ no contiene potencias de primo, la función $f(x) = (\psi(x) - x)/x$ es continua allí. Aplicando la propiedad mencionada al inicio a $f(x) = (\psi(x) - x)/x$ en los intervalos $[e^{\delta_1}, e^{\delta_2}]$ y $[e^{\delta_2}, e^{\delta_3}]$, encontramos que existen z_1, z_2 con $z_1 \in (e^{\delta_1}, e^{\delta_2})$, y $z_2 \in (e^{\delta_2}, e^{\delta_3})$ tales que $\psi(z_1) = z_1$ y $\psi(z_2) = z_2$, pero al no haber potencias de primos en el intervalo $[e^{\delta_1}, e^{\delta_3}] = [e^{\delta_1}, e^{\delta_2}] \cup [e^{\delta_2}, e^{\delta_3}]$ se obtiene que $\psi(z_1) = \psi(z_2)$. Esto nos lleva a que $z_1 = z_2$, lo cual es imposible, pues $z_1 \in (e^{\delta_1}, e^{\delta_2})$, y $z_2 \in (e^{\delta_2}, e^{\delta_3})$. De forma análoga podemos llegara a esta contradicción suponiendo que los ceros se acumulan a la derecha de x_0. Por lo tanto, para $x \geq 1$, los ceros de $W(x)$ no se acumulan. $\square$

3.4.4. Demostración final

Probamos ahora, como lo hemos prometido, que $\alpha = 0$. Antes de ir a la demostración recordemos que, del Teorema 3.4.10 existe una constante M tal que

$$\left|\int_{x_1}^{x_2} W(u)du\right| \leq M, \tag{3.31}$$

para todo $x_1, x_2 \geq 0$. Además, el Teorema 3.4.9 también nos garantiza la existencia de una constante N tal que

$$|W(x_2) - W(x_1)| \leq N|x_2 - x_1|, \tag{3.32}$$

para todo $x_1, x_2 > 0$. Por conveniencia vamos a considerar que $MN > 1$. Usando que $\alpha \leq 1$, se sigue que $\alpha^2/MN < 1$ de donde obtenemos la siguiente desigualdad

$$\frac{1}{2} < 1 - \frac{\alpha^2}{2MN}. \tag{3.33}$$

Ahora procedemos a mostrar que $\alpha = 0$. Tomemos $r > \alpha$. De la definición de α existe un x_r tal que

$$|W(x)| \leq r, \tag{3.34}$$

siempre que $x \geq x_r$. Si tuviéramos $W(x) \neq 0$ para todo $x \geq x_r$, entonces, como vimos en el Teorema 3.4.10, la desigualdad (3.31) nos garantiza que $\displaystyle\int_{x_r}^{x} |W(u)|du \leq M$, para todo $x \geq x_r$. Luego

$$\begin{aligned}
\beta = \lim_{x\to\infty} \frac{1}{x}\int_0^x |W(u)|du &= \lim_{x\to\infty} \frac{1}{x}\int_0^{x_r} |W(u)|du + \lim_{x\to\infty} \frac{1}{x}\int_{x_r}^{x} |W(u)|du \\
&\leq \left(\int_0^{x_r} |W(u)|du\right) \lim_{x\to\infty} \frac{1}{x} + M \lim_{x\to\infty} \frac{1}{x} \\
&= 0,
\end{aligned}$$

lo que mostraría que $\beta = 0$ y en consecuencia $\alpha = 0$, por el Teorema 3.4.8.

Supongamos entonces que $W(x)$ tiene ceros arbitrariamente grandes. Sean $a, b \geq x_r$ ceros de $W(x)$, con $a < b$. Mostraremos que

$$\int_a^b |W(x)|dx < (b-a)r\left(1 - \frac{\alpha^2}{2MN}\right). \tag{3.35}$$

Inicialmente, analizaremos el caso en el que a y b son ceros consecutivos de $W(x)$. Para esto vamos a estudiar los casos; $b-a \geq 2M/r$, $b-a \leq 2r/N$ y $2r/N < b-a < 2M/r$.

Comenzamos analizando el caso $b-a \geq 2M/r$. Como estamos considerando a y b ceros sucesivos de $W(x)$, entonces $W(x) \neq 0$ para todo x con $a < x < b$. Luego, se sigue de (3.31) que $\int_a^b |W(x)|dx \leq M \leq \frac{1}{2}(b-a)r$. Además, si multiplicamos la desigualdad (3.33) por $(b-a)r$ se obtiene que $\frac{1}{2}(b-a)r < (b-a)r\left(1-\alpha^2/(2MN)\right)$. En consecuencia

$$\int_a^b |W(x)|dx \leq \frac{1}{2}(b-a)r < (b-a)r\left(1-\frac{\alpha^2}{2MN}\right).$$

Ahora, consideremos el caso $b-a \leq 2r/N$. Para todo x con $a \leq x \leq b$, se obtiene de (3.32) que $|W(x)-W(a)| \leq N(x-a)$, y como $W(a)=0$, en realidad tenemos $|W(x)| \leq N(x-a)$. Análogamente, $|W(b)-W(x)| \leq N(b-x)$ y $W(b)=0$ implica que $|W(x)| \leq -N(x-b)$. Esto muestra que para todo x en el intervalo $[a,b]$, la curva $|W(x)|$ se encuentra encerrada por las rectas $y=N(x-a)$, $y=-N(x-b)$ y $y=0$, las cuales forman una región triangular de altura $\frac{1}{2}(b-a)N$ y base $b-a$.

En la Figura 3.3 representamos gráficamente esta situación.

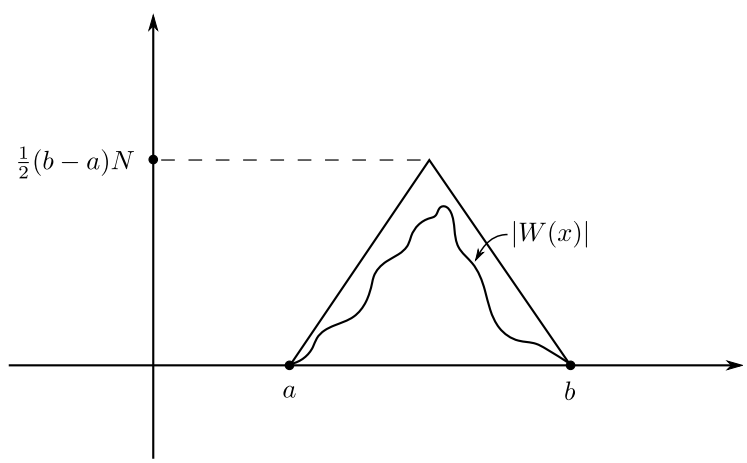

Figura 3.3: Región triangular que encierra a $|W(x)|$

Podemos ver que el área bajo la curva $|W(x)|$ para x entre a y b es menor o igual que el área del triángulo de altura $\frac{1}{2}(b-a)N$ y base $b-a$. De esto, teniendo en cuenta que $\frac{1}{2}(b-a)N \leq r$, obtenemos $\int_a^b |W(x)|dx \leq \frac{1}{2}(b-a)r$. Usando este hecho,

multiplicando nuevamente la desigualdad (3.33) por $(b-a)r$, encontramos que

$$\int_a^b |W(x)|dx \leq \frac{1}{2}(b-a)r < (b-a)r\left(1-\frac{\alpha^2}{2MN}\right).$$

Finalmente, consideramos el caso en que $2r/N < b-a < 2M/r$. Razonando como en el caso anterior, para una distancia r/N desde cada extremo, como se muestra la Figura 3.4,

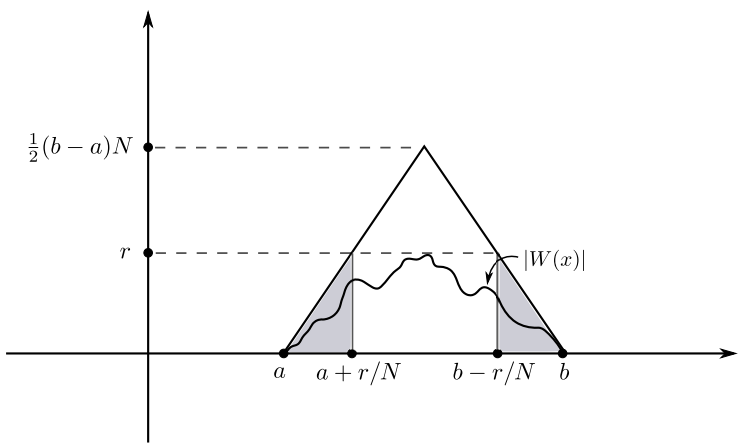

Figura 3.4: Región triangular que encierra a $|W(x)|$

donde se ha determinado la posición de r evaluando la recta $y = N(x-a)$ para $x = a + r/N$. De la figura vemos claramente que el área de las regiones sombreadas es $\frac{r^2}{2N}$, por lo que

$$\int_a^{a+r/N} |W(x)|dx + \int_{b-r/N}^b |W(x)|dx \leq \frac{r^2}{N}.$$

Por otro lado, se sigue de (3.34) que

$$\int_{a+r/N}^{b-r/N} |W(x)|dx \leq r\int_{a+r/N}^{b-r/N} dx = \left(b-a-\frac{2r}{N}\right)r.$$

De estas dos desigualdades, y la hipótesis de que $2r/N < b-a < 2M/r$, se sigue que

$$\begin{aligned}\int_a^b |W(x)|dx &= \int_a^{a+r/N} |W(x)|dx + \int_{b-r/N}^b |W(x)|dx + \int_{a+r/N}^{b-r/N} |W(x)|dx \\ &\leq \frac{r^2}{N} + \left(b-a-\frac{2r}{N}\right)r = (b-a)r\left(1-\frac{r}{N(b-a)}\right) \\ &\leq (b-a)r\left(1-\frac{r^2}{2MN}\right) < (b-a)r\left(1-\frac{\alpha^2}{2MN}\right).\end{aligned}$$

De esta manera hemos mostrado que cuando $a, b \geq x_r$ son ceros consecutivos de $W(x)$, se tiene

$$\int_a^b |W(x)|dx < (b-a)r\left(1 - \frac{\alpha^2}{2MN}\right). \tag{3.36}$$

En el caso en que a y b no sean ceros consecutivos de $W(x)$, el Teorema 3.4.11 nos permite construir una partición $z_0, z_1, \ldots, z_n$ del intervalo cerrado $[a, b]$ de tal forma que los z_i con $0 < i < n$ sean precisamente los ceros intermedios entre a y b. De esta forma, para todo $i = 1, 2, \ldots, n$, tenemos que $W(x) \neq 0$ en los intervalos abiertos (z_{i-1}, z_i). Así, el resultado (3.36) nos garantiza que

$$\int_{z_{i-1}}^{z_i} |W(x)|dx \leq (z_i - z_{i-1})r(1 - \alpha^2/(2MN)),$$

para todo $i = 1, 2, \ldots, n$. Luego

$$\begin{aligned}\int_a^b |W(x)|dx &= \sum_{i=1}^n \int_{z_{i-1}}^{z_i} |W(x)|dx \leq r\left(1 - \frac{\alpha^2}{2MN}\right)\sum_{i=1}^n (z_i - z_{i-1}) \\ &= r\left(1 - \frac{\alpha^2}{2MN}\right)(z_n - z_0) = (b-a)r\left(1 - \frac{\alpha^2}{2MN}\right).\end{aligned}$$

Esto termina la prueba de nuestra afirmación.

Ahora, sea $y \geq x_r$. Usando nuevamente el Teorema 3.4.11 podemos escoger x_0 como el primer cero de $W(x)$ a la derecha de x_r y y_0 el cero más grande a la izquierda de y, entonces, por (3.35) y (3.31) se obtiene que

$$\begin{aligned}\int_0^y |W(x)|dx &= \int_0^{x_0} |W(x)|dx + \int_{x_0}^{y_0} |W(x)|dx + \int_{y_0}^{y} |W(x)|dx \\ &< \int_0^{x_0} |W(x)|dx + (y_0 - x_0)r\left(1 - \frac{\alpha^2}{2MN}\right) + M \\ &\leq \int_0^{x_0} |W(x)|dx + yr\left(1 - \frac{\alpha^2}{2MN}\right) + M,\end{aligned}$$

donde se ha usado que $y_0 - x_0 \leq y$, pues $0 < x_0 \leq y_0 \leq y$. Dividiendo por y se obtiene finalmente que

$$\frac{1}{y}\int_0^y |W(x)|dx < \frac{1}{y}\int_0^{x_0} |W(x)|dx + r\left(1 - \frac{\alpha^2}{2MN}\right) + \frac{M}{y},$$

haciendo que $y \to \infty$, llegamos a que $\beta \leq r(1 - \alpha^2/(2MN))$. Recordando que $\beta \geq \alpha$, tenemos

$$\alpha \leq r\left(1 - \frac{\alpha^2}{2MN}\right),$$

dado que esto es válido para todo $r > \alpha$, debe ser válido para $r = \alpha$, es decir, tenemos $\alpha \leq \alpha(1 - \alpha^2/(2NM))$, de donde resulta que $\alpha^3 \leq 0$, y como $\alpha \geq 0$, necesariamente debemos tener $\alpha = 0$.

Bibliografía

[1] Apostol, T., *Introduction to analytic number theory*, Springer-Verlag, New York, 1976.

[2] Hadamard, J., *Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques*, Bull. Soc. Math. France, 1896.

[3] Levinson, N., *A Motivated Account of an Elementary Proof of the Prime Number Theorem*, The American Mathematical Monthly, Vol. 76, 1969.

[4] Selberg, A., *An Elementary Proof of the Prime-Number Theorem*, Annals of Mathematics, Second Series, Vol. 50, 1949

[5] Tatuzawa, T. and Iseki, K., *On Selberg's Elementary Proof of the Prime-Number Theorem*. Comm. by Z. Suetuna, M.Z.A., 1951.

[6] De la Vallée Poussin, Ch. *Recherches analytiques sur la théorie des nombres premiers*. Ann. Soc. Sci. Bruxelles, 1896.