Sobre el teorema de los números primos en progresiones aritmética

Mateo Andrés Manosalva Amaris

Director:
John Jaime Rodriguez

Facultad de Ciencias
Departamento de Matemáticas
Bogotá, D.C, Colombia
10 de marzo de 2025

Abstract

The prime number theorem is the assetion that in the limit, the quotient $\frac{\pi(x)\log x}{x}$ goes to 1, which means that $\pi(x) \sim \frac{x}{\log x}$, where $\pi(x)$ is the prime counting function. In arithmetic progressions $a + kq$ with $(a, q) = 1$, we have that $\pi(a, q, x)$; the prime counting function restricted to the progression, has the asymptotic behavior $\pi(a, q, x) \sim \frac{x}{\varphi(q)\log x}$, meaning that primes are uniformly distributed among the residue classes modulo q. In this work, we will present the proof of this result, the underlying ideas, and applications. For this, we will make use of Tauberian theory, which will allow us to present a detailed and concise proof, followed by studying the non-vanishing of $L(\chi, s)$ and some properties of Dirichlet characters and series.

Resumen

El teorema de los números primos nos dice que en el límite, el cociente $\frac{\pi(x)\log x}{x}$ tiende a 1, es decir, que $\pi(x) \sim \frac{x}{\log x}$ donde $\pi(x)$ es la función contadora de primos. En progresiones aritméticas $a+kq$ con $(a, q) = 1$, tenemos que $\pi(a, q, x)$; la función contadora restringida a la progresión, tiene el comportamiento asintótico $\pi(a, q, x) \sim \frac{x}{\varphi(q)\log x}$, es decir, los primos se distribuyen uniformemente en las clases de residuos módulo q. En este trabajo se presentará la prueba de este resultado, las ideas subyacentes y aplicaciones. Para esto, haremos uso de la teoría Tauberiana, lo que nos permitirá presentar una prueba detallada y corta, que se seguirá estudiando la no nulidad de $L(\chi, s)$ y algunas propiedades de los caracteres y series de Dirichlet.

Agradecimientos

Por muchos años solía creer imposible que algún día me graduaría de matemático. Al momento de entregar esta tesis vienen a mí muchos recuerdos del largo y duro proceso que fue para un muchacho de una región, aparentemente olvidada para el país, llegar a Bogotá y adaptarse a las exigencias académicas de esta universidad. Sin lugar a dudas hay algo de cierto en lo que pensaba, no podría haber llegado aquí sin el apoyo de muchas personas en mi vida, esta tesis no hubiera sido posible si estas personas no me hubieran apoyado de la forma en que lo hicieron.

Así pues, quisiera comenzar agradeciendo a mi familia, primero a mis padres (Marco y Raiza) por su apoyo incondicional y por creer en mí durante estos largos años, papá, sin ti nada de esto hubiera sido posible.

También quisiera agradecer a mi hermano James por haber sido otro padre para mí y a mi hermano Juan, quien siempre fue mi ejemplo, mi compañía y mi mejor amigo, fue gracias a él que conocí la educación superior. Quisiera agradecer a mis hermanos (Luis y Fabio) por su amistad y su apoyo a lo largo de los años, a Jessica y Kamila por sacarme de la casa cuando probablemente la matemática ya me tenía el cerebro estallado, también quisiera agradecerle a mi tía Marleny por haber sido un gran apoyo para mi padre y para mí. Gracias a todos, nunca han dejado de estar a mi lado, sus consejos y su afecto los atesoro en mi corazón.
Quisiera agradecer a Karolina por su apoyo, amor y comprensión a lo largo de los años, por haber creído en mí cuando yo no lo hice. Sin lugar a dudas, sin ella este trabajo tampoco hubiera sido posible.

También quiero agradecer a mis amigos, a los viejos (Andrés, Iván y Viuche) con quienes he compartido casi toda la carrera y a los nuevos (Alejandra, Santiago, Sergio y Sandra). Su amistad y su apoyo a lo largo de los años ha sido fundamental, gracias por todas las charlas sobre matemáticas y por divagar conmigo sobre cualquier cosa, por las risas, por todas las veces que salimos luego de clases y por la cantidad no contable de noches que trasnochamos estudiando, gracias a ustedes soy mucho mejor matemático de lo que pensé que podría ser y la carrera no ha sido todo un infierno. Me honra haber conocido tan buenos matemáticos y haber hecho tan buenos amigos a lo largo de los años.
Quisiera agradecerles particularmente a Santiago y Alejandra por estos últimos años en los que fueron indispensables, por haber sido mis compañeros de trabajo en la parte más difícil de la carrera y por la amistad que hemos construido.

Adicionalmente quisiera expresar mi gratitud a Sergio y Santiago por su constante ayuda a revisar este largo trabajo, particularmente a Santiago, aunque no se exprese en forma puntual en las páginas de este texto, muchas de sus ideas fueron de gran ayuda cuando me encontraba estancado en una demostración.

Finalmente quiero agradecer a los profesores del departamento de matemáticas, particularmente al profesor John Jaime Rodriguez, por haber sido un director paciente, comprensivo y por todo lo que aprendí de él a largo del tiempo, sin lugar a duda llevó a que esta tesis sea lo que es.

Gracias a todos.

Contenido

4 | Primos en progresiones aritmética

Introducción

> *La matemática posee no solo verdad, sino también belleza suprema; una belleza fría y austera, como aquella de la escultura, sin apelación a ninguna parte de nuestra naturaleza débil, sin los adornos magníficos de la pintura o la música, pero sublime y pura, y capaz de una perfección severa como solo las mejores artes pueden presentar*
>
> **— Bertrand Russel**

Alrededor del año 300 a.C. Euclides prueba que hay infinitos números primos, establece que si los primos son finitos, entonces el producto $p_1 \dots p_n + 1$ no es divisible por ningún primo $p_1, \dots, p_n$, de esta manera siempre se puede construir un número primo adicional. En el siglo XVIII Euler prueba que hay infinitos primos usando la divergencia de la serie armónica, si asumimos que hay un número finito de números primos, entonces el siguiente producto es finito:

$$\prod_p \frac{1}{1-p^{-1}}.$$

Ahora note que el término del producto es a lo que converge una serie geométrica y dado que $|p^{-1}| < 1$, entonces

$$\begin{aligned}\infty > \prod_p \frac{1}{1-p^{-1}} &= \prod_p \left(\sum_{k=0}^{\infty} \frac{1}{p^k}\right) = \prod_p \left(1 + \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \dots\right) \\ &= \sum_{n=1}^{\infty} \frac{1}{n} \\ &= \infty,\end{aligned}$$

ya que todo número natural puede escribirse de manera única como producto de potencias de primos, esto nos lleva a una evidente contradicción. Euler consigue este argumento ya que venía de estudiar problemas similares, como la convergencia de la serie

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

La idea que tuvo Euler para este problema provenía de estudiar la serie de Taylor de la función $\sin x$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm \dots,$$

así

$$\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} \pm \dots$$

En este punto es donde hace un salto de fe, pensando que el polinomio de Taylor se puede escribir como un producto infinito si lo factorizamos sobre sus raíces, ie. Las raíces de $\frac{\sin x}{x}$, asume que lo que ocurre para polinomios finitos también se tiene para polinomios infinitos, obteniendo que

$$\begin{aligned}\frac{\sin x}{x} &= 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} \pm \dots \\ &= \left(1 + \frac{x}{\pi}\right)\left(1 - \frac{x}{\pi}\right)\left(1 + \frac{x}{2\pi}\right)\left(1 - \frac{x}{2\pi}\right)\left(1 + \frac{x}{3\pi}\right)\left(1 - \frac{x}{3\pi}\right)\dots \\ &= \left(1 - \frac{x^2}{\pi^2}\right)\left(1 - \frac{x^2}{2^2\pi^2}\right)\left(1 - \frac{x^2}{3^2\pi^2}\right)\dots,\end{aligned}$$

luego comparando el coeficiente de x^2 en la serie con el de el producto:

$$\frac{1}{3!} = \frac{1}{\pi^2}\left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots\right).$$

Esta idea que le daría la "solución" al problema se formaliza a través del teorema de factorización de Weierstrass. Euler seguiría estudiando este problema por mucho tiempo y lo generalizaría a través de la serie absolutamente convergente

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad s > 1.$$

Tiempo después encuentra una fórmula para obtener los valores de esta función en los números pares, ie. $\zeta(2s)$ y también obtuvo su desarrollo como producto:

$$\zeta(s) = \prod_{p} \frac{1}{1 - p^{-s}}.$$

Esto le permitió demostrar la divergencia de la serie $\sum_{p} \frac{1}{p}$, un argumento directo y totalmente analítico de que hay infinitos números primos.

Estas ideas llamaron la atención de dos matemáticos muy importantes, Dirichlet y Riemann. Dirichlet usó estas ideas para probar su teorema de progresiones aritmética, Riemann por otro lado estudió íntimamente la función $\zeta(s)$, le asignó a s un número complejo y también la llevó a tener su fama actual al lanzar su conocida conjetura, pero, ¿esto qué tiene que ver con el teorema de los números primos?.

Conjeturado de manera independiente por Gauss (1792) y Legendre (1798), el teorema de los números primos nos permite entender el comportamiento asintótico de la función contadora de primos $\pi(x)$, nos dice que para números grandes, la cantidad de primos menores que x se puede aproximar por $\frac{x}{\log x}$, escrito de manera formal

$$\lim_{x\to\infty} \frac{\pi(x)\log x}{x} = 1 \quad \text{o en notación asintótica} \quad \pi(x) \sim \frac{x}{\log x}.$$

Una interpretación heurística de este teorema viene de estudiar la densidad de un conjunto de números naturales. Dado $N \subseteq \mathbb{N}$, la densidad natural de N la definimos como:

$$d = \lim_{n\to\infty} \frac{|\{m \leqslant n : m \in N\}|}{n}, \quad \text{siempre que exista el límite.}$$

Note que estudiar la probabilidad de, por ejemplo, que un entero sea divisible por un primo p será equivalente a calcular la densidad del conjunto de enteros que cumple esta propiedad, veamos esto. Dado n, sea c el número de enteros $m \leqslant n$ tal que $p \mid m$, sabemos por un simple conteo que:

$$\frac{n}{p} - 1 \leqslant c \leqslant \frac{n}{p} + 1.$$

Luego, $d = \lim_{n\to\infty} \frac{c}{n} = \frac{1}{p}$ por el criterio de comparación. Esto nos dice que la probabilidad de que un entero no sea divisible por p es $1 - \frac{1}{p}$, sabemos además que este evento es excluyente, así... La probabilidad de que un número sea primo viene dada por:

$$\prod_{p<n} \left(1 - \frac{1}{p}\right).$$

Para estimar el crecimiento asintótico de este producto una buena idea sería invertirlo, En efecto:

$$\prod_{p<n} \frac{1}{1-p^{-1}} = \prod_{p<n} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \ldots\right) = \sum_{k<n} \frac{1}{k} = H(n),$$

en el siguiente capítulo veremos justamente que $H(n) \sim \log n$, esto nos dice que la probabilidad de que un número sea primo es $(\log n)^{-1}$, más aún:

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

Pero esto no es una prueba del TNP, entonces ¿cómo se puede demostrar algo así?, el camino a seguir en un principio es sorprendente y viene del estudio de la función de $\zeta(s)$, vista como función de variable compleja absolutamente convergente si $\Re(s) > 1$. El primero en mostrar que estudiar esta función daba un camino hacia una prueba del teorema de los números primos fue Riemann en su famoso articulo "Sobre la cantidad de primos menores que una magnitud dada"[1]. Allí Riemann presentaría muchas ideas, pero no las desarrollaría y fue el trabajo de los matemáticos en los siguientes 50 años llegar a una demostración, trabajo que culminaría en las demostraciones Hadammar y de la Vallée Poussin que aparecen en 1896, la prueba, vendría del hecho de que $\zeta(1+it) \neq 0$, es decir, la función ζ no se anula en la recta vertical de los complejos con parte real 1, algo sencillamente maravilloso.

Veremos al final de este trabajo la forma en que este teorema se extiende a progresiones aritmética $a + kq$ con $(a, q) = 1$, donde

$$\pi(a, q, x) \sim \frac{x}{\varphi(q) \log x},$$

con φ denotando la función Phi de Euler, y como hay $\varphi(q)$ clases generadoras de primos, entonces los primos se distribuyen uniformemente en las clases módulo q. Esta es una versión más fuerte que el TNP original, y será nuestro objetivo conseguirla.

Sería ideal que el lector de este trabajo esté familiarizado con algunos conceptos del Análisis y del Álgebra, particularmente hago énfasis en un curso de Análisis II y de Variable Compleja, esto hará que la lectura sea mucho más agradable.

Lista de símbolos, notación:

Aquí se esclarecen algunas herramientas de notación importantes para este trabajo.

- $\sum_{p}$ Denota que la suma se hace sobre el conjunto de los números primos, análogamente $\prod_{p}$.
- Denota $\log x$, la función logaritmo natural.
- $f(x) = O(g(x))$
  Existe una constante $M > 0$ tal que $|f(x)| \leqslant Mg(x)$ para todo x en un dominio específico.
- $f(x) \ll g(x)$: $\quad f(x) = O(g(x))$
- $f(x) = o(g(x))$: $\quad \lim_{x\to\infty} \dfrac{f(x)}{g(x)} = 0$
- $f(x) \asymp g(x)$: $\quad f(x) \ll g(x)$ y $g(x) \ll f(x)$
- $f(x) \sim g(x)$ denota $\lim_{x\to\infty} \dfrac{f(x)}{g(x)} = 1$
- Denotamos por (a, b) el máximo común divisor de a y b.
- $\mathbb{1}_{\mathcal{K}}$ denota la función indicadora (característica) de $\mathcal{K}$.
- Tomaremos $\mathbb{N} = \mathbb{Z}^{+}$: $\quad$ el conjunto de los enteros positivos.
- $a^k \parallel b$: $\quad a^k \mid b$ y $a^{k+1} \nmid b$
- e^x o $\exp(x)$ denota la función exponencial.
- $\{x\} = x - \lfloor x \rfloor$: la función parte fraccionaria.
- $\chi(n)$: un carácter de Dirichlet.
- $\pi(x)$: la función contadora de primos.
- $\pi(a, n, x)$: la función contadora de primos en una progresión aritmética,

$$\pi(a, n, x) = \sum_{p\equiv a \bmod n} 1$$

- $\operatorname{li}(x) = \int_2^x \dfrac{dt}{\log t}$: la función logaritmo integral.

- $\sum_{p \equiv a(n)}$ en este trabajo denota $\sum_{p \equiv a \text{ mód } n}$.
- Sea $f : (a, b) \longrightarrow \mathbb{R}$ una función real, denotamos:

$$f(c-) = \lim_{x \to c^-} f(x),\ c \in (a, b]$$
$$f(c+) = \lim_{x \to c^+} f(x),\ c \in [a, b).$$

Nota. Algunas de estas herramientas de notación se abordarán de manera más detallada más adelante.

Preliminares

> *Hasta el día de hoy, los matemáticos han intentado en vano descubrir algún orden en la secuencia de números primos, y tenemos razones para creer que es un misterio al que la mente humana nunca penetrará*
>
> **— Leonhard Euler**

Para comenzar con este capítulo presentaremos el teorema fundamental de la aritmética (TFA), una pieza crucial en cualquier trabajo sobre teoría de números.

Teorema 1.1 (TFA)**.** Todo entero $n > 1$ se puede escribir como producto de primos de manera única salvo el orden de los factores, es decir:

$$n = \prod_{j=1}^{m} p_j^{k_j}.$$

Escribiremos $p^m \parallel n$ siempre que si $p^m \mid n$ entonces $p^{m+1} \nmid n$, es decir, p^m es la potencia exacta que divide a n, esto nos permite escribir el TFA como:

$$n = \prod_{p^m \parallel n} p^m.$$

1.1 Funciones aritmética

Definición. Una función aritmética es una función con dominio los naturales y codominio $\mathbb{R}$ o $\mathbb{C}$, es decir a es función aritmética si

$$a : \mathbb{N} \to \mathbb{F},$$

con $\mathbb{F} = \mathbb{C}$ o $\mathbb{F} = \mathbb{R}$.

Esta definición nos muestra que las funciones aritmética no son más que sucesiones de números reales o complejos, en algunos casos será útil considerarlas de esta manera y de manera análoga a las sucesiones las denotaremos como a_n, donde cada a_n representa $a(n)$. Veamos algunos ejemplos importantes:

- **Función constante k:**

$$k(n) = k, \text{ para todo } n \in \mathbb{N}.$$

- **Función unidad:**

$$e(n) = \begin{cases} 1 & n = 1 \\ 0 & n \neq 1. \end{cases}$$

- **Función número de divisores:** $\tau(n)$, el número de divisores positivos de n (incluyendo 1 y n)

$$\tau(n) = \sum_{j|n} 1.$$

- **Función suma de divisores:** $\sigma(n)$, la suma de los divisores positivos de n

$$\sigma(n) = \sum_{j|n} j.$$

- **Función de Möbius:** $\mu(n)$, se define como

$$\mu(n) = \begin{cases} 1 & n = 1 \\ 0 & \text{si n no es libre de cuadrados} \\ (-1)^k & \text{si n tiene k factores primos.} \end{cases}$$

- **Función phi de Euler:** $\varphi(n)$, el número de enteros positivos $m \leqslant n$ que son primos relativos a n ($(m, n) = 1$)

$$\varphi(n) = \sum_{\substack{m=1 \\ (m,n)=1}}^{n} 1.$$

- **Función de Von Mangoldt:** $\Lambda(n)$, se define como

$$\Lambda(n) = \begin{cases} \log p & n = p^m \\ 0 & \text{en otro caso.} \end{cases}$$

- **Función identidad:** $N(n)$, la función identidad se define como:

$$N(n) = n.$$

Por la naturaleza de $\mathbb{N}$, existen dos clases importantes de funciones aritmética, las funciones aditivas y multiplicativas:

- Las funciones aditivas que satisfacen

$$f(mn) = f(m) + f(n) \quad \text{siempre que } (m, n) = 1,$$

- las funciones multiplicativas que satisfacen

$$f(mn) = f(m)f(n) \text{ siempre que } (m,n) = 1.$$

Si una función aditiva o multiplicativa satisface la propiedad para cualquier par de números naturales m y n, se dirá que la función es completamente aditiva o completamente multiplicativa, respectivamente, las funciones aditivas y multiplicativas están determinadas por sus valores en las potencias de los números primos.

Demostración. Supongamos que f es aditiva y $n > 1$, por el TFA:

$$f(n) = f\left(\prod_{p^m \| n} p^m\right) = \sum_{p^m \| n} f(p^m).$$

Ahora, si f es multiplicativa:

$$f(n) = f\left(\prod_{p^m \| n} p^m\right) = \prod_{p^m \| n} f(p^m).$$

□

Si además la función es completamente multiplicativa:

$$f(n) = f\left(\prod_{i=1}^{m} p_i^{k_i}\right) = \prod_{i=1}^{m} f(p_i)^{k_i},$$

lo que también ocurre para funciones completamente aditivas, cambiando el producto por una suma. Una propiedad adicional que será útil para caracterizar estas funciones es que si f es aditiva y no idénticamente nula, entonces para algún n, $f(1 \cdot n) = f(1) + f(n)$, así $f(1) = 0$, análogamente si f es multiplicativa y no idénticamente nula $f(1 \cdot n) = f(1)f(n)$, $f(1) = 1$.

Ahora veamos que aunque la función de Von Mangoldt, parece extraña, su definición es natural y nos permite obtener una versión logarítmica del teorema fundamental de la aritmética.

Teorema 1.2. Dado $n \in \mathbb{N}$, $n > 1$ entonces:

$$\log(n) = \sum_{j|n} \Lambda(j).$$

Demostración. Note que si $n > 1$, entonces por el TFA:

$$\log(n) = k_1 \log(p_1) + \ldots + k_m \log(p_m),$$

donde los p_j de la igualdad son los primos de su descomposición y k_j sus potencias respectivas. Así, esta igualdad nos dice que en el cálculo de $\log(n)$ solo importan los valores del log en los divisores primos o potencias de primos, luego:

$$\log(n) = \sum_{j|n} \Lambda(j).$$

□

Sin embargo, la principal motivación para introducir la función de Von Mangoldt es que sus sumas parciales $\sum_{n\leqslant x} \Lambda(n)$ son la suma ponderada de las potencias de primos $p^m \leqslant x$, tomando como peso $\log p$, el peso correcto para compensar la densidad de primos. No es difícil demostrar que las potencias p^m con $(m \geqslant 2)$ contribuyen poco en la suma anterior.

De hecho, estudiar el comportamiento asintótico de la suma anterior resultará equivalente a estudiar el de la función de contadora de primos $\pi(x)$; aún más, el TNP es equivalente a la afirmación

$$\lim_{x\to\infty} \frac{1}{x} \sum_{n\leqslant x} \Lambda(n) = 1.$$

Esta equivalencia nos dará el camino a la prueba del teorema de los números primos, lo que la convierte en una función aritmética muy importante.[2]

Definición. Las funciones $\psi(x)$ y $\vartheta(x)$ de Chevyshev se definen como sigue:

$$\vartheta(x) = \sum_{p\leqslant x} \log p, \qquad \psi(x) = \sum_{n\leqslant x} \Lambda(n).$$

1.1.1. La función de Möbius

Es natural preguntarse por la definición de la función de Möbius, ya que de todas parece ser la más extraña, uno se preguntaría si hay una forma de motivarla...

Consideremos la función:

$$L(x) = \sum_{n\leqslant x} \log(n).$$

Note que aplicando el teorema anterior

$$L(x) = \sum_{n\leqslant x} \log(n) = \sum_{n\leqslant x} \sum_{j|n} \Lambda(j). \tag{1.1}$$

Vamos a aplicar una técnica muy útil y frecuente en teoría de números, el cambio de orden de sumación, para esto vamos a cambiar n y j de orden en la doble suma (1.1) y conservaremos la condición $j \mid n$.

$$\begin{aligned} L(x) &= \sum_{n\leqslant x} \log(n) = \sum_{n\leqslant x}\sum_{j|n} \Lambda(j) \\ &= \sum_{j\leqslant x}\sum_{\substack{n\leqslant x \\ j|n}} \Lambda(j) \\ &= \sum_{j\leqslant x} \Lambda(j) \sum_{\substack{n\leqslant x \\ j|n}} 1, \end{aligned}$$

ahora, ¿cuántos enteros positivos $n \leqslant x$ hay tal que $j \mid n$?, pues exactamente $\frac{x}{j}$, así

$$L(x) = \sum_{j\leqslant x} \Lambda(j) \sum_{m\leqslant \frac{x}{j}} 1,$$

y cambiando nuevamente el orden de sumación

$$\begin{aligned} L(x) &= \sum_{m\leqslant x}\sum_{j\leqslant \frac{x}{m}} \Lambda(j) \\ &= \sum_{m\leqslant x} \psi\left(\frac{x}{m}\right). \end{aligned}$$

Esta identidad la abordaremos más adelante, pero de momento sabemos que podemos escribir a $L(x)$ en términos de $\psi(x)$, ¿y si queremos lo opuesto?, ie. a $\psi(x)$ en términos de $L(x)$, ¿podemos **invertir** el papel de las funciones?. Vamos a abordar esta pregunta poniéndola en un contexto más general.

Siguiendo a [3], supongamos $F(x)$ y $G(x)$ funciones aritmética con $G(x) = \sum_{n\leqslant x} F\left(\frac{x}{n}\right)$, tenemos que:

$$G\left(\frac{x}{2}\right) = F\left(\frac{x}{2}\right) + F\left(\frac{x}{4}\right) + F\left(\frac{x}{6}\right) + \ldots,$$

así:

$$G(x) - G\left(\frac{x}{2}\right) = F(x) + F\left(\frac{x}{3}\right) + F\left(\frac{x}{5}\right) + \ldots,$$

podemos pensar que continuar restando los términos $G\left(\frac{x}{j}\right)$ nos permitirá obtener la inversión, sin embargo el término $G\left(\frac{x}{3}\right)$ contiene a $F\left(\frac{x}{6}\right)$, por tanto

$$G(x) - G\left(\frac{x}{2}\right) - G\left(\frac{x}{3}\right) = F(x) + F\left(\frac{x}{5}\right) - F\left(\frac{x}{6}\right) + F\left(\frac{x}{7}\right) + \ldots,$$

así, en los siguientes pasos debemos eliminar $-F\left(\frac{x}{6}\right)$. Esto se lograría sumando $G\left(\frac{x}{6}\right)$ y no restándolo. La suma anterior nos muestra además que no necesitamos restar

$G\left(\frac{x}{4}\right)$ pues $F\left(\frac{x}{4}\right)$ ya desapareció al restar $G\left(\frac{x}{2}\right)$.

Así, podemos intuir que necesitamos multiplicar $G(\frac{x}{j})$ en cada sumando, por una función que nos de el signo adecuado (sume y reste, según se necesite) o anule el término, como ocurre en el caso de $G\left(\frac{x}{4}\right)$. Denotemos esta función que estamos buscando como $\mu(x)$. Si suponemos que existe dicha función, entonces

$$F(x) = \sum_{j \leqslant x} \mu(j) G\left(\frac{x}{j}\right). \tag{1.2}$$

Además, ya tenemos algunos valores de μ, $\mu(1) = 1$, $\mu(2) = \mu(3) = -1, \mu(4) = 0$ y $\mu(6) = 1$. Podemos de momento darnos cuenta que estos valores parecen coincidir con los que obtendríamos al evaluar la función de Möbius, lo cual no es ninguna coincidencia, sin embargo aún no podemos afirmar que son en esencia la misma función. Note que por la definición de G

$$G\left(\frac{x}{j}\right) = \sum_{k \leqslant \frac{x}{j}} F\left(\frac{x}{jk}\right), \tag{1.3}$$

por tanto al reemplazar (1.3) en (1.2), obtenemos:

$$\begin{aligned} F(x) &= \sum_{j \leqslant x} \mu(j) \sum_{jk \leqslant x} F\left(\frac{x}{jk}\right) = \sum_{jk \leqslant x} \mu(j) F\left(\frac{x}{jk}\right) \\ &= \sum_{n \leqslant x} F\left(\frac{x}{n}\right) \sum_{jk = n} \mu(j). \end{aligned}$$

Finalmente

$$F(x) = F(x) + \sum_{1 < n \leqslant x} F\left(\frac{x}{n}\right) \sum_{jk = n} \mu(j). \tag{1.4}$$

Para obtener la inversión necesitamos que la doble suma en (1.4) se anule, y dado que no tenemos condiciones sobre F, la función μ debe cumplir que si $n \neq 1$

$$\sum_{jk = n} \mu(j) = 0$$

En efecto, *la función que cumple esta propiedad es... la función de Möbius*.

Teorema 1.3. Sea $n \geqslant 1$, entonces

$$\sum_{d|n} \mu(d) = e(n).$$

Antes de continuar con la prueba de este resultado notemos que la suma en (1.2) en realidad no recorre los $j \leqslant x$, sino los j que son divisores de x ya que G es función aritmética y por tanto $\frac{x}{j}$ es necesariamente un número natural. Así

$$F(x) = \sum_{j|x} \mu(j) G\left(\frac{x}{j}\right) \tag{1.5}$$

Esta suma sobre los divisores de n llevará el nombre de convolución o producto de Dirichlet y nos permitirá darle al conjunto de las funciones aritmética una estructura de monoide abeliano, estas ideas sin embargo las estudiaremos en la siguiente sección. Ahora continuemos con la prueba.

Demostración. Si $n = 1$, entonces $1 = e(1) = \mu(1)$, si $n \neq 1$, entonces por el teorema fundamental de la aritmética $n = \prod_{i=1}^{k} p_i^{\alpha_i}$, note que los únicos divisores d tales que $\mu(d) \neq 0$ son los que toman la forma $d = p_{i_1} \dots p_{i_j}$ donde $\mathcal{K} = \{i_1, \dots, i_j\} \subseteq \{1, \dots, k\}$, en este caso $\mu(d) = (-1)^{|\mathcal{K}|}$. Necesitamos saber cuántas veces va a aparecer este valor en la suma, es decir dado un $0 \leqslant r \leqslant k$ fijo, ¿cuántos subconjuntos de $\{1, \dots, k\}$ tienen cardinal r?, exactamente $\binom{k}{r}$. Así la suma toma la forma:

$$\begin{aligned}
\sum_{d|n} \mu(d) &= 1 + \sum_{i} \mu(p_i) + \sum_{i,j} \mu(p_i p_j) + \dots + \mu(p_1 \dots p_k) \\
&= 1 - k + \binom{k}{2} + \dots + (-1)^k \\
&= \sum_{r=1}^{k} \binom{k}{r} (-1)^r = (1-1)^k = 0.
\end{aligned}$$

□

Aplicando esto a la función $L(x)$, obtenemos que

$$\psi(x) = \sum_{j|n} \mu(j) L\left(\frac{n}{j}\right).$$

La fórmula en (1.5) se conoce como inversión de Möbius, las ideas aquí sin embargo fueron abordadas de manera informal, para poder presentar un argumento riguroso, necesitamos, como se menciono antes, introducir la convolución de Dirichlet, que además nos permitirá obtener propiedades importantes de algunas de las funciones aritmética que hemos presentado en esta sección.

1.2 Convolución de Dirichlet

Siguiendo las ideas de la sección anterior, presentamos la siguiente definición:

Definición. Sean f y g funciones aritméticas. Definimos la convolución o producto de Dirichlet como

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

Algunos resultados del capítulo anterior se pueden escribir en términos de convolución, por ejemplo, el TFA se puede presentar como

$$\log n = \sum_{j|n} \Lambda(j) = \Lambda * 1,$$

donde 1, denota la función constante 1, también $\psi(x) = \mu * L$, pero la convolución no solo se introduce como una manera de simplificar notación, como mencionamos antes, esta tiene propiedades importantes que nos permitirán darle una estructura algebraica a las funciones aritmética.

Teorema 1.4. Sean f y g funciones aritméticas. Entonces se cumple lo siguiente

- $f * g = g * f$.
- $(f * g) * h = f * (g * h)$.
- $e * f = f * e = f$.

Demostración. Primero note que $\sum_{j|n} f(j) g\left(\frac{n}{j}\right) = \sum_{jk=n} f(j) g(k)$, ya que en ambos casos la suma recorre los divisores de n, luego

$$\begin{aligned}(f * g)(n) &= \sum_{j_1 j_2 = n} f(j_1)\, g(j_2) = \sum_{j_1 j_2 = n} g(j_1)\, f(j_2) \\ &= (g * f)(n).\end{aligned}$$

Ya que no importa el orden en el la suma recorra los divisores, lo que prueba la conmutatividad. Ahora, recordemos que $e(n) = 1$ si $n = 1$ y $e(n) = 0$ si $n \neq 1$, tenemos que

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

y como $e\left(\frac{n}{j}\right) = 0$ si $j \neq n$, los términos de la suma son cero excepto cuando $j = n$,

$$(e * f)(n) = (f * e)(n) = \sum_{j|n} f(j) e\left(\frac{n}{j}\right) = f(n) = f.$$

Para probar la asociatividad, considere $N = g * h$ y $M = f * g$, luego

$$\begin{aligned}
(f * N)(n) &= \sum_{j_1 j_2 = n} f(j_1)\, N(j_2) \\
&= \sum_{j_1 j_2 = n} f(j_1) \left(\sum_{j_3 j_4 = j_2} g(j_3)\, h(j_4) \right) \\
&= \sum_{j_1 j_3 j_4 = n} f(j_1)\, g(j_3)\, h(j_4) \\
&= \sum_{j_1 j_3 j_4 = n} f(j_3)\, g(j_4)\, h(j_1) \\
&= \sum_{j_1 j_2 = n} \left(\sum_{j_3 j_4 = j_2} f(j_3)\, g(j_4) \right) h(j_1) \\
&= \sum_{j_1 j_2 = n} M(j_2)\, h(j_1) \\
&= (M * h)(n).
\end{aligned}$$

□

Hemos probado en particular que la función e es el elemento neutro de la convolución, sabemos además que $\mu * 1 = e$, es decir la función de Möbius tiene inverso multiplicativo, con estas nuevas herramientas podemos presentar una prueba corta y rigurosa de la fórmula de inversión de Möbius (1.5).

Teorema 1.5 (Fórmula de inversión de Möbius)**.** Sean f y g funciones aritmética, entonces $f = g * 1$ si y solo si $g = \mu * f$.

Demostración. Note que $f = g * 1$ si y solo si $\mu * f = \mu * g * 1 = g * \mu * 1 = g * e = g$ □

Sin embargo, no toda función aritmética tiene inverso multiplicativo, el caso más evidente es tomar la función constante $N = 0$, note que para toda f, $f * N = N$. Esto nos lleva a la pregunta: ¿bajo qué condiciones una función aritmética tiene inverso multiplicativo?, la respuesta podría venir de estudiar las características que no permiten que N lo tenga... A saber, N *se anula en todo punto*, ¿bastaría con que esta función no se anule en todo su dominio para que tenga inversa?, o ¿en algún punto en particular?, la respuesta nos viene del siguiente teorema, basta con que la función no se anule en 1 para poder garantizar además la *unicidad*.

Teorema 1.6. Sea f una función aritmética tal que $f(1) \neq 0$. Entonces existe una única función aritmética g tal que $f * g = e$.

Demostración. Note que si $n = 1$, entonces $f(1)g(1) = e(1) = 1$, así $g(1) = \dfrac{1}{f(1)}$, ahora supongamos que g se ha definido para todos lo valores $1 < k < n$, en efecto si $f * g(n) = 0$ obtenemos

$$0 = \sum_{d|n} g(d) f\left(\frac{n}{d}\right) = \sum_{\substack{d|n \\ d<n}} f\left(\frac{n}{d}\right) g(d) + f(1)g(n), \tag{1.6}$$

así

$$g(n) = \frac{-1}{f(1)} \sum_{\substack{d|n \\ d<n}} f\left(\frac{n}{d}\right) g(d).$$

Esto nos define g de manera recursiva, lo que concluye el resultado por inducción matemática.

□

Esto nos permite dotar a estas funciones aritmética de una estructura de grupo Abeliano, ya que si $f(1) \neq 0$ y $g(1) \neq 0$, entonces $f * g(1) = f(1)g(1) \neq 0$.

Teorema 1.7. Sean f y g funciones aritméticas multiplicativas, entonces $f * g$ también es multiplicativa.

Demostración. Sean $x, y \in \mathbb{N}$ tal que $(x, y) = 1$. Note que cada divisor $d \mid xy$ puede escribirse de manera única como $d = mn$ donde $m \mid x$ y $n \mid y$, además $(m, n) = 1$ y $\left(\frac{x}{m}, \frac{y}{n}\right) = 1$, por lo tanto

$$\begin{aligned}
(f * g)(xy) &= \sum_{d|xy} f(d) g\left(\frac{xy}{d}\right) \\
&= \sum_{\substack{m|x \\ n|y}} f(mn) g\left(\frac{xy}{mn}\right) \\
&= \sum_{\substack{m|x \\ n|y}} f(m) g\left(\frac{x}{m}\right) f(n) g\left(\frac{y}{n}\right) \\
&= \sum_{m|x} f(m) g\left(\frac{y}{m}\right) \sum_{n|y} f(n) g\left(\frac{y}{n}\right) \\
&= (f * g)(x)(f * g)(y),
\end{aligned}$$

así $f * g$ es multiplicativa. □

Teorema 1.8. Si f es multiplicativa, entonces $g = f^{-1}$ también es multiplicativa.

Donde f^{-1} denota su inversa, presentaremos una prueba siguiendo a [2]

Demostración. Queremos ver que:

$$g(n_1 n_2) = g(n_1) g(n_2) \quad \text{si} \quad (n_1, n_2) = 1. \tag{1.7}$$

Procedamos por inducción matemática. Sea $n = n_1 n_2$, si $n_1 n_2 = 1$, entonces $n_1 = n_2 = 1$, luego

$$g(1 \cdot 1) = g(1) = \frac{1}{f(1)} = 1 = g(1)g(1).$$

Supongamos ahora que g satisface (1.7) para todo $k_1k_2 \geqslant 2$ tal que $k_1k_2 < n$ y sean n_1 y n_2 tales que $n_1n_2 = n$ y $(n_1, n_2) = 1$, por (1.6) tenemos que:

$$\begin{aligned}
0 &= \sum_{d|n_1n_2} f(d)g\left(\frac{n_1n_2}{d}\right) \\
&= \sum_{\substack{d_1|n_1 \\ d_2|n_2 \\ d_1d_2<n}} f(d_1)f(d_2)g\left(\frac{n_1}{d_1}\right)g\left(\frac{n_2}{d_2}\right) + g(n_1n_2) \\
&= \sum_{\substack{d_1|n_1 \\ d_2|n_2}} f(d_1)f(d_2)g\left(\frac{n_1}{d_1}\right)g\left(\frac{n_2}{d_2}\right) + g(n_1n_2) - g(n_1)g(n_2) \\
&= (f * g)(n_1)(f * g)(n_2) + (g(n_1n_2) - g(n_1)g(n_2)),
\end{aligned}$$

luego:

$$g(n_1)g(n_2) = e(n_1)e(n_2) + g(n_1n_2),$$

y como $n_1n_2 \geqslant 2$, entonces $n_1 \geqslant 2$ o $n_2 \geqslant 2$, así $e(n_1)e(n_2) = 0$, por tanto $g(n_1n_2) = g(n_1)g(n_2)$. □

Corolario 1.9. Sea $\mathcal{M}$ el conjunto de funciones aritmética multiplicativas, entonces $(\mathcal{M}, *)$ es un grupo Abeliano.

1.2.1. Propiedades de algunas funciones aritmética

Finalizaremos esta sección con algunas propiedades importantes de las funciones aritmética que definimos el inicio del capítulo.

La función φ de Euler

La propiedad del teorema (1.3) nos permite manipular sumas con condiciones de coprimalidad, es decir sumas sobre los n que son coprimos con un entero k fijo. Considere el conjunto $C_k = \{n \mid (n, k) = 1\}$, note que la función característica del conjunto C_k es:

$$\mathbb{1}_{C_k}(n) = \sum_{d|(n,k)} \mu(d) = e((n, k)).$$

Una aplicación de esto nos permite obtener la siguiente propiedad de la función φ de Euler:

$$\begin{aligned}\varphi(n) &= \sum_{\substack{m\leqslant n\\(m,n)=1}} 1 = \sum_{m\leqslant n} \mathbb{1}_{C_n}(m)\\ &= \sum_{m\leqslant n}\sum_{d|(m,n)} \mu(d)\\ &= \sum_{d|n}\mu(d)\sum_{\substack{m\leqslant n\\ d|m}} 1\\ &= \sum_{d|n}\mu(d)\frac{n}{d}\\ &= \mu * N(n) = n\sum_{d|n}\frac{\mu(d)}{d}.\end{aligned}$$

Es claro que la función N es multiplicativa por definición, luego esta propiedad nos permite probar que la función φ es multiplicativa, por el teorema (1.7) basta ver que en efecto μ lo es.

Teorema 1.10. La función μ es multiplicativa.

Demostración. Supongamos que $n = n_1n_2$, si $n = 1$, entonces $n_1 = n_2 = 1$, $\mu(n) = \mu(n_1)\mu(n_2) = 1$. Ahora supongamos que $\mu(k) = \mu(k_1)\mu(k_2)$, para todo $k = k_1k_2$ tal que $1 < k < n$ y $(k_1, k_2) = 1$ y sean n_1, n_2 tales que $n_1n_2 = n$ y $(n_1, n_2) = 1$, tenemos que

$$\begin{aligned}0 &= \sum_{d|n_1n_2}\mu(d)\\ &= \sum_{\substack{d_1|n_1\\ d_2|n_2\\ d_1d_2<n}}\mu(d_1)\mu(d_2) + \mu(n_1n_2)\\ &= \sum_{d_1|n_1}\mu(d_1)\sum_{d_2|n_2}\mu(d_2) + \mu(n_1n_2) - \mu(n_1)\mu(n_2)\\ &= \mu(n_1n_2) - \mu(n_1)\mu(n_2).\end{aligned}$$

Así, por el principio de inducción matemática se sigue el resultado. □

Corolario 1.11. La función $\varphi(n)$ tiene las siguientes propiedades:

i) $\varphi(mn) = \varphi(m)\varphi(n)$ si $(m, n) = 1$

ii) $\varphi(p^n) = p^n - p^{n-1}$

iii) $\varphi(n) = n\prod_{p|n}\left(1 - \frac{1}{p}\right)$

Demostración. i) Como $\varphi = \mu * N$, se sigue del teorema anterior.

ii) Note que si $n = p^k$, entonces

$$\begin{aligned}\varphi(p^k) &= p^k \sum_{j|p^k} \frac{\mu(j)}{j} \\ &= p^k \left(1 + \frac{\mu(p)}{p} + \frac{\mu(p^2)}{p^2} + \ldots + \frac{\mu(p^k)}{p^k}\right) \\ &= p^k \left(1 - \frac{1}{p}\right) = p^k - p^{k-1}.\end{aligned}$$

iii) Sea $n > 1$, por el TFA se sigue que

$$\begin{aligned}\varphi(n) &= \varphi\left(\prod_{p^m \| n} p^m\right) \\ &= \prod_{p^m \| n} \varphi(p^m) \\ &= \prod_{p^m \| n} p^m - p^{m-1} \\ &= n \prod_{p|n} \left(1 - \frac{1}{p}\right).\end{aligned}$$

□

Al inicio del capítulo mencionamos que las funciones aritmética están totalmente determinadas por sus valores en las potencias de primos, esta propiedad nos permite probar de manera sencilla afirmaciones del estilo $f * g = h$, siempre que f, g y h sean funciones multiplicativas, basta ver que $f * g(p^m) = h(p^m)$, veamos un ejemplo:

Teorema 1.12. La función $\varphi(n)$ satisface la propiedad:

$$n = \sum_{j|n} \varphi(j)$$

Demostración. La afirmación se puede escribir como $N = \varphi * 1$, como estas funciones son multiplicativas, entonces basta ver que la identidad se tiene en las potencias de primos, en efecto

$$\begin{aligned}\sum_{j|p^m} \varphi(j) &= \varphi(1) + \varphi(p) + \varphi(p^2) + \varphi(p^3) + \ldots + \varphi(p^m) \\ &= 1 + (p - 1) + (p^2 - p) + (p^3 - p^2) + /\ldots/ + (p^m - p^{m-1}) \\ &= p^m.\end{aligned}$$

□

Esta identidad también puede probarse usando propiedades de la convolución

$$\sum_{j|n} \varphi(j) = \varphi * 1(n) = (N * \mu) * 1(n) = N * (\mu * 1)(n) = N * e(n) = n.$$

Las funciones número y suma de divisores

También podemos aplicar la convolución para obtener también propiedades de la funciones suma y número de divisores, por ejemplo:

$$\sigma(n) = \sum_{j|n} j = N * 1(n) \quad \text{y} \quad \tau(n) = \sum_{j|n} 1 = 1 * 1(n). \tag{1.8}$$

Note que $\sigma * \varphi = (N * 1) * (\mu * N) = (N * e) * N = N * N$, en efecto:

$$\sigma * \varphi(n) = N * N(n) = \sum_{j|n} j \cdot \frac{n}{j} = n \sum_{j|n} 1 = n\tau(n).$$

Con lo que obtenemos una propiedad interesante que relaciona estas 3 funciones aritmética, pero además (1.8) nos dice también que las funciones σ y τ son multiplicativas por ser convolución de funciones multiplicativas. Observemos una última propiedad de estas funciones, que nos permite caracterizar la noción de primalidad.

Proposición 1.13. Un entero n es primo si y solo si $\sigma(n) + \varphi(n) = n\tau(n)$.

Demostración. Si n es primo, entonces $\varphi(n) = n - 1$, $\sigma(n) = n + 1$ y $\tau(n) = 2$, luego es claro que $\sigma(n) + \varphi(n) = n\tau(n)$. Veamos ahora que si n no es primo entonces no se sigue el teorema

Si n no es primo entonces $\varphi(n) < n - 1$, ahora note que

$$\sigma(n) = \sum_{j|n} j = 1 + \sum_{\substack{j|n \\ j>1}} j \leqslant 1 + n(\tau(n) - 1).$$

Luego:

$$\sigma(n) + \varphi(n) < n - 1 + 1 + n\tau(n) - n = n\tau(n).$$

□

1.3 Sumación Parcial

Los resultados que hemos podido obtener hasta el momento solo se centran en casos finitos, en sumas sobre los divisores de un entero n fijo o sobre los k que son primos relativos a n, estos son casos privilegiados, nuestro objetivo sigue siendo el TNP, para esto necesitamos poder obtener relaciones asintóticas, algunas funciones como μ o φ aparentan comportamientos caóticos al graficarlas en función de n y por tanto no tiene

mucho sentido estudiar un comportamiento asintótico para ellas, sin embargo, algunas funciones aritmética $f(n) = a_n$ tienen buen comportamiento en la media, en el sentido de que sus sumas parciales:

$$A(x) = \sum_{n \leqslant x} a_n.$$

Tienden a ser suaves conforme $x \to \infty$ y frecuentemente podemos estudiarlas de manera precisa, ejemplo de esto son $\pi(x)$ o $\psi(x)$. En esta sección estudiaremos algunos de los métodos principales para obtener dichas estimaciones, estimaciones que nos darán un camino a la prueba del TNP.

Definición. Sea f una función real definida en $[a, b]$. Suponga que $f(x+)$ y $f(x-)$ existen para todo $x \in (a, b)$. Definimos

- $f(x) - f(x-)$: Salto a izquierda de f en x.
- $f(x+) - f(x)$: Salto a derecha de f en x.
- $[f(x) - f(x-)] + [f(x+) - f(x)] = f(x+) - f(x-)$: Salto de f en x.

Nota. Diremos que $f(x) = O(g(x))$ cuando existen constantes $M > 0$ y x_0 tales que para todo $x > x_0$, se cumple que $|f(x)| \leqslant M|g(x)|$, es decir, nuestro dominio específico es $x > x_0$.

Una **fórmula asintótica** para $f(x)$ es una expresión de la forma $f(x) \sim g(x)$, mientras que una **estimación asintótica** para $f(x)$ es una expresión del tipo $f(x) = g(x) + O(R(x))$, donde $g(x)$ representa el término principal y $R(x)$ el término de error. La notación O grande nos permite entonces controlar el error de estimaciones asintóticas de manera precisa, cosa que no ocurre con o pequeña, recordemos que:

$$f(x) = o(g(x)) : \quad \lim_{x \to \infty} \frac{f(x)}{g(x)} = 0.$$

Esto intuitivamente nos dice que $f(x)$ "crece más lento" que $g(x)$, pero esto no nos da tanta información acerca de f con O grande, ya que no sabemos la "velocidad" con la que este cociente tiende a 0.

Las estimaciones con O también son más fáciles de trabajar y manipular que las estimaciones con o. Por ejemplo, las O se pueden "sacar" de integrales o sumas, siempre que las funciones involucradas sean no negativas, mientras que tales manipulaciones generalmente no se permiten con las o.

1.3.1. La integral de Riemann-Stieltjes

Sean $P = \{x_0, x_1, ..., x_n\} \in \mathcal{P}[a, b]$ (las particiones del intervalo $[a, b]$) y $t_k \in [x_{k-1}, x_k]$ cualesquiera. Una suma de la forma

$$S(P, f) = \sum_{k=1}^{n} f(t_k)\Delta x_k,$$

se denomina **suma de Riemann** de f en $[a, b]$, recordamos que $\Delta x_k = x_k - x_{k-1}$.

La integral usual se define como el límite de las sumas de Riemann, esta definición nos da una correspondencia importante: Toda integral de Riemann se puede ver como el límite de una suma, una serie.

Sin embargo, no toda suma se puede ver como una integral bajo esta definición, si tenemos una con una condición de sumación sobre, por ejemplo, los números primos, no hay un camino claro para expresarla como una integral, la solución a este problema viene de generalizar la integral de Riemann.

La propiedad deseada la obtendremos de la integral de Riemann-Stieltjes.

Definición.

i) Sean $P = \{x_0, x_1, ..., x_n\} \in \mathcal{P}[a, b]$ (las particiones del intervalo $[a, b]$) y $t_k \in [x_{k-1}, x_k]$ cualesquiera.

Una suma de la forma

$$S(P, f, \alpha) = \sum_{k=1}^{n} f(t_k)\Delta\alpha_k,$$

se denomina **suma de Riemann-Stieltjes** de f con respecto a α en $[a, b]$.

ii) Decimos que f es **Riemann-Integrable** con respecto a α en $[a, b]$, y escribimos "$f \in \mathcal{R}(\alpha)$ en $[a, b]$", si existe $A \in \mathbb{R}$ que satisface que
Para todo $\varepsilon > 0$ existe $P_\varepsilon \in \mathcal{P}[a, b]$ tal que si para toda $P \supset P_\varepsilon$ y para cualquier elección de puntos $t_k \in [x_{k-1}, x_k]$, entonces

$$| S(P, f, \alpha) - A | < \varepsilon.$$

Donde $\Delta\alpha_k = \alpha(x_k) - \alpha(x_{k-1})$, cuando A existe, es único y se denota por

$$\int_a^b f d\alpha \quad \text{o} \quad \int_a^b f(x) d\alpha(x).$$

La función f es llamada *integrando* y la función α es llamada *integrador*.

Note que la integral de Riemann no es más que un caso particular de la de Riemann-Stieltjes, cuando $\alpha(x) = x$, esta integral y sus propiedades se suelen estudiar en un curso

de Análisis II, aquí recordaremos algunas de las más importantes.

Teorema 1.14 ([4], Teorema 7.11)**.** Sea α una función escalonada definida en $[a, b]$ con salto α_k en x_k.

Sea f una función definida en $[a, b]$ tal que f y α no sean ambas discontinuas a la derecha o a la izquierda de cada x_k. Entonces $\int_a^b f d\alpha$ existe y se tiene que:

$$\int_a^b f(x)d\alpha(x) = \sum_{k=1}^{n} f(x_k)\,\alpha_k.$$

La prueba de esto se encuentra en [4], sin embargo, nos será útil explorar la idea.

Note que α es constante en los intervalos (x_{k-1}, x_k), luego por intervalos la integral es 0 ya que para cualquier suma de Riemann-Stieltjes, $S(P, f, \alpha)$=0. Así, para conocer el valor de la integral entonces solo tendríamos que sumar el valor que toma alrededor de cada x_k.

Supongamos que α tiene salto α_k en un punto $x_k \in [a, b]$, no es difícil ver que:

$$\int_a^b f d\alpha = f(x_k)[\alpha(x_k+) - \alpha(x_k-)] = f(x_k)\alpha_k. \tag{1.9}$$

La prueba de esto se encuentra también en [4] [Teorema 7.9], en efecto si repetimos esto para cada x_k obtenemos la suma deseada.

Sea f continua en $[0, N]$, el teorema anterior nos permite expresar la suma $\sum_{n=1}^{N} a_n f(n)$ como una integral de Riemann Stieltjes, a saber

$$\sum_{n=1}^{N} a_n f(n) = \int_0^N f(x)dA(x).$$

Ya que al aplicar (1.9) tomando como integrador las sumas parciales tenemos que $A(n+) - A(n-) = A(n) - A(n-1) = a_n$, luego en cada paso la integral toma el valor $a_n f(n)$.

Nota. El límite inferior de la integral puede ser cualquier número en el intervalo [0, 1) y el superior cualquier número en [N, N+1) sin afectar el valor de la integral, la siguiente notación es empleada

$$\sum_{n=1}^{N} a_n f(n) = \int_0^N f(x)dA(x) = \int_{1^-}^N f(x)dA(x)$$

Ejemplo.

- Si se toma $\alpha(x) = [x]$, entonces

$$\int_a^b f(x)d[x] = \sum_{a<n\leqslant b} f(n).$$

- Si se toma $\alpha(x) = \pi(x)$, la función contadora de primos, que tiene un salto de 1 en cada p primo, entonces

$$\int_a^b f(x)d\pi(x) = \sum_{a<p\leqslant b} f(p).$$

1.3.2. Algunas propiedades de la integral de Riemann-Stieltjes

Teorema 1.15 (Integración por partes)**.** Si $f \in \mathcal{R}(\alpha)$ en $[a, b]$ entonces $\alpha \in \mathcal{R}(f)$ en $[a, b]$ y

$$\int_a^b f(x)d\alpha(x) + \int_a^b \alpha(x)df(x) = f(b)\alpha(b) - f(a)\alpha(a).$$

Bajo ciertas condiciones una integral de Riemann-Stieltjes se puede reducir a una integral de Riemann usual, esto es muy útil puesto que estamos más familiarizados con el cálculo de estas, dichas condiciones son presentadas en el siguiente teorema:

Teorema 1.16. Sea $f \in R(\alpha)$ en $[a, b]$, donde $\alpha \in C^1[a, b]$, entonces, $\int_a^b f(x)\alpha'(x)dx$ existe y

$$\int_a^b f\, d\alpha = \int_a^b f(x)\alpha'(x)dx.$$

Las pruebas de estos teoremas se encuentran también en [4], por lo que no las presentaremos aquí para no extendernos demasiado en la teoría de esta sección, con estas propiedades ya podemos presentar una prueba del teorema de sumación de Abel.

Teorema 1.17 (Sumación parcial de Abel)**.** Sea a_n función aritmética y $f \in C^1[1, x]$, entonces

$$\sum_{n\leqslant x} a_n f(n) = A(x)f(x) - \int_1^x A(t)f'(t)dt.$$

Demostración. Tenemos que:

$$\sum_{n\leqslant x} a_n f(n) = \int_0^x f(t)dA(t) = f(x)A(x) - \int_0^x A(t)df(t),$$

ya que $A(x) = 0$ para todo $x \in [0,1)$, en efecto

$$\begin{aligned}\sum_{n\leqslant x} a_n f(n) &= f(x)A(x) - \int_1^x A(t)df(t)\\ &= f(x)A(x) - \int_1^x A(t)f'(t)dt,\end{aligned}$$

por el teorema 1.16. □

1.4 Algunas estimaciones básicas

Vamos a ver algunas aplicaciones de la teoría que hemos presentado, algunas de estas ya las habíamos mencionado antes, por ejemplo, el comportamiento asintótico de las sumas de la serie armónica.

Vamos a ver que $\sum_{n\leqslant x} \frac{1}{n} = \log x + \gamma + O\left(\frac{1}{x}\right)$, donde $\gamma = \lim_{n\to\infty}\left(\sum_{n=1}^{n} \frac{1}{n} - \log n\right)$ es conocida como la constante de Euler-Mascheroni.

Sea $a_n = 1$ y $f(x) = \frac{1}{x}$, así $A(x) = [x]$ y al usar la sumación parcial tenemos que

$$\begin{aligned}\sum_{n\leqslant x} \frac{1}{n} &= \frac{[x]}{x} + \int_1^x \frac{[t]}{t^2}dt\\ &= \frac{x-\{x\}}{x} + \int_1^x \frac{t-\{t\}}{t^2}dt.\end{aligned}$$

Donde $\{x\}$ denota la parte fraccionaria de x, es decir $\{x\} = x - [x]$, note que $\{x\} = O(1)$, luego

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

Basta ver que 1-$\int_1^\infty \frac{\{x\}}{x^2}dx = \gamma$ ya que la otra integral también es del orden de $O\left(\frac{1}{x}\right)$, en efecto

$$\begin{aligned}\gamma &= \lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{1}{k}-\log n\right)\\ &= \lim_{n\to\infty}\left(1+\sum_{1<k\leqslant n}\frac{1}{k}-\log n\right)\\ &= 1+\lim_{n\to\infty}\left(\int_1^n\frac{1}{x}d[x]-\int_1^n\frac{1}{x}dx\right)\\ &= 1+\lim_{n\to\infty}\left(\int_1^n\frac{[x]}{x^2}dx-\int_1^n\frac{1}{x}dx\right)\\ &= 1-\int_1^\infty\frac{\{x\}}{x^2}dx.\end{aligned}$$

Esto concluye el resultado que mencionamos en la introducción, $H(n)\sim\log n$. Veamos otro ejemplo:

Ejemplo. Estimación de las sumas parciales de $\log n$:

Sea $a_n = 1$ y $f(x) = \log x$, así $A(x) = [x]$, luego:

$$\begin{aligned}\sum_{n\leqslant x}\log n &= [x]\log x-\int_1^x\frac{[t]}{t}dt\\ &= (x-O(1))\log x-\int_1^x\frac{t-O(1)}{t}dt\\ &= x\log x-O(\log x)-(x-1)+O(\log x)\\ &= x\log x-x+O(\log x).\end{aligned}$$

1.4.1. Una equivalencia importante

Vamos a obtener finalmente que el TNP es equivalente a la afirmación $\psi(x)\sim x$

Definición. Sea $x\in\mathbb{N}$, con $x>1$, definimos la función contadora de primos $\pi(x)$ como:

$$\pi(x)=\sum_{p\leqslant x}1.$$

Como vimos antes, la fórmula de sumación de Abel tiene un gran poder teórico que explotaremos en distintos lugares de este trabajo. Por lo pronto ella será esencial para para estimar $\vartheta(x)$ y $\pi(x)$, de donde obtendremos la equivalencia deseada.

Teorema 1.18. Dado $x \geqslant 2$

$$\vartheta(x) = \pi(x)\log x - \int_2^x \frac{\pi(t)}{t}dt,$$
$$\pi(x) = \frac{\vartheta(x)}{\log x} + \int_2^x \frac{\vartheta(t)}{t\log^2 t}dt.$$

Demostración. Sea $\mathbb{1}_p$ la función característica del conjunto de primos. Tenemos entonces las siguientes igualdades:

$$\vartheta(x) = \sum_{n\leqslant x} \mathbb{1}_p(n)\log(n) \quad \text{y} \quad \pi(x) = \sum_{n\leqslant x} \mathbb{1}_p(n).$$

Fijemos $x \geqslant 2$. Por la fórmula de sumación de Abel y como $\pi(t) = 0$ para todo $t < 2$, tenemos que

$$\begin{aligned}\vartheta(x) &= \pi(x)\log(x) - \int_1^x \frac{\pi(t)}{t}dt \\ &= \pi(x)\log(x) - \int_2^x \frac{\pi(t)}{t}dt.\end{aligned}$$

Notando que $\pi(x) = \sum_{n\leqslant x} \frac{\mathbb{1}_p(n)\log(n)}{\log(n)}$ y que $\vartheta(t) = 0$ para todo $t < 2$ tenemos que

$$\begin{aligned}\pi(x) &= \frac{\vartheta(x)}{\log(x)} + \int_1^x \frac{\vartheta(t)}{t\log^2(t)}dt \\ &= \frac{\vartheta(x)}{\log(x)} + \int_2^x \frac{\vartheta(t)}{t\log^2(t)}dt.\end{aligned}$$

□

Ahora vamos a establecer una conexión entre las dos funciones de Chebyshev que definimos antes, en efecto

$$\psi(x) = \sum_{n\leqslant x}\Lambda(n) = \sum_{m=1}^{\infty}\sum_{p^m\leqslant x}\Lambda(p^m) = \sum_{m=1}^{\infty}\sum_{p\leqslant x^{\frac{1}{m}}}\log p.$$

Note que la suma sobre m realmente es finita porque la suma sobre p se detiene cuando $x^{\frac{1}{m}} < 2$, es decir, cuando $m > \log_2(x)$, entonces

$$\psi(x) = \sum_{m\leqslant\log_2 x}\sum_{p\leqslant x^{\frac{1}{m}}}\log p = \sum_{m\leqslant\log_2 x}\vartheta\left(x^{\frac{1}{m}}\right). \tag{1.10}$$

Teorema 1.19. Si $x > 0$ se tiene que:

$$0 \leqslant \frac{\psi(x)}{x} - \frac{\vartheta(x)}{x} \leqslant \frac{\log^2 x}{\sqrt{x}\log 4}.$$

Demostración. Por (1.10) se sigue que:

$$\psi(x) = \sum_{2 \leqslant m \leqslant \log_2 x} \vartheta\left(x^{\frac{1}{m}}\right) + \vartheta(x),$$

ya que si $m = 1$ entonces $\vartheta(x^{\frac{1}{m}}) = \vartheta(x)$, luego

$$\psi(x) - \vartheta(x) = \sum_{2 \leqslant m \leqslant \log_2 x} \vartheta\left(x^{\frac{1}{m}}\right) \geqslant 0.$$

Ahora por definición de ϑ

$$\vartheta(x) = \sum_{p \leqslant x} \log p \leqslant x \log x,$$

entonces

$$\begin{aligned} 0 \leqslant \psi(x) - \vartheta(x) &\leqslant \sum_{2 \leqslant m \leqslant \log_2 x} x^{\frac{1}{m}} \log x^{\frac{1}{m}} \\ &\leqslant \sqrt{x} \sum_{2 \leqslant m \leqslant \log_2 x} \log x^{\frac{1}{m}} \\ &\leqslant \sqrt{x}(\log_2(x) \log(\sqrt{x})) \\ &= \frac{\sqrt{x} \log^2 x}{\log(4)}, \end{aligned}$$

y dividiendo por x obtenemos el resultado □

Teorema 1.20. La afirmación $\psi(x) \sim x$ es equivalente a $\vartheta(x) \sim x$.

Demostración. Por el teorema anterior:

$$0 \leqslant \frac{\psi(x)}{x} - \frac{\vartheta(x)}{x} \leqslant \frac{\log^2(x)}{\sqrt{x} \log(4)}$$

y como $\lim_{x \to \infty} \frac{\log^2(x)}{\sqrt{x} \log(4)} = 0$, entonces cuando $x \to \infty$ se tiene que $\frac{\psi(x)}{x} - \frac{\vartheta(x)}{x} = 0$, lo que concluye el resultado. □

Teorema 1.21. Las siguientes afirmaciones son equivalentes:

i) $\pi(x) \sim \dfrac{x}{\log(x)}$.

ii) $\vartheta(x) \sim x$.

iii) $\psi(x) \sim x$.

Por el teorema anterior basta ver que i es equivalente a ii

Demostración. Por el teorema 1.18 y dado que estamos trabajando con aproximaciones asintóticas, podemos asumir que $x \geqslant 2$, se sigue que

$$\frac{\vartheta(x)}{x} = \frac{\pi(x)\log(x)}{x} - \frac{1}{x}\int_2^x \frac{\pi(t)}{t}dt. \tag{1.11}$$

$$\frac{\pi(x)\log(x)}{x} = \frac{\vartheta(x)}{x} + \frac{\log(x)}{x}\int_2^x \frac{\vartheta(t)}{t\log^2(t)}dt. \tag{1.12}$$

Basta con ver que las integrales de (1.11) y (1.12) van a 0 , cuando $x \to \infty$.

$(\rightarrow)$ Por hipótesis $\frac{\pi(x)}{x}\left(\frac{1}{\log(x)}\right)^{-1} = 1$ cuando $x \to \infty$. Esto es equivalente a decir que $\frac{\pi(t)}{t} = O\left(\frac{1}{\log(t)}\right)$. Luego para todo $x \geqslant 2$ positivo fijo tenemos que

$$\frac{1}{x}\int_2^x \frac{\pi(t)}{t}dt = O\left(\frac{1}{x}\int_2^x \frac{1}{\log(t)}dt\right).$$

Note que

$$\begin{aligned}\int_2^x \frac{1}{\log(t)}dt &\leqslant \int_2^{\sqrt{x}} \frac{1}{\log(t)}dt + \int_{\sqrt{x}}^x \frac{1}{\log(t)}dt \\ &\leqslant \frac{\sqrt{x}}{\log(2)} + \frac{x-\sqrt{x}}{\log(\sqrt{x})},\end{aligned}$$

luego

$$\frac{1}{x}\int_2^x \frac{1}{\log(t)}dt \leqslant \frac{1}{\sqrt{x}\log(2)} + \frac{1}{\log(\sqrt{x})} - \frac{1}{\sqrt{x}\log(\sqrt{x})},$$

así, cuando $x \to \infty$, $\frac{1}{x}\int_2^x \frac{1}{\log(t)}dt = 0$.

$(\leftarrow)$ Análogamente $\vartheta(t) = O(t)$. Por tanto

$$\frac{\log(x)}{x}\int_2^x \frac{\vartheta(t)}{\log^2(t)}dt = O\left(\frac{\log(x)}{x}\int_2^x \frac{1}{\log^2(t)}dt\right),$$

La integral en O se puede acotar de manera análoga a la anterior:

$$\begin{aligned}\int_2^x \frac{1}{\log^2(t)}dt &= \int_2^{\sqrt{x}} \frac{1}{\log^2(t)}dt + \int_{\sqrt{x}}^x \frac{1}{\log^2(t)}dt \\ &\leqslant \frac{\sqrt{x}}{\log^2(2)} + \frac{x-\sqrt{x}}{\log^2(\sqrt{x})}\end{aligned}.$$

Multiplicando ambos lados por $\frac{\log(x)}{x}$, podemos ver que si $x \to \infty$, $\frac{\log(x)}{x}\int_2^x \frac{1}{t\log^2(t)} = 0$. □

Finalizaremos con una aplicación interesante de la fórmula de sumación de Abel:

Proposición 1.22 (Fórmula de Stirling)**.** Si n es un entero positivo, entonces

$$n! = C\sqrt{n}n^n e^{-n}\left(1 + O\left(\frac{1}{n}\right)\right).$$

Demostración. Por el teorema 1.17:

$$\begin{aligned}\sum_{k\leqslant n} \log k &= n\log n - n + 1 + \int_1^n \frac{\{t\}}{t}dt \\ &= n\log n - n + 1 + \frac{1}{2}\log n + \int_1^n \frac{s(t)}{t}dt,\end{aligned}$$

donde $s(t) = \{t\} - \frac{1}{2}$, note que aplicando integración por partes

$$\int_1^n \frac{s(t)}{t} = \left.\frac{S(t)}{t}\right|_1^n + \int_1^n \frac{S(t)}{t^2}dt,$$

donde $S(t) = \displaystyle\int_1^t s(y)dy$, tenemos que $\left.\dfrac{S(t)}{t}\right|_1^n = 0$ ya que $s(t)$ es una función periódica de periodo 1 y $\displaystyle\int_k^{k+1} s(t)dt = 0$ para todo entero k. Note que $|S(t)| \leqslant \dfrac{1}{2}$, de esto se sigue que

$$\begin{aligned}\int_1^n \frac{s(t)}{t}dt = \int_1^n \frac{S(t)}{t^2} &= \int_1^\infty \frac{S(t)}{t^2}dt - \int_n^\infty \frac{S(t)}{t^2}dt \\ &= c + O\left(\frac{1}{n}\right). \qquad \left(|S(t)| \leqslant \frac{1}{2}\right)\end{aligned}$$

Por tanto:

$$\sum_{k\leqslant n} \log k = n\log n - n + \frac{1}{2}\log n + c + O\left(\frac{1}{n}\right),$$

en efecto

$$\begin{aligned}n! &= exp\left(\sum_{k\leqslant n} \log k\right) \\ &= exp\left(n\log n - n + \frac{1}{2}\log n + c + O\left(\frac{1}{n}\right)\right) \\ &= C\sqrt{n}n^n e^{-n}\left(1 + O\left(\frac{1}{n}\right)\right).\end{aligned}$$

□

Se puede ver que $C = \sqrt{2\pi}$, sin embargo no abordaremos eso en este trabajo.

1.5 Series de Dirichlet

Dada una función aritmética f, se le pueden asignar a esta dos tipos de series importantes:

$$E(z) = \sum_{n=1}^{\infty} f(n)z^n,$$
$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s},$$

estas son llamadas funciones generatrices de f. Las series de potencia generatrices fueron introducidas por Euler con el propósito de estudiar problemas de naturaleza aditiva, esto ya que tienen la siguiente propiedad:

$$E_f(z)E_g(z) = \sum_{n=1}^{\infty} h(n)z^n,$$

donde

$$h(n) = \sum_{a+b=n} f(a)g(b).$$

Note que $h(n)$ se parece la convolución de Dirichlet, pero en una versión aditiva, por eso estas series toman un papel importante para este tipo de problemas.

Sin embargo, nuestro interés son las series $F(s)$, estas fueron introducidas por Dirichlet en su trabajo sobre primos en progresiones aritmética y son particularmente útiles cuando f es una función multiplicativa ya que entre muchas cosas, en el producto de dos de estas aparece la convolución multiplicativa que vimos antes, por lo que a través de estas podemos obtener propiedades de $f(n)$ como en la sección 1.2.

Con lo que se mencionó antes parece que toda la teoría se conecta, y que de cierta forma esto solo es una forma diferente de atacar los mismos problemas, sin embargo, este no es el caso, las propiedades analíticas de una serie de Dirichlet, vista como función de variable compleja s pueden ser explotadas para obtener información importante acerca del comportamiento de las sumas parciales de funciones aritmética

$$\sum_{n \leqslant x} f(n)$$

y que no podemos obtener con la sumación parcial.

Definición. Sea f una función aritmética, la serie

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s},$$

es llamada *serie de Dirichlet asociada a* f.

La variable s es usualmente escrita como $s = \sigma + it$ con $\sigma = \Re(s)$ y $t = \mathrm{Im}(s)$, la serie de Dirichlet más popular es la función zeta de Riemann $\zeta(s)$, definida como la serie de Dirichlet asociada a la constante 1:

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \quad (\sigma > 1)$$

1.5.1. Propiedades algebraicas

Teorema 1.23. Sean f y g las funciones aritmética asociadas a $F(s)$ y $G(s)$. Sea $h = f * g$ la convolución de f y g y $H(s)$ su serie de Dirichlet asociada. Si $F(s)$ y $G(s)$ convergen absolutamente en algún punto s, entonces también $H(s)$ y $H(s) = F(s)G(s)$.

Demostración. Note que por la convergencia absoluta:

$$\begin{aligned} F(s)G(s) &= \sum_{k=1}^{\infty} \frac{f(k)}{k^s} \sum_{m=1}^{\infty} \frac{g(m)}{m^s} \\ &= \sum_{k=1}^{\infty} \sum_{m=1}^{\infty} \frac{f(k)g(m)}{k^s m^s} \\ &= \sum_{n=1}^{\infty} \frac{1}{n^s} \sum_{km=n} f(k)g(m) = \sum_{n=1}^{\infty} \frac{f * g(n)}{n^s}, \end{aligned}$$

además:

$$\begin{aligned} \sum_{n=1}^{\infty} \left| \frac{h(n)}{n^s} \right| &\leqslant \sum_{n=1}^{\infty} \frac{1}{|n^s|} \sum_{km=n} |f(k)||g(m)| \\ &= \left(\sum_{k=1}^{\infty} \left| \frac{f(k)}{k^s} \right| \right) \left(\sum_{m=1}^{\infty} \left| \frac{g(m)}{m^s} \right| \right) \\ &< \infty. \end{aligned}$$

□

Es muy importante tener en cuenta que esto no se puede garantizar sin la convergencia absoluta, ya que los reordenamientos que realizamos en la prueba no son posibles.

Corolario 1.24. Sean f una función aritmética con serie de Dirichlet asociada $F(s)$, g tal que $f * g = e$ y $G(s)$ la serie de Dirichlet asociada a g, entonces $G(s) = \dfrac{1}{F(s)}$ en cualquier punto s en el que $F(s)$ y $G(s)$ sean ambas absolutamente convergentes.

Demostración. Note que la serie de Dirichlet asociada a e es 1, luego por el teorema anterior:

$$1 = \sum_{n=1}^{\infty} \frac{e(n)}{n^s} = \sum_{n=1}^{\infty} \frac{f * g(n)}{n^s} = F(s)G(s).$$

□

Nota. La convergencia absoluta de $F(s)$ no implica la de la serie de Dirichlet asociada con su inversa. Por ejemplo, la función definida por $f(1) = 1$, $f(2) = -1$, y $f(n) = 0$ para $n \geqslant 3$ tiene la serie de Dirichlet $F(s) = 1 - 2^{-s}$, que converge para todo s. Sin embargo, la serie de Dirichlet de la inversa de Dirichlet de f es $\frac{1}{F(s)} = (1 - 2^{-s})^{-1} = \sum_{k=0}^{\infty} \frac{1}{2^{ks}}$, que converge absolutamente en $\sigma > 0$, pero no en el semiplano $\sigma \leqslant 0$, ya que es una serie geométrica. [2]

Si tomamos una serie de potencias $F(x)$, sabemos por lo menos tres cosas de ella:

- F tiene un disco de convergencia,
- En el interior del disco, F converge absolutamente.
- En el interior del disco, F es una función analítica [4]

Nuestro objetivo ahora es lograr obtener propiedades similares para series de Dirichlet.

1.5.2. Propiedades analíticas

Primero note que $x^s = e^{s \log x} = e^{(\sigma + it) \log x} = x^{\sigma} e^{it \log x}$, por tanto $|x^s| = |x^{\sigma}|$ ya que

$$\begin{aligned} \left|e^{i\theta}\right| &= |\cos(\theta) + i \sin(\theta)| \\ &= 1, \end{aligned}$$

luego $\left|e^{it \log x}\right| = 1$. Esto muestra que la convergencia absoluta de una serie de Dirichlet está determinada únicamente por σ.

Ejemplo (Convergencia de $\zeta(s)$)**.** Note que

$$\sum_{n=1}^{\infty} \left|\frac{1}{n^s}\right| = \sum_{n=1}^{\infty} \frac{1}{n^{\sigma}}.$$

Si $\sigma = 1$ sabemos que la serie diverge, en otro caso, la serie converge si y solo si la integral:

$$\int_1^{\infty} \frac{1}{x^{\sigma}} dx = \lim_{n \to \infty} \frac{x^{1-\sigma}}{1-\sigma}\Big|_1^n = \lim_{n \to \infty} \frac{n^{1-\sigma}}{1-\sigma} - \frac{1}{1-\sigma}.$$

Luego, es claro que $\zeta(s)$ converge absolutamente si $\sigma > 1$ y diverge si $\sigma \leqslant 1$.

Definición. Sea Ω un abierto en $\mathbb{C}$ y $f : \Omega \to \mathbb{C}$, la *derivada* de f en $z_0 \in \Omega$ está dada por

$$f'(z_0) = \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

Si el límite existe. Diremos que f es *diferenciable* en z_0.

Si f es diferenciable en todo punto de una vecindad de z_0, entonces diremos que f es *holomorfa* en z_0. Si f es holomorfa en todo punto de un abierto $U \subseteq \Omega$, diremos que f es holomorfa en U. Finalmente si f es holomorfa en Ω, diremos que es holomorfa.

Por otro lado, f se dice analítica en z_0, si existe un abierto $U \subseteq \Omega$ tal que $z_0 \in U$ y

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n,$$

para todo $z \in U$, si f es analítica en todo punto de U, entonces decimos que es analítica en U, si f es analítica en todo el plano complejo diremos que es una función entera.

El conjunto de puntos $s = \sigma + it$ con $\sigma > a$ es llamado semiplano, veremos que toda serie de Dirichlet tiene un semiplano $\sigma > \sigma_c$ en el que la serie converge y otro semiplano $\sigma > \sigma_a$ en el que la serie converge absolutamente. Las constantes σ_c y σ_a son llamas abscisa de convergencia y abscisa de convergencia absoluta respectivamente. Si la serie de Dirichlet converge para todo $s \in \mathbb{C}$, decimos que $\sigma_c = -\infty$, si no converge para todo $s \in \mathbb{C}$, decimos que $\sigma_c = \infty$, de manera análoga lo haremos con la convergencia absoluta.[5]

Teorema 1.25. Sea $f : \mathbb{N} \to \mathbb{C}$, supongamos que la serie

$$F(s) = \sum_{n=1}^{\infty} \left| \frac{f(n)}{n^s} \right|,$$

no converge, ni diverge para todo $s \in \mathbb{C}$, entonces existe un $\sigma_a \in \mathbb{R}$ tal que $F(s)$ converge para todo $\sigma > \sigma_a$ y diverge para todo $\sigma < \sigma_a$

Demostración. Sea $|n^s| = n^\sigma$, si $\sigma > \sigma_0$, entonces $|n^s| \geqslant n^{\sigma_0}$, así $|f(n)n^{-s}| \leqslant |f(n)|n^{-\sigma_0}$. Por el criterio de comparación, si $F(s)$ converge para $s = \sigma_0 + it_0$, entonces converge para todo s con parte real $\sigma \geqslant \sigma_0$. Considere el conjunto

$$A : \left\{ a \in R : \sum_{n=1}^{\infty} \left| \frac{f(n)}{n^a} \right| < \infty \right\}.$$

Como $F(s)$ no diverge para todo $s \in \mathbb{C}$, entonces A es no vacío, además tiene un mínimo ya que $F(s)$ no converge para todo $s \in C$, sea σ_a el mínimo de A. Si $s \in C$ tiene parte real $\sigma < \sigma_a$, entonces $\sigma \notin A$, $F(s)$ diverge. Análogamente si s tiene parte real $\sigma > \sigma_a$, entonces $\sigma > a$ para algún $a \in A$, $F(s)$ converge.

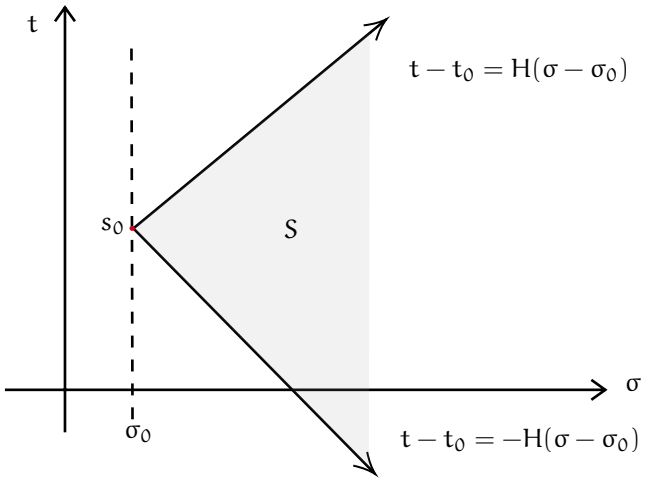
□

Teorema 1.26. Sean Ω un abierto en $\mathbb{C}$, (f_n) una sucesión de funciones analíticas definidas en Ω, y $f : \Omega \to \mathbb{C}$. Si (f_n) converge absolutamente a f en cualquier subconjunto compacto de Ω, entonces f es analítica en Ω y la sucesión (f'_n) también converge uniformemente a f' en cualquier subconjunto compacto de Ω.

No entraremos en los detalles de esta prueba aquí, el lector interesado la puede consultar en [6].

Teorema 1.27. Suponga que la serie $F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$ converge en el punto $s = s_0$ y sea $H > 0$, entonces $F(s)$ converge uniformemente en S, donde S es el conjunto:

$$S = \{s \in \mathbb{C} : \sigma > \sigma_0 \text{ y } |t - t_0| \leqslant H(\sigma - \sigma_0)\}.$$



Demostración. Sean $H > 0$, $\varepsilon > 0$ dado y $1 < M < N$, como la serie converge en $s = s_0$, entonces

$$\sum_{M<n\leqslant N} f(n)n^{-s_0}n^{s_0-s} = \int_{M^-}^{N} t^{s_0-s}dA(t) = \int_{M^-}^{N} t^{s_0-s}d(A(t) - A(M)),$$

con $A(t) = \sum_{n\leqslant t} f(n)n^{-s_0}$, luego

$$\begin{aligned}\sum_{M<n\leqslant N} f(n)n^{-s} &= \frac{A(N) - A(M)}{N^{s-s_0}} - \int_{M}^{N} (A(t) - A(M))d(t^{s_0-s}) \\ &= \frac{A(N) - A(M)}{N^{s-s_0}} - (s_0 - s)\int_{M}^{N} (A(t) - A(M))t^{s_0-s-1}dt.\end{aligned}$$

Tenemos que $A(t)$ converge cuando $t \to \infty$, luego $A(N) - A(M)$ y $A(t) - A(M)$ convergen a 0 cuando $M, N \to \infty$. Así, para M, N suficientemente grandes y $\sigma > \sigma_0$

$$|A(N) - A(M)| < \frac{\varepsilon}{(H+2)}, \quad |A(t) - A(M)| < \frac{\varepsilon}{H+2},$$

de esto se sigue que:

$$\begin{aligned}\left|\sum_{M<n\leqslant N}\frac{f(n)}{n^s}\right| &< \frac{\varepsilon}{H+2}+\frac{\varepsilon}{H+2}\left|s-s_0\right|\int_M^\infty t^{\sigma_0-\sigma-1}dt\\ &= \frac{\varepsilon}{H+2}+\frac{\varepsilon\left|s-s_0\right|}{(H+2)(\sigma-\sigma_0)M^{\sigma-\sigma_0}}\\ &\leqslant \left(1+\frac{|s-s_0|}{\sigma-\sigma_0}\right)\frac{\varepsilon}{H+2}.\end{aligned}$$

Queremos ver que $\left|\sum_{M<n\leqslant N} f(n)n^{-s}\right| < \varepsilon$, así el teorema se sigue de el teorema de Cauchy para convergencia uniforme de series, en efecto

$$|s-s_0| = |(\sigma-\sigma_0)+i(t-t_0)| \leqslant \sigma-\sigma_0+|t-t_0| \leqslant \sigma-\sigma_0+H(\sigma-\sigma_0) \leqslant (H+1)(\sigma-\sigma_0),$$

así:

$$\left|\sum_{M<n\leqslant N}\frac{f(n)}{n^s}\right| < (1+(H+1))\frac{\varepsilon}{H+2} = \varepsilon.$$

□

Note que haciendo H suficientemente grande obtenemos que $F(s)$ converge para todo s en el semiplano $\sigma > \sigma_0$, geométricamente lo que hacemos es abrir el cono de la región S para obtener este semiplano.[7]

Teorema 1.28. Supongamos que $F(s) = \sum_{n=1}^{\infty} f(n)n^{-s}$ no converge para todo $s \in \mathbb{C}$ ni diverge para todo $s \in \mathbb{C}$, entonces existe $\sigma_c \in \mathbb{R}$ tal que $F(s)$ converge para todo s con $\sigma > \sigma_0$ y diverge para todo $\sigma < \sigma_0$.

Demostración. Como mencionamos antes, tomando H suficientemente grande vemos que si $F(s)$ converge en $s = s_0$, entonces por el teorema anterior $F(s)$ converge para todo $\sigma > \sigma_0$, por tanto, de manera análoga a la prueba del teorema 1.25, consideremos el conjunto:

$$A = \left\{ a \in \mathbb{R} : \sum_{n=1}^{\infty} f(n)n^{-a} \text{ converge} \right\}.$$

Nuevamente consideremos σ_c el mínimo de A, si s tiene parte real $\sigma > \sigma_c$, entonces $\sigma > a$ para algún $a \in A$, luego $F(s)$ converge, si s tiene parte real $\sigma < \sigma_c$, entonces $\sigma < \frac{\sigma+\sigma_c}{2} < \sigma_c$ y como σ_c es el mínimo, $\frac{\sigma+\sigma_c}{2} \notin A$, lo que implica que $F(s)$ diverge. □

Ejemplo. Supongamos que la suma

$$\sum_{n\leqslant x} f(n) \text{ está acotada para todo } x \in [1,\infty).$$

Tenemos que n^{-s} para $\sigma > 0$ es decreciente, luego dado $\sigma_0 > 0$, tenemos que $\sum_{n=1}^{\infty} f(n)n^{-\sigma_0}$ converge por el criterio de Dirichlet y por el teorema anterior $F(s)$ converge para todo S con $\sigma > \sigma_0$, haciendo σ_0 arbitrariamente pequeño, obtenemos que $F(s)$ converge en el semiplano $\sigma > 0$.

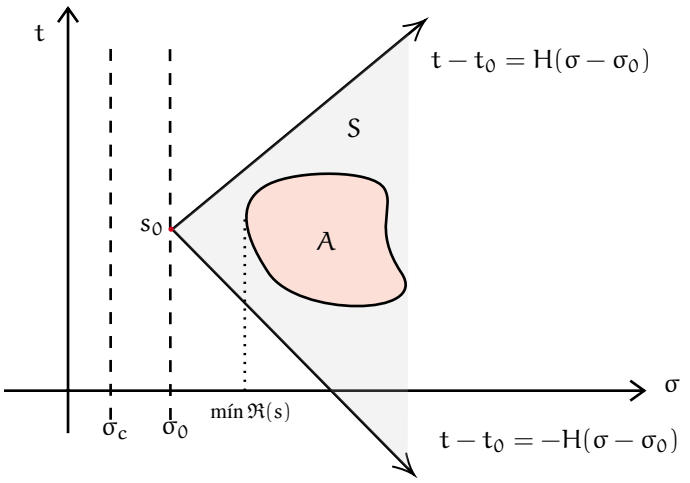
El siguiente resultado nos da una propiedad que esperábamos, podemos derivar término a término:

Corolario 1.29. Sea $F(s) = \sum_{n=1}^{\infty} f(n)n^{-s}$ para todo $s \in \mathbb{C}$ con $\sigma > \sigma_c$, entonces $F(s)$ es analítica en el semiplano $\sigma > \sigma_c$ y $F'(s)$ se puede obtener derivando término a término, es decir:

$$F'(s) = -\sum_{n=1}^{\infty} \frac{f(n)\log n}{n^s} \quad \text{para todo } s \text{ con } \sigma > \sigma_c.$$

Demostración. Sea A un compacto en el semiplano $\sigma > \sigma_c$, tomemos s_0 con parte real σ_0 tal que $\sigma_c < \sigma_0 < \min\{\Re(s) : s \in A\}$, como $F(s)$ converge en s_0, entonces por el Teorema 1.27, $F(s)$ converge uniformemente en la región $S = \{s \in \mathbb{C} : \sigma > \sigma_0 \text{ y } |t-t_0| \leqslant H(\sigma-\sigma_0)\}$. Sabemos que A es acotado porque es compacto, luego existe un $H > 0$ tal que $A \subset S$, basta tomar un H suficientemente grande. De esto se sigue que $F(s)$ converge uniformemente en cualquier compacto A del semiplano $\sigma > \sigma_c$, por el Teorema 1.26, $F(s)$ es analítica en el semiplano $\sigma > \sigma_0$, $F'(s)$ se puede obtener derivando término a término:

$$F'(s) = -\sum_{n=1}^{\infty} f(n)(\log n)n^{-s}.$$



□

Nota. Por lo anterior, $F'(s)$ tiene la misma abscisa de convergencia que $F(s)$, note que el corolario anterior también se tiene si consideramos s en el semiplano de convergencia absoluta, la prueba es exactamente la misma, resaltamos esto porque además nos dice que $F'(s)$ tiene la misma abscisa de convergencia absoluta que $F(s)$, esto nos será útil después.[8]

Corolario 1.30. La función zeta de Riemann es analítica en el semiplano $\sigma > 1$. Luego

$$\zeta'(s) = -\sum_{n=1}^{\infty} \frac{\log n}{n^s} \text{ para todo } s \in \mathbb{C} \text{ con } \sigma > 1.$$

Recordemos que el TFA en términos de convolución nos dice que $\log n = \Lambda * 1$, por tanto

$$\zeta'(s) = -\sum_{n=1}^{\infty} \frac{\log n}{n^s} = -\sum_{n=1}^{\infty} \frac{\Lambda * 1(n)}{n^s} = -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \sum_{k=1}^{\infty} \frac{1}{k^s}.$$

El paso anterior, en el que separamos la convolución como un producto depende de que ambas series sean absolutamente convergentes en el punto s. No tenemos de momento una abscisa de convergencia absoluta para la serie de Dirichlet asociada a la función de Von Mangolth, Afortunadamente, no es difícil ver que también $\sigma_a = 1$. Note que para todo $n \in \mathbb{N}$, $|\Lambda(n)| \leqslant \log n$ y como la serie de Dirichlet de $\log(n)$ es $-\zeta'(s)$ que tiene abscisa de convergencia absoluta $\sigma_a = 1$, entonces

$$\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \quad \text{converge absolutamente si } \sigma > 1.$$

El paso que hicimos es válido en el semiplano $\sigma > 1$, por tanto:

$$\frac{\zeta'(s)}{\zeta(s)} = -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}$$

Esta es la derivada logarítmica de $\zeta(s)$, una serie de Dirichlet que será muy importante en la prueba del teorema de los números primos. El siguiente teorema nos da una relación importante entre σ_c y σ_a.

Teorema 1.31. Sea $F(s)$ una serie de Dirichlet, entonces $\sigma_c \leqslant \sigma_a \leqslant \sigma_c + 1$

Demostración. Es claro que si $F(s)$ es absolutamente convergente, entonces $\sigma_a \geqslant \sigma_c$, veamos la otra desigualdad. Note que la serie $\sum_{n=1}^{\infty} f(n)n^{-\sigma_c-\varepsilon}$ converge para todo $\varepsilon > 0$, por tanto $f(n)n^{-\sigma_c-\varepsilon}$ converge a 0 cuando $n \to \infty$. Es decir, existe un $N > 0$ tal que $|f(n)| < n^{\sigma_c+\varepsilon}$ para todo $n \geqslant N$, luego

$$\left| \frac{f(n)}{n^{\sigma+1+2\varepsilon}} \right| < \frac{1}{n^{1+\varepsilon}}.$$

Tenemos que la serie $\sum_{n=1}^{\infty} n^{-1-\varepsilon}$ converge, luego por criterio de comparación, la serie

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^{\sigma_c+2\varepsilon+1}} \quad \text{converge absolutamente,}$$

esto es, $\sigma_a \leqslant \sigma_c + 1 + 2\varepsilon$, para todo $\varepsilon > 0$, luego obtenemos $\sigma_a \leqslant \sigma_c + 1$. □

Teorema 1.32 (Unicidad)**.** Sean $F(s)$ y $G(s)$ las series de Dirichlet asociadas a f y g respectivamente, si $F(s)$ y $G(s)$ tienen abscisa de convergencia finita y además $F(s) = G(s)$ para todo s con σ suficientemente grande. Entonces $f(n) = g(n)$.

Demostración. Sea $h(n) = f(n) - g(n)$ y $H(s) = F(s) - G(s)$ su serie de Dirichlet, por hipótesis, existe un σ_0 tal que $H(s)$ converge en el semiplano $\sigma > \sigma_0$ y $H(s)$ es idénticamente nula en este semiplano, queremos ver que $h(n) = n$ para todo $n \in \mathbb{N}$. Supongamos que $h(n)$ no es idénticamente nula y sea N el mínimo entero positivo tal que $h(n) \neq 0$, entonces

$$H(s) = \frac{h(N)}{N^s} + \sum_{n=N+1}^{\infty} \frac{h(n)}{n^s},$$

$$h(N) = N^s H(s) - N^s \sum_{n=N+1}^{\infty} h(n) n^{-s}.$$

De esto se sigue que para todo $\sigma > \sigma_0$, $h(N) = -N^\sigma \sum_{n=N+1}^{\infty} h(n) n^{-\sigma}$. y por tanto

$$|h(N)| \leqslant \sum_{n=N+1}^{\infty} |h(n)| \left(\frac{N}{n}\right)^{\sigma},$$

tomando $\sigma = \sigma_0 + \lambda$ con $\lambda \geqslant 0$, tenemos que para todo $n \geqslant N+1$

$$\left(\frac{N}{n}\right)^{\sigma} = \left(\frac{N}{n}\right)^{\lambda} \left(\frac{N}{n}\right)^{\sigma_0} \leqslant \left(\frac{N}{N+1}\right)^{\lambda} \left(\frac{N}{n}\right)^{\sigma_0},$$

luego:

$$\begin{aligned} |h(N)| &\leqslant \left(\frac{N}{N+1}\right)^{\lambda} \sum_{n=N+1}^{\infty} |h(n)| \left(\frac{N}{n}\right)^{\sigma_0} = N^{\sigma_0} \left(\frac{N}{N+1}\right)^{\lambda} \sum_{n=N+1}^{\infty} |h(n)| n^{-\sigma_0} \\ &= C \left(\frac{N}{N+1}\right)^{\lambda}, \end{aligned}$$

donde C es una constante que no depende de λ dado que la serie $\sum_{n=N+1}^{\infty} |h(n)| n^{-\sigma_0}$ converge.

Note que

$$\lim_{\lambda\to\infty} C\left(\frac{N}{N+1}\right)^{\lambda} = 0.$$

Luego $h(N) = 0$, contradicción. $f(n) = g(n)$ para todo $n \in \mathbb{N}$.

□

Con estas propiedades nos basta para lo que queremos mostrar en este trabajo, sin embargo hay muchas otras propiedades analíticas de las series de Dirichlet que son útiles en otros contextos y además muy interesantes, algunas fuentes en las que se pueden consultar son [8] y [7].

1.5.3. El producto de Euler

Ya habíamos mencionado antes que las series de Dirichlet son particularmente útiles cuando f es una función multiplicativa, ya que en este caso vemos a ver que

$$F(s) = \prod_p \left(1 + \frac{f(p)}{p^s} + \frac{f\left(p^2\right)}{p^{2s}} + + \frac{f\left(p^3\right)}{p^{3s}} + \cdots\right),$$

siempre que las series sobre los primos y el producto converjan absolutamente. Adicionalmente, si f es completamente multiplicativa

$$F(s) = \prod_p (1 - f(p))^{-1}.$$

Este resultado es llamado producto de Euler, curiosamente no porque Euler lo haya demostrado, sino porque el primer resultado de este estilo fue dado por Euler, para la función $\zeta(s)$, $s > 1$. Como sabemos, este resultado es muy importante para poder probar la divergencia de la serie de los recíprocos de los números primos, por lo que requerimos la versión general para poder presentar una prueba del teorema de Dirichlet en nuestro siguiente capítulo.

Para este punto debemos asumir algunos resultados de convergencia de productos infinitos, resultados que son bien conocidos, sin embargo, vamos a recordarlos por completitud. Dicho esto no presentaremos pruebas, estos resultados se estudian bien a detalle en [4], allí se encuentran pruebas detalladas y el lector interesado puede consultarlas.

Definición. Decimos que el producto $\prod_{n=1}^{\infty} a_n$ converge si:

- Existe $k \in \mathbb{N}$ tal que $a_n \neq 0$ para todo $n \geqslant k$ y
- $\lim_{m\to\infty} \prod_{n=k}^{m} a_n$ existe y es no nulo.

De la definición anterior tenemos que un producto $\prod_{n=1}^{\infty} a_n$ converge a 0 si y solo si existe

un n tal que $a_n = 0$ y hay solo una cantidad finita de a_n de esta forma, así tomando $N = \text{máx}\{n : a_n = 0\}$ el $\lim_{m\to\infty} \prod_{k=N+1}^{m} a_N$ existe y es no nulo. Por ejemplo, el producto

$$(0)(1)(1)(1)(1)\cdots \quad \text{converge a 0,}$$

pero los productos $(0)(1)(0)(1)(0)(1)\cdots$ y $(0)(1)(2)(3)(4)\cdots$ no convergen [5]. De manera análoga a lo que ocurre con series, si el producto $\prod_{n=1}^{\infty} a_n$ converge, entonces a_n converge a 1. Nos interesa caracterizar la convergencia de los productos de la forma $\prod_{n=1}^{\infty}(1 + a_n)$, note que por lo anterior a_n converge a 0.

Definición. Sea a_n una sucesión de números complejos. El producto $\prod_{n=1}^{\infty}(1 + a_n)$ es absolutamente convergente si $\prod_{n=1}^{\infty}(1 + |a_n|)$ es convergente.

En productos infinitos también la convergencia absoluta implica convergencia.

Teorema 1.33. Si la serie $\sum_{n=1}^{\infty} |a_n|$ converge, entonces el producto $\prod_{n=1}^{\infty}(1 + |a_n|)$ converge.

Como mencionamos, no se presentará una prueba de este resultado, el lector puede consultarla en [4].

Teorema 1.34. Sea f una función aritmética multiplicativa si $\sum_{n=1}^{\infty} f(n)$ converge absolutamente, entonces

$$\sum_{n=1}^{\infty} f(n) = \prod_{p} \left(1 + f(p) + f\left(p^2\right) + \cdots\right).$$

Si f es completamente multiplicativa, entonces

$$\sum_{n=1}^{\infty} f(n) = \prod_{p} (1 - f(p))^{-1}.$$

Demostración. Note que:

$$\sum_{p} \left|f(p) + f(p^2) + \ldots\right| \leqslant \sum_{p} |f(p)| + \sum_{p} \left|f(p^2)\right| + \ldots \leqslant \sum_{n=2}^{\infty} |f(n)|. \qquad (1.13)$$

La serie en la derecha de (1.13) converge, entonces $\prod_p \left(1 + f(p) + f\left(p^2\right) + \cdots\right)$ converge absolutamente. Como f es multiplicativa tenemos que

$$\begin{aligned} P(x) &= \prod_{p\leqslant x}\left(1 + f(p) + f(p^2) + \ldots\right) = \prod_{p\leqslant x}\left(1 + \sum_{k=1}^{\infty} f(p^k)\right) \\ &= 1 + \sum_{k_1=1}^{\infty} f(p_1^{k_1}) + \sum_{k_2=1}^{\infty} f(p_2^{k_2}) + \sum_{k_1=1}^{\infty} f(p_1^{k_1}) \sum_{k_2=1}^{\infty} f(p_2^{k_2}) + \ldots \\ &= \sum_{k_1=0}^{\infty} \ldots \sum_{k_t=0}^{\infty} (f(p_1^{k_1}) \ldots f(p_t^{k_t})) \\ &= \sum_{k_1=0}^{\infty} \ldots \sum_{k_t=0}^{\infty} (f(p_1^{k_1} \ldots p_t^{k_t})). \end{aligned} \tag{1.14}$$

En efecto podemos obtener el término 1 tomando $k_1 = k_2 = \ldots = k_t = 0$ ya que $f(1) = 1$ porque f es multiplicativa. Note ahora que dado el conjunto A de los enteros positivos con todos sus factores primos menores o iguales a x, el producto se puede escribir de manera compacta como

$$P(x) = \sum_{n\in A} f(n),$$

luego, dado B el conjunto de los enteros positivos con al menos un factor primo mayor que x

$$\left|\sum_{n=1}^{\infty} f(n) - P(x)\right| \leqslant \sum_{n\in B} |f(n)| \leqslant \sum_{n>x} |f(n)| = \sum_{n=1}^{\infty} |f(n)| - \sum_{n\leqslant x} |f(n)|.$$

Lo que converge a 0 cuando $x \to \infty$, luego $P(x)$ converge a $\sum_{n=1}^{\infty} f(n)$ cuando $x \to \infty$. Si f es completamente multiplicativa, $f(p^k) = f(p)^k$, aplicando esto a (1.14) obtenemos lo deseado:

$$P(x) = \prod_p \left(1 + f(p) + (f(p))^2 + (f(p))^3 + \cdots\right) = \prod (1 - f(p))^{-1},$$

ya que las series $\sum_{k=0}^{\infty} f(p_t^{k_t})$ en (1.14) se vuelven geométricas, cada una convergiendo a $(1 - f(p_t))^{-1}$, $|f(p_t)| < 1$, luego

$$\sum_{n=1}^{\infty} f(n) = \prod_{n=1}^{\infty} (1 - f(p))^{-1}.$$

□

Corolario 1.35 (Producto de Euler)**.** Sea f una función aritmética y $F(s)$ su serie de Dirichlet, supongamos que $F(s)$ converge absolutamente para $\sigma > \sigma_a$. Si f es multiplicativa

$$\sum_{n=1}^{\infty} f(n)n^{-s} = \prod_{p} \left(1 + f(p)p^{-s} + f\left(p^2\right)p^{-2s} + \cdots\right).$$

Si f es completamente multiplicativa

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

y cada producto es absolutamente convergente para $\sigma > \sigma_a$.

La prueba es esencialmente trivial, basta considerar

$$F(s) = \sum_{n=1}^{\infty} f_s(n) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s},$$

y aplicar el teorema anterior a f_s, note que si f es multiplicativa o completamente multiplicativa, también f_s.

Corolario 1.36. Sea $s \in \mathbb{C}$ tal que $\sigma > 1$, entonces

$$\zeta(s) = \prod_{p} \left(1 - \frac{1}{p^s}\right)^{-1}.$$

Observemos lo siguiente

$$\begin{aligned}
\frac{1}{\zeta(s)} &= \prod_{p} \left(1 - \frac{1}{p^s}\right) \\
&= \left(1 - \frac{1}{2^s}\right)\left(1 - \frac{1}{3^s}\right)\left(1 - \frac{1}{5^s}\right)\cdots \\
&= 1 - \sum_{p} p^{-s} + \sum_{\substack{p,q \\ p\neq q}} p^{-s}q^{-s} - \ldots \\
&= \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}.
\end{aligned}$$

Hemos probado unicidad en series de Dirichlet, esto nos permite obtener una prueba alternativa de $\mu * 1 = e$. En efecto

$$\sum_{n=1}^{\infty} \frac{e(n)}{n^s} = 1 = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}\zeta(s) = \sum_{n=1}^{\infty} \frac{\mu * 1(n)}{n^s}.$$

Otro ejemplo de esto es:

$$\begin{aligned}
\frac{\zeta(s-1)}{\zeta(s)} &= \prod_p \frac{(1-p^{-s})}{(1-p^{-s+1})} \\
&= \prod_p \left(1-\frac{1}{p^s}\right)\left(\sum_{k=0}^{\infty}\frac{p^k}{p^{ks}}\right) \\
&= \prod_p \left(1-\frac{1}{p^s}\right)\left[1+\frac{p}{p^s}+\frac{p^2}{p^{2s}}+\ldots\right] \\
&= \prod_p \left(\left[1+\frac{p}{p^s}+\frac{p^2}{p^{2s}}+\ldots\right]-\left[\frac{1}{p^s}+\frac{p}{p^{2s}}+\frac{p^2}{p^{3s}}+\ldots\right]\right) \\
&= \prod_p \left[1+\left(1-\frac{1}{p}\right)\left(\frac{p}{p^s}+\frac{p^2}{p^{2s}}+\ldots\right)\right] \\
&= \prod_p \left[1+\frac{p-1}{p^s}+\frac{p^2-p}{p^{2s}}+\ldots\right] \\
&= \sum_{n=1}^{\infty} f(n)n^{-s},
\end{aligned}$$

donde $f(n)$ es tal que $f(p^k) = p^k - p^{k-1}$, luego, $f(n) = \varphi(n)$, ahora note que

$$\zeta(s-1) = \sum_{n=1}^{\infty}\frac{1}{n^{s-1}} = \sum_{n=1}^{\infty}\frac{n}{n^s},$$

así:

$$\frac{\zeta(s-1)}{\zeta(s)} = \sum_{n=1}^{\infty}\frac{\varphi(n)}{n^s} = \sum_{n=1}^{\infty}\frac{n}{n^s}\sum_{n=1}^{\infty}\frac{\mu(n)}{n^s} = \sum_{n=1}^{\infty}\frac{\mu * N(n)}{n^s},$$

por el teorema de unicidad, obtenemos que $\mu * N = \varphi$. Esto muestra el poder teórico de las series de Dirichlet, cuando tenemos propiedades adecuadas, podemos obtener mucha información de las funciones aritmética asociadas. Pero esto no es todo, mostraré una propiedad adicional, toda serie de Dirichlet se puede ver como una integral de Riemann-Stieltjes y bajo ciertas condiciones esta integral tomará la forma de una transformada de Laplace. Este hecho tomará particular importancia en la prueba del TNP.

Lema 1.37 (Kronecker)**.** Sean f una función aritmética y $s \in \mathbb{C}$ con $\sigma > 0$ tal que $F(s)$ converge, entonces

$$\lim_{x\to\infty}\frac{1}{x^s}\sum_{n\leqslant x} f(n) = 0.$$

Demostración. Sea f fija pero arbitraria y $s \in \mathbb{C}$ con $\sigma > 0$. Consideremos

$$S(x) = \sum_{n\leqslant x} f(n), \quad D(x) = \sum_{n\leqslant x}\frac{f(n)}{n^s},$$

por hipótesis $D(x)$ converge a un D cuando $x \to \infty$, queremos ver que $\lim_{x\to\infty} \dfrac{S(x)}{x^s} = 0$. Dado $\varepsilon > 0$, existe un $x_0 \geqslant 1$, tal que para todo $x \geqslant x_0$

$$|D(x) - D| < \frac{\varepsilon}{2}.$$

Sea $x \geqslant x_0$, aplicando sumación parcial a $S(x)$ tenemos que

$$\begin{aligned} S(x) = \sum_{n\leqslant x} \frac{f(n)}{n^s} n^s &= \int_{1^-}^{x} t^s d(D(t)) = D(x)x^s - \int_1^x D(t) s t^{s-1} dt \\ &= \int_0^x D(x) s t^{s-1} dt - \int_1^x D(t) s t^{s-1} dt. \end{aligned}$$

Note que $D(t) = 0$ para todo $t \in [0, 1)$, luego en la segunda integral podemos cambiar el límite inferior por 0, obteniendo que

$$\begin{aligned} |S(x)| &= \left| \int_0^x (D(x) - D(t)) s t^{s-1} dt \right| \\ &\leqslant \int_0^x |D(x) - D(t)|\, |s|\, t^{\sigma-1} dt, \end{aligned}$$

para $x_0 \leqslant t \leqslant x$ tenemos que:

$$|D(t) - D(x)| \leqslant |D(t) - D| + |D - D(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

y si $0 \leqslant t \leqslant x_0$, entonces

$$\begin{aligned} |D(t) - D(x)| &\leqslant |D(t)| + |D| + |D - D(x)| \\ &< \sum_{n\leqslant x_0} \left| \frac{f(n)}{n^s} \right| + |D| + \frac{\varepsilon}{2} = M, \end{aligned}$$

donde M depende de ε, pero no de x. De esto se sigue que

$$\begin{aligned} |S(x)| &\leqslant \varepsilon \int_{x_0}^{x} |s| t^{\sigma-1} dt + M \int_0^{x_0} |s| t^{\sigma-1} dt \\ &\leqslant \frac{|s|}{\sigma} \left(\varepsilon \left(x^\sigma - x_0^\sigma \right) + M x_0^\sigma \right), \end{aligned}$$

así

$$\left| \frac{S(x)}{x^s} \right| \leqslant \frac{|s|}{\sigma} \left(\varepsilon + \frac{M x_0^\sigma}{x^\sigma} \right). \tag{1.15}$$

Como M no depende de x, entonces tomando $x \to \infty$, el último término en (1.15) tiende a 0, luego para todo $\varepsilon > 0$

$$\lim_{x\to\infty} \left| \frac{S(x)}{x^s} \right| \leqslant \frac{\varepsilon |s|}{\sigma},$$

y como ε es arbitrario, $\lim_{x\to\infty} \frac{S(x)}{x^s} = 0$.

□

Sea $F(s)$ una serie de Dirichlet con abscisa de convergencia $\sigma_c > 0$, para todo $s \in \mathbb{C}$ con $\sigma > \sigma_c$

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{f(n)}{n^s} &= \lim_{N\to\infty} \int_{1^-}^{N} x^{-s} d(S(x)) \\ &= \lim_{N\to\infty} \frac{S(N)}{N^s} - \lim_{N\to\infty} \int_1^N S(x) d(x^{-s}) \\ &= s\int_1^{\infty} \frac{S(x)}{x^{s+1}} dx.\end{aligned}$$

El término de borde se anula por el lema de Kronecker. Así obtenemos una representación integral para cualquier serie de Dirichlet con abscisa de convergencia positiva. Aplicando esto a $\zeta(s)$ obtenemos una extensión analítica de la función zeta de Riemann al semiplano $\sigma > 0$ con un polo simple en $s = 1$ con residuo 1.

Note que tomando $x = e^t$, $dx = e^t dt$

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = s\int_0^{\infty} S(e^t) e^{-st} dt.$$

La serie de Dirichlet toma forma de transformada de Laplace. Aunque esto parezca una propiedad cualquiera, Korevaar y Zagier aplicaron esto para obtener una prueba corta del TNP en la que no es necesario todo el poder del teorema Tauberiano de Wiener-Ikehara [9], esta prueba es una modificación de la prueba de Newman[10], donde el teorema analítico que se emplea se obtiene de manera más sencilla en virtud de lo anterior.

1.6 La función zeta de Riemann

Teorema 1.38 (Propiedades)**.** Sea $s \in \mathbb{C}$ tal que $\sigma > 1$, entonces

i) $\zeta(s) \neq 0$

ii) $\zeta'(s) = -\sum_{n=1}^{\infty} \frac{\log n}{n^s}$

iii) $\frac{\zeta'(s)}{\zeta(s)} = -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}$

Las propiedades (ii) y (iii) ya las hemos probado antes, basta ver (i).

Demostración. Como $\sigma > 1$, entonces

$$\zeta(s) = \prod_p \frac{1}{1-p^{-s}},$$

donde el producto es absolutamente convergente, note que

$$\frac{1}{1-p^{-s}} = \frac{p^s}{p^s-1} = 1 + \frac{1}{p^s-1},$$

entonces el producto no tiene factores nulos, luego no converge a 0, $\zeta(s) \neq 0$. □

Teorema 1.39. Sea $s \in \mathbb{C}$ tal que $\sigma > 1$, entonces

$$\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \frac{\{t\}}{t^{s+1}}dt.$$

Demostración. Tenemos que $\zeta(s) = s\int_1^\infty \frac{[t]}{t^{s+1}}dt$, como $[t] = t - \{t\}$, entonces

$$\begin{aligned}\zeta(s) &= s\int_1^\infty \frac{t-\{t\}}{t^{s+1}} \\ &= s\int_1^\infty \frac{1}{t^s}dt - s\int_1^\infty \frac{\{t\}}{t^{s+1}}dt \\ &= \frac{s}{s-1} - s\int_1^\infty \frac{\{t\}}{t^{s+1}}dt.\end{aligned}$$

□

Teorema 1.40 (Extensión analítica)**.** La representación integral de $\zeta(s)$ nos da una extensión analítica al semiplano $\sigma > 0$ con un polo simple en $s = 1$ con residuo 1.

Demostración. Note que $\frac{s}{s-1} = 1 + \frac{1}{s-1}$ es analítica en el semiplano $\sigma > 0$ excepto por un polo simple en $s = 1$ con residuo 1. Basta ver que la función

$$f(s) = \int_1^\infty \frac{\{t\}}{t^{s+1}}dt,$$

es analítica en el semiplano $\sigma > 0$. Para cada $n \in \mathbb{N}$, sea $f_n(s) = \int_1^n \frac{\{t\}}{t^{s+1}}dt$, tenemos que

$$f_n(s) = \int_1^n \{t\}e^{-(s+1)\log t}dt = \int_1^n \sum_{n=0}^\infty \frac{\{t\}(-\log t)^n(s+1)^n}{n!}dt,$$

donde la suma dentro de la integral es la serie de Taylor de e^x que sabemos tiene radio de convergencia infinito. Queremos ver que f_n es una sucesión de funciones analíticas, para lo que basta ver que podemos introducir la integral en la serie de potencias, además queremos ver que $f_n \to f$ uniformemente en cualquier subconjunto compacto de $\sigma > 0$. En efecto

$$\sum_{m=0}^{\infty}\left|\frac{\{t\}(-\log t)^m(s+1)^m}{m!}\right| \leqslant \sum_{m=0}^{\infty}\frac{(|\log t|(|s|+1))^m}{m!}$$
$$= e^{|\log t|\,(|s|+1)} = t^{|s|+1},$$

entonces

$$\int_1^n \sum_{m=0}^{\infty}\left|\frac{\{t\}(-\log t)^m(s+1)^m}{m!}\right| dt \leqslant \int_1^n t^{|s|+1}dt < \infty,$$

luego por convergencia absoluta

$$f_n(s) = \sum_{m=0}^{\infty}\frac{(s+1)^m}{m!}\int_1^n \{t\}(-\log t)^m dt.$$

Sea $s \in \mathbb{C}$ con $\sigma \geqslant \delta > 0$. Dado $\varepsilon > 0$, sea $N = (\delta\varepsilon)^{-\frac{1}{\delta}}$, entonces para todo $n > N$:

$$|f(s) - f_n(s)| \leqslant \int_n^{\infty}\left|\frac{\{t\}}{t^{s+1}}\right| dt \leqslant \int_n^{\infty}\frac{1}{t^{\sigma+1}} \leqslant \frac{1}{\sigma n^{\sigma}}$$
$$\leqslant \frac{1}{\delta n^{\delta}}$$
$$< \frac{1}{\delta(\delta\varepsilon)^{-\frac{1}{\delta}\delta}} = \varepsilon.$$

como N no depende de s, f_n converge uniformemente a f en el semiplano $\sigma \geqslant \delta$, en particular sobre cualquier compacto de $\sigma > 0$, por el Teorema 1.26 $f(s)$ es analítica en el semiplano $\sigma > 0$. □

1.7 El método de Dirichlet de la hipérbola

En algunas situaciones requerimos estimar expresiones del estilo $\sum_{n\leqslant x}(f*g)(n)$, estas son sumas parciales de una convolución de dos funciones aritmética, por ejemplo, si queremos estimar

$$\sum_{n\leqslant x}\tau(n),$$

tenemos que

$$\sum_{n\leqslant x}\tau(n) = \sum_{d\leqslant x}\sum_{k\leqslant x/d} 1$$
$$= \sum_{d\leqslant x}\left[\frac{x}{d}\right] \qquad ,$$
$$= \sum_{d\leqslant x}\left(\frac{x}{d} + O(1)\right) = x\log x + O(x).$$

sin embargo, dado que $\tau(n) = (1*1)(n)$ podemos obtener un resultado mejor aún

Teorema 1.41 (Dirichlet)**.** Para todo $x \geqslant 2$ se tiene

$$\sum_{n\leqslant x} \tau(n) = x(\log x + 2\gamma - 1) + O(\sqrt{x}).$$

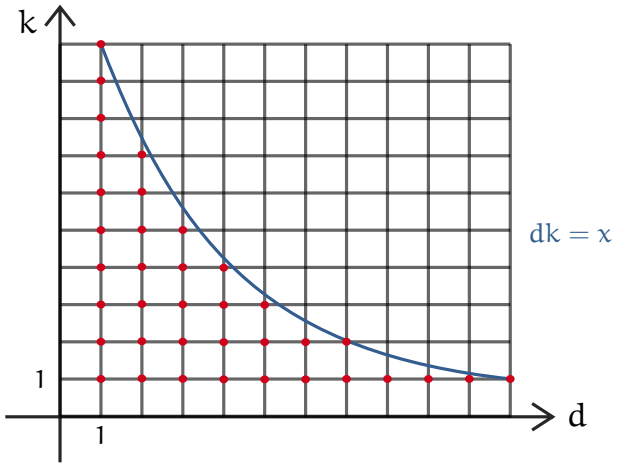
Para lograr esto introducimos el método de Dirichlet de la hipérbola, note que la expresión

$$\sum_{n\leqslant x} (f * g)(n)$$

se puede reescribir como sigue

$$\begin{aligned}\sum_{n\leqslant x} (f * g)(n) &= \sum_{n\leqslant x} \sum_{dk=n} f(d)g(k) \\ &= \sum_{dk\leqslant x} f(d)g(k),\end{aligned}$$

donde la suma sobre $dk \leqslant x$ representa la suma sobre los puntos $(k, d) \in \mathbb{N} \times \mathbb{N}$ que están por debajo de la hipérbola $dk = x$ como en la siguiente figura

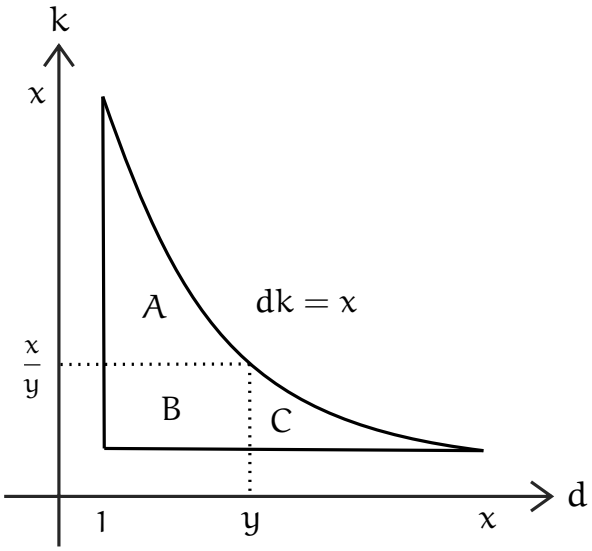


Sumar de manera adecuada nos permite hacer mejores estimaciones, tenemos el siguiente teorema.

Teorema 1.42 (Método de Dirichlet de la hipérbola)**.** Sean f y g funciones aritmética. Dado $y > 0$, tenemos que

$$\sum_{n\leqslant x} (f * g)(n) = \sum_{d\leqslant y} f(d) \sum_{k\leqslant \frac{x}{d}} g(k) + \sum_{k\leqslant \frac{x}{y}} g(k) \sum_{d\leqslant \frac{x}{k}} f(d) - \left(\sum_{k\leqslant \frac{x}{y}} g(k)\right)\left(\sum_{d\leqslant y} f(d)\right)$$

La siguiente figura ilustra la idea



en particular podemos tomar $y = \sqrt{x}$ de lo que obtenemos simetría, el rectángulo de la figura se vuelve un cuadrado de lado $\sqrt{x}$ y esto es conveniente para realizar algunas estimaciones, en particular en este trabajo tomaremos siempre $y = \sqrt{x}$, sin embargo presentaremos la prueba general.

La prueba es inmediada de considerar la suma sobre $A \cup B$, sumarle la suma $B \cup C$ y restar la suma sobre B, sin embargo veremos una más algebraica

Demostración. Podemos de forma directa calcular esta suma como sigue

$$\begin{aligned}\sum_{dk\leqslant x} f(d)g(k) &= \sum_{\substack{dk\leqslant x\\ d\leqslant y}} f(d)g(k) + \sum_{\substack{dk\leqslant x\\ d>y}} f(d)g(k)\\ &= \sum_{d\leqslant y} f(d) \sum_{k\leqslant \frac{x}{d}} g(k) + \sum_{k<\frac{x}{y}} g(k) \sum_{y<d\leqslant \frac{x}{k}} f(d).\end{aligned}$$

Note que a la izquierda ya tenemos el primer término del teorema. Ahora podemos reescribir la última suma de la derecha para obtener lo deseado, en efecto

$$\begin{aligned}\sum_{k<\frac{x}{y}} g(k)\left(\sum_{d\leqslant \frac{x}{k}} f(d) - \sum_{d\leqslant y} f(d)\right) &= \sum_{k<\frac{x}{y}} g(k) \sum_{d\leqslant \frac{x}{k}} f(d) - \left(\sum_{k<\frac{x}{y}} g(k)\right)\left(\sum_{d\leqslant y} f(d)\right)\\ &= \sum_{k\leqslant \frac{x}{y}} g(k) \sum_{d\leqslant \frac{x}{k}} f(d) - \left(\sum_{k\leqslant \frac{x}{y}} g(k)\right)\left(\sum_{d\leqslant y} f(d)\right).\end{aligned}$$

□

1.7.1. El problema de Dirichlet

Sea $\Delta(x)$ el término de error en el teorema de Dirichlet, definido como

$$\Delta(x) = \sum_{n \leqslant x} \tau(n) - x \log x - (2\gamma - 1)x.$$

Por el teorema de Dirichlet, se tiene que $\Delta(x) = O(\sqrt{x})$. La estimación de $\Delta(x)$ constituye el problema de Dirichlet. Dicho problema es de gran interés, no solo por ser un gran desafío sin resolver, sino principalmente debido a que el intento de abordarlo nos lleva a otros problemas profundos interrelacionados con distintas cuestiones en la teoría de números, incluyendo la hipótesis de Riemann. Por lo que cualquier avance significativo podría tener implicaciones extensas, afectando a una amplia gama de otros problemas en el campo. La mayoría de los resultados conocidos son estimaciones de la forma $\Delta(x) = O(x^s)$ para alguna constante s.

El teorema de Dirichlet permite afirmar que se puede elegir $s = 1/2$. Por otro lado, Hardy demostró a inicios del siglo XX que la estimación no es válida para s menor que $1/4$. La conjetura predominante es que $1/4$ podría ser el exponente óptimo para la estimación del error, pero esta hipótesis permanece sin demostración. Hace casi un siglo, G.F. Voronoii estableció que se podría tomar $s = 1/3$, y a pesar de los grandes esfuerzos de numerosos investigadores, no se ha logrado un avance significativo; el récord actual para s se encuentra cerca de $0,3149$.

Dicho esto veamos una prueba del teorema.

Demostración. *Teorema 1.41*. Tenemos que

$$\sum_{n \leqslant x} \tau(n) = \sum_{d \leqslant x} \sum_{k \leqslant x/d} 1,$$

aplicando el método de la hipérbola con $y = \sqrt{x}$, separamos la suma en S_1, S_2 y S_3 como sigue

$$\begin{aligned} S_1 &= \sum_{d \leqslant \sqrt{x}} \sum_{k \leqslant x/d} 1 \\ &= \sum_{d \leqslant \sqrt{x}} \left[\frac{x}{d}\right] = \sum_{d \leqslant \sqrt{x}} \left(\frac{x}{d} + O(1)\right) \\ &= x\left(\log \sqrt{x} + \gamma + O\left(\frac{1}{\sqrt{x}}\right)\right) + O(\sqrt{x}) \\ &= x\left(\frac{1}{2}\log x + \gamma\right) + O(\sqrt{x}), \end{aligned}$$

simétricamente

$$\begin{aligned} S_2 &= \sum_{k\leqslant\sqrt{x}} \sum_{d\leqslant x/k} 1 \\ &= \sum_{k\leqslant\sqrt{x}} \left[\frac{x}{k}\right] = \sum_{k\leqslant\sqrt{x}} \left(\frac{x}{k} + O(1)\right) \\ &= x\left(\log\sqrt{x} + \gamma + O\left(\frac{1}{\sqrt{x}}\right)\right) + O(\sqrt{x}) \\ &= x\left(\frac{1}{2}\log x + \gamma\right) + O(\sqrt{x}) \end{aligned}$$

y por último

$$\begin{aligned} S_3 &= \sum_{d\leqslant\sqrt{x}} \sum_{k\leqslant\sqrt{x}} 1 \\ &= [\sqrt{x}]^2 = (\sqrt{x} - \{\sqrt{x}\})^2 \\ &= x + O(\sqrt{x}). \end{aligned}$$

Sumando $S_1 + S_2 + S_3$ obtenemos lo requerido.

□

El teorema de Dirichlet

> *...Dirichlet creó una parte nueva en las matemáticas, la aplicación de las series infinitas que Fourier ha introducido en la teoría del calor en la exploración de las propiedades de los números primos. Él ha descubierto una variedad de teoremas que ... son los pilares de las nuevas teorías*
>
> **— C. G. J. Jacobi**

El teorema de Dirichlet afirma que dados $a, n \in \mathbb{N}$ tal que $(a, n) = 1$, hay infinitos primos de la forma $a, a + n, a + 2n, a + 3n, \ldots$, el primer resultado sobre la infinitud de los números primos se remonta a Euclides. Supongamos que hay una cantidad finita de primos, podemos contarlos... $p_1, p_2, \ldots, p_n$, note que $p_1 p_2 \ldots p_n + 1$ es primo ya que si $p_i \mid p_1 p_2 \ldots p_n + 1$ para algún $1 \leqslant i \leqslant n$, entonces

$$1 = p_i(K - (p_1 \ldots p_{i-1} p_{i+1} \ldots p_n)),$$

luego $p_i \mid 1$, una contradicción, es decir, siempre podemos construir un primo p_{n+1} con los n primos anteriores, entonces son infinitos.

Intentemos replicar este argumento para probar que hay infinitos primos de la forma $4k + 1$, supongamos que hay finitos primos de la forma $4k + 1$, digamos $p_1, p_2, \ldots, p_n$, debemos construir un nuevo primo de la forma $4k + 1$ para que funcione el argumento de Euclides, sin embargo la expresión $p_1 p_2 \ldots p_n + 1$ no siempre es de la forma $4k + 1$, por ejemplo $5 \times 13 + 1 = 66$, que es congruente a 2 módulo 4, de hecho con esta expresión siempre conseguimos pares. Requerimos una expresión nueva, por ejemplo, podríamos hacer $2p_1 \ldots p_n + 1$, pero también falla, note que $2 \times 5 \times 13 + 1 = 131$ que es un primo de la forma $4k + 3$.

Proposición 2.1. Sea $n \in \mathbb{Z}$, todo divisor primo impar de n^2+1 es de la forma $4k+1$.

Demostración. Suponga que existe $p = 4k+3$ primo tal que $p \mid n^2+1$, entonces $n^2 \equiv -1 \pmod{p}$, luego por el pequeño teorema de Fermat

$$(\text{mód } p) : 1 \equiv n^{p-1} \equiv (n^2)^{2k+1} \equiv (-1)^{2k+1} \equiv -1,$$

así $2 \equiv 0 \pmod{p}$, esto es $p = 2$, contradicción.

□

Con el teorema anterior sabemos que la expresión que buscamos es $N = (2p_1 \ldots p_n)^2 + 1$ ya que de esta forma obtenemos un número cuyos divisores primos son de la forma $4k+1$, basta ver que ningún p_i con $1 \leqslant i \leqslant n$ divide a N.

Supongamos que $p_i \mid N$, luego $p_i(K - 4p_1^2 \dots p_i \dots p_n^2) = 1$ una contradicción, entonces N es un primo de la forma $4k+1$. Continuar replicando este argumento es inviable cuando trabajamos con primos módulo un entero n arbitrario, además no nos sirve para atacar el panorama general, la prueba de este teorema llegaría de una idea totalmente distinta...

Teorema 2.2 (Euler)**.** La serie $\sum_p \frac{1}{p}$ diverge.

Demostración. Por el producto de Euler:

$$\log(\zeta(s)) = \sum_p^{\infty} \left(\sum_{k=1}^{\infty} \frac{1}{k(p)^{ks}} \right) = \sum_p \frac{1}{p^s} + \sum_p \left(\sum_{k=2}^{\infty} \frac{1}{kp^{ks}} \right), \quad \Re(s) > 1$$

note que:

$$\begin{aligned} \sum_p \left(\sum_{k=2}^{\infty} \frac{1}{kp^{ks}} \right) &\leqslant \sum_p \left(\sum_{k=2}^{\infty} \frac{1}{p^{ks}} \right) \\ &\leqslant \sum_p \frac{1}{p^s} \left(\frac{1}{p^s - 1} \right) \\ &\leqslant \sum_p \frac{1}{p^s(p^s-1)} \leqslant \sum_{n=1}^{\infty} \frac{1}{n^s(n^s-1)}. \end{aligned}$$

[1]La última serie converge siempre que $\Re(s) > \frac{1}{2}$, entonces por la divergencia de la serie armónica, tomando el límite cuando $s \to 1^+$

$$\infty = \lim_{s \to 1^+} \log(\zeta(s)) = \lim_{s \to 1^+} \sum_p \frac{1}{p^s} + \sum_p \left(\sum_{k=2}^{\infty} \frac{1}{kp^{ks}} \right) = \sum_p \frac{1}{p} + O(1),$$

lo que nos dice que la suma de los recíprocos de los primos diverge y por lo tanto podemos decir que los primos son infinitos. □

La idea de Dirichlet es replicar este argumento de Euler para probar que hay infinitos primos de la forma $4k + 1$, para ello define la siguiente serie

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s},$$

donde, $\chi(n)$ es la función indicadora:

$$\chi(a) = \begin{cases} 0 & \text{si } a \text{ es par} \\ 1 & \text{si } a \equiv 1 \text{ mód } 4 \\ -1 & \text{si } a \equiv 3 \text{ mód } 4 \end{cases}$$

[1]La convergencia de esta serie nos llegó haciendo cuentas en una clase de estructuras algebraicas, Santiago me dijo “Mateo esa serie es geométrica”.

note que χ es completamente multiplicativa, entonces

$$L(s,\chi) = \prod_p \frac{1}{1-\chi(p)p^{-s}}.$$

Siguiendo los pasos de la prueba anterior obtenemos

$$\begin{aligned}\log L(s,\chi) &= \sum_p \frac{\chi(p)}{p^s} + \sum_p \sum_{k=2}^{\infty} \frac{\chi(k)}{kp^{ks}} \\ &= \sum_{p\equiv 1(4)} \frac{1}{p^s} - \sum_{p\equiv 3(4)} \frac{1}{p^s} + g_1(s,\chi),\end{aligned}$$

donde no es difícil ver que $g_1(s,\chi)$ es convergente cuando $s \to 1^+$ ya que $|\chi(n)| \leqslant 1$, además

$$\log \zeta(s) = \sum_p \frac{1}{p^s} + g(s),$$

así

$$\begin{aligned}\log \zeta(s) + \log L(s,\chi) &= 2\sum_{p\equiv 1(4)} \frac{1}{p^s} + \left(\frac{1}{2^s} + g(s) + g_1(s,\chi)\right), \\ \log \zeta(s) - \log L(s,\chi) &= 2\sum_{p\equiv 3(4)} \frac{1}{p^s} + \left(\frac{1}{2^s} + g(s) - g_1(s,\chi)\right).\end{aligned}$$

Tomando $\lim_{s\to 1^+}$ como antes se verifica que hay infinitos primos de la forma $4k+1$ y $4k+3$ ya que el término restante converge, sin embargo debemos garantizar algo, que $L(1,\chi)$ converge y es distinto de 0. Para esto note que

$$\sum_{n=1}^{\infty} \frac{\chi(n)}{n} = \sum_{n=1}^{\infty} \frac{(-1)^n}{2n+1} = \arctan(1) = \frac{\pi}{4}.$$

Este camino parece fructífero, intentemos replicar esto en un caso general, sea $f(n)$ la función característica de la progresión aritmética, es decir

$$f(n) = \begin{cases} 1, & n \equiv a \pmod m \\ 0, & n \not\equiv a \pmod m \end{cases}$$

en el caso de que $f(n)$ sea completamente multiplicativa tendríamos un producto de Euler

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_p \left(1 - \frac{f(p)}{p^s}\right)^{-1}, \quad \Re(s) > 1$$

y así por argumentos análogos a los de Euler se tendría que

$$\log\left(\sum_{n=1}^{\infty} \frac{f(n)}{n^s}\right) = \sum_{p\equiv a(m)} \frac{1}{p^s} + O(1)$$

Lamentablemente, $f(n)$ generalmente no es multiplicativa

Para resolver este inconveniente estudiaremos los caracteres de un grupo y en particular los caracteres de Dirichlet, estos veremos que poseen propiedades de ortogonalidad, lo que nos permitirá hacer ¡Análisis de Fourier! y representar a la función f característica de la progresión en su serie de Fourier como una combinación lineal finita de funciones completamente multiplicativas (los caracteres), con esto en mente veamos primero unos preliminares sobre caracteres que necesitaremos en la prueba.

2.1 Caracteres y el teorema de Dirichlet

Primero vamos a presentar definición formal de la idea de carácter.

Definición. Sea G un grupo, χ es un carácter de G si $\chi : G \to \mathbb{C}^\times$ y satisface que para todo $a, b \in G$, $\chi(ab) = \chi(a)\chi(b)$, es decir, un homomorfismo de G en $\mathbb{C}^\times$.

El homomorfismo trivial que mapea a todo $g \in G$ al 1 lo llamaremos "carácter trivial" denotado χ_0.

Definición. Dado un grupo G, definimos el conjunto de todos los caracteres de G, denotado como $\widehat{G}$. También definimos la multiplicación en $\widehat{G}$ como

$$\chi_1\chi_2(g) = \chi_1(g)\chi_2(g) \quad \text{para todo } \chi_1, \chi_2 \in \widehat{G} \quad \text{y } g \in G$$

Decimos $e(x)$ para denotar $e^{2\pi i x}$. Por ejemplo, $e(1/n)$ es una raíz n-ésima de la unidad. Una caracterización de los caracteres de un grupo cíclico finito es la siguiente.

Teorema 2.3. Sea G un grupo cíclico de orden n generado por g, $G = \langle g \rangle$, entonces hay exactamente n caracteres $\chi_0, \ldots, \chi_{n-1}$ de G, dados por $\chi_m(g^k) = e((mk)/n)$ para todo $0 \leqslant m \leqslant n-1$ y $k \in \mathbb{Z}$, esto es

$$\chi_0(g^k) = 1, \chi_1(g^k) = e(k/n), \chi_2(g^k) = e(2k/n), \ldots, \chi_{n-1}(g^k) = e((n-1)k/n).$$

Demostración. Sea χ un carácter de G, tenemos que $\chi(g)^n = \chi(g^n) = \chi(1) = 1$, esto es que $\chi(g)$ es una raíz $n-$ésima de la unidad, luego $\chi(g) = e(m/n)$ para algún $0 \leqslant m \leqslant n-1$, dado que G es cíclico y generado por g, entonces χ está totalmente determinado por $\chi(g)$ y $\chi(g)^k = \chi(g^k) = e((km)/n)$ para todo $k \in \mathbb{Z}$, esto es $\chi = \chi_m$, Así si χ es un carácter de G, χ debe ser uno de los $\chi_0, \ldots, \chi_{n-1}$.Ahora veamos que $\chi_0, \ldots, \chi_{n-1}$ están bien definidos y son caracteres distintos de G.

En efecto si $g^{k_1} = g^{k_2}$ entonces $k_1 \equiv k_2 \pmod{n}$, luego $\chi_m(g^{k_1}) = e((mk_1)/n) = e((mk_2)/n) = \chi_m(g^{k_2})$ para todo $0 \leqslant m \leqslant n-1$, por tanto $\chi_0, \ldots, \chi_{n-1}$ están bien

definidos en G. Dados $0 \leqslant m_1, m_2 \leqslant n-1$ con $m_1 \neq m_2$, entonces

$$\chi_{m_1}(g) = e\left(\frac{m_1}{n}\right) \neq e\left(\frac{m_2}{n}\right) = \chi_{m_2}(g),$$

como χ_{m_1} y χ_{m_2} están totalmente determinados por $\chi_{m_1}(g)$ y $\chi_{m_2}(g)$, entonces $\chi_0, \ldots \chi_{n-1}$ son todos distintos.

Finalmente debemos ver que dado χ_m con $0 \leqslant m \leqslant n-1$, χ_m es homomorfismo de G en $\mathbb{C}^\times$, en efecto dados $a, b \in G$, $a = g^{k_1}$ y $b = g^{k_2}$. Note que

$$\begin{aligned}
\chi_m(ab) &= \chi_m(g^{k_1} g^{k_2}) \\
&= \chi_m(g^{k_1+g_2}) \\
&= e\left(\frac{m(k_1+k_2)}{n}\right) \\
&= e\left(\frac{mk_1}{n}\right) e\left(\frac{mk_2}{n}\right) \\
&= \chi_m(a)\chi_m(b).
\end{aligned}$$

Esto es, χ_m es homomorfismo. □

Teorema 2.4. Sea G un grupo, el conjunto $\widehat{G}$ es un grupo abeliano bajo la multiplicación.

Demostración. El carácter principal χ_0 de G es la identidad de $\widehat{G}$ ya que dado $a \in g$ $\chi_m\chi_0(a) = \chi_m(a)\chi_0(a) = \chi_m(a)$ con $0 \leqslant m \leqslant n-1$, además note que la función $\chi^{-1} = 1/\chi(g)$ es también un carácter y por lo tanto tenemos inversos en $\widehat{G}$, Además $\widehat{G}$ es cerrado bajo la multiplicación, note que:

$$\begin{aligned}
\chi_{m_1}\chi_{m_2}(ab) &= \chi_{m_1}(ab)\chi_{m_2}(ab) \\
&= \chi_{m_1}(a)\chi_{m_1}(b)\chi_{m_2}(a)\chi_{m_2}(b) \\
&= \chi_{m_1}\chi_{m_2}(a)\chi_{m_1}\chi_{m_2}(b).
\end{aligned}$$

La conmutatividad y asociatividad se siguen de manera análoga usando la conmutatividad y asociatividad de $\mathbb{C}^\times$.

□

Nota. Por el teorema anterior, dado que $\widehat{G}$ es grupo, lo llamaremos grupo dual de G.

La inversa de χ en la prueba del anterior, denotada como χ^{-1} en algunos casos se escribe también como $\overline{\chi}$, la razón de esto es que si G es un grupo abeliano finito entonces $|\chi(g)| = 1$, así $\chi^{-1}(g) = 1/\chi(g) = \overline{\chi(g)}$ por propiedad de la norma en los complejos. Así la función $\overline{\chi}$ definida como $\overline{\chi}(g) = \overline{\chi(g)}$ y la función χ^{-1} son iguales. [5]

Teorema 2.5. Sea $G = \langle g \rangle$ un grupo cíclico de orden n, entonces:

(i) Dado un carácter χ de G

$$\sum_{g \in G} \chi(g) = \begin{cases} n, & \text{si } \chi = \chi_0 \\ 0, & \text{e.o.c} \end{cases}$$

(ii) Dado $a \in G$

$$\sum_{\chi \in \widehat{G}} \chi(a) = \begin{cases} n, & \text{si } a = 1 \\ 0, & \text{e.o.c} \end{cases}$$

(iii) $\widehat{G}$ es un grupo cíclico generado por χ_1.

Demostración. (i) Note que dado un carácter χ no trivial de G, existe un $0 \leqslant m \leqslant n-1$ tal que $\chi = \chi_m$. Si $m \neq 0$

$$\sum_{g \in G} \chi(g) = \sum_{k=0}^{n-1} \chi_m\left(g_0^k\right) = \sum_{k=0}^{n-1} e\left(\frac{mk}{n}\right) = \sum_{k=0}^{n-1} \left(e\left(\frac{m}{n}\right)\right)^k = \frac{1 - e(m)}{1 - e\left(\frac{m}{n}\right)},$$

y $1 - e(m) = 1 - \cos(2\pi m) = 0$ para todo $m = 1, \ldots, n-1$, si $m = 0$ entonces

$$\sum_{k=0}^{n-1} \chi_m\left(g_0^k\right) = \sum_{k=0}^{n-1} 1 = n.$$

(ii) Sea $a \in G$, $a = g^k$ para algún $k = 0, 1, \ldots, n-1$, en efecto

$$\sum_{\chi \in \widehat{G}} \chi(a) = \sum_{m=0}^{n-1} \chi_m(a) = \sum_{m=0}^{n-1} \chi_m\left(g^k\right) = \sum_{m=0}^{n-1} e\left(\frac{mk}{n}\right),$$

análogamente se ve que esta suma es n si $k = 0$ y es 0 si $k \neq 0$, y el valor $k = 0$ corresponde a $g^0 = 1$. Para (iii) note que por el teorema 2.3

$$\widehat{G} = \{\chi_0, \chi_1, \ldots, \chi_{n-1}\} = \left\{\chi_1^0, \chi_1, \chi_1^2, \ldots, \chi_1^{n-1}\right\} = \langle \chi_1 \rangle .$$

Esto completa la prueba.

□

Observe que el teorema anterior nos permite establecer lo siguiente, dado G un grupo cíclico de orden n, entonces G es isomorfo a su grupo dual $\widehat{G}$, esto no es exclusivo de los grupos cíclicos, sabemos que todo grupo abeliano finito se puede escribir como suma directa de grupos cíclicos, Dirichlet probó lo siguiente

Teorema 2.6 (Dualidad)**.** Sea G un grupo abeliano finito de orden n, entonces G tiene exactamente n caracteres. Más aún, sea

$$G = \mathbb{Z}/m_1\mathbb{Z} \oplus \mathbb{Z}/m_2\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/m_\ell\mathbb{Z},$$

entonces los caracteres de G son de la forma

$$\chi((a_1, a_2, \ldots, a_\ell)) = e\left(\frac{a_1k_1}{m_1}\right) e\left(\frac{a_2k_2}{m_2}\right) \cdots e\left(\frac{a_\ell k_\ell}{m_\ell}\right),$$

donde $k_i = 0, 1, \ldots, m_i - 1$ para todo $i = 1, 2, \ldots, \ell$. Adicionalmente, $\widehat{G}$ es un grupo abeliano finito de orden n, $|\widehat{G}| = |G| = m_1m_2 \cdots m_k$, y $\widehat{G}$ es isomorfo a G:

$$\begin{aligned} \widehat{G} &\simeq \widehat{\mathbb{Z}/m_1\mathbb{Z}} \times \widehat{\mathbb{Z}/m_2\mathbb{Z}} \times \cdots \times \widehat{\mathbb{Z}/m_\ell\mathbb{Z}} \\ &\simeq \mathbb{Z}/m_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/m_\ell\mathbb{Z} \simeq G. \end{aligned} \tag{2.1}$$

Demostración. Primero consideremos el caso en que G es la suma directa de $\mathbb{Z}/m_i\mathbb{Z}$ para cada $i = 1, 2, 3, \ldots, \ell$, sea $e_i = (0, 0, \ldots, 1, 0, 0, \ldots, 0)$ donde 1 está en la posición i. Entonces e_i tiene orden m_i y por lo tanto $\chi(e_i)$ es una m_i-ésima raíz de la unidad, así pues

$$\begin{aligned} \chi((a_1, \ldots, a_\ell)) &= \chi(a_1e_1 + \ldots + a_\ell e_\ell) \\ &= \chi(e_1)^{a_1} \ldots \chi(e_2)^{a_2} \ldots \chi(e_\ell)^{a_\ell} \\ &= e\left(\frac{a_1k_1}{m_1}\right) e\left(\frac{a_2k_2}{m_2}\right) \ldots e\left(\frac{a_\ell k_\ell}{m_\ell}\right), \end{aligned}$$

donde $k_i = 0, 1, \ldots, m_i - 1$ para todo $i = 1, 2, \ldots, \ell$. De esto además vemos que χ se escribe como un producto de caracteres de $\mathbb{Z}/m_1\mathbb{Z}, \ldots, \mathbb{Z}/m\mathbb{Z}$, esto es, se tiene (2.1) y $|\widehat{G}| = |G| = m_1m_2 \cdots m_k$ dado que hay m_i posibles elecciones para k_i. Como G es abeliano finito entonces:

$$G \simeq \mathbb{Z}/m_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/m_\ell\mathbb{Z}.$$

lo que completa la prueba. □

2.1.1. Ortogonalidad de los caracteres

Los caracteres de un grupo abeliano finito satisfacen relaciones de ortogonalidad, cuando $G = (\mathbb{Z}/m\mathbb{Z})^\times$ estas relaciones de ortogonalidad nos permiten hacer análisis de Fourier, tenemos transformada y representación en "serie de Fourier", en nuestro caso esta serie será en realidad una suma finita, por lo que no nos preocuparemos por la convergencia de la misma. Las relaciones de ortogonalidad que veremos en esta sección son generalizaciones del teorema 2.5, recordemos que la idea es poder expresar la función característica de una clase de residuos módulo m como combinación lineal de caracteres, funciones completamente multiplicativas.

Teorema 2.7. Sea G un grupo abeliano finito. Entonces

(i) Dado un carácter χ de G

$$\sum_{g\in G}\chi(g)=\begin{cases}|G|, & \text{si } \chi=\chi_0\\ 0, & \text{e.o.c .}\end{cases}$$

(ii) Dado $g\in G$

$$\sum_{\chi\in\widehat{G}}\chi(g)=\begin{cases}|G|, & \text{si } g=1\\ 0, & \text{e.o.c .}\end{cases}$$

Demostración. Si $\chi=\chi_0$ entonces $\chi(g)=1$ para todo $g\in G$ por tanto la suma (i) es igual a $|G|$. Dado $\chi\neq\chi_0$, existe $h\in G$ tal que $\chi(h)\neq 1$, luego

$$\chi(h)\sum_{g\in G}\chi(g)=\sum_{g\in G}\chi(hg)=\sum_{g\in G}\chi(g),$$

ya que la suma se hace sobre todos los elementos del grupo, hg solo me permuta los elementos de la suma, así

$$(\chi(h)-1)\sum_{g\in G}\chi(g)=0,$$

y como $\chi(h)\neq 1$, obtenemos lo requerido. De manera análoga obtenemos (ii), si $g=1$, entonces $\chi(g)=1$ para todo $\chi\in\widehat{G}$ y la suma (ii) es igual a $|\widehat{G}|=|G|$ por el teorema 2.6. Suponga que $g\neq 1$, entonces existe $\psi\in\widehat{G}$ tal que $\psi(g)\neq 1$. Nuevamente

$$\psi(g)\sum_{\chi\in\widehat{G}}\chi(g)=\sum_{\chi\in\widehat{G}}(\psi\chi)(g)=\sum_{\chi\in\widehat{G}}\chi(g).$$

De donde se sigue el resultado.

□

Corolario 2.8 (Ortogonalidad)**.** Sea G un grupo abeliano finito. Entonces

(i) Si χ y ψ son caracteres de G

$$\sum_{g\in G}\psi(g)\overline{\chi}(g)=\begin{cases}|G|, & \text{si } \psi=\chi;\\ 0, & \text{e.o.c.}\end{cases}$$

(ii) Si g y h son elementos de G

$$\sum_{\chi\in\widehat{G}}\chi(g)\overline{\chi}(h)=\begin{cases}|G|, & \text{si } g=h\\ 0, & \text{e.o.c.}\end{cases}$$

Nota. Un lector familiarizado con algunos conceptos del análisis de Fourier podría notar que la suma presentada en (i) es muy similar al producto interno en $\ell^2(\mathbb{Z})$, esto no es para nada coincidencia, en esencia son el mismo producto interno.

Sea G un grupo abeliano finito se puede verificar que el espacio de funciones $f : G \to \mathbb{C}$, es un espacio vectorial de dimensión finita con el producto interno

$$\langle g, h\rangle = \sum_{g\in G} f(g)\overline{h(g)}.$$

Además podemos ver a $\widehat{G}$ como un subconjunto del espacio, ahora sí procedamos con la prueba.

Demostración. La suma en (i) se puede escribir como $\sum_{g\in G} \left(\psi\chi^{-1}\right)(g)$ y

$$\psi\chi^{-1} = \chi_0 \quad \text{si y solo si} \quad \psi = \chi.$$

La suma en (ii) es igual a $\sum_{x\in\widehat{G}} \chi\left(gh^{-1}\right)$, luego el resultado es una consecuencia inmediata del teorema anterior.

□

Procedamos formalmente suponiendo que tenemos

$$f(g) = \frac{1}{|G|}\sum_{\psi\in\widehat{G}} \widehat{f}(\psi)\psi(g),$$

por la ortogonalidad

$$\begin{aligned}\langle f, \chi\rangle &= \frac{1}{|G|}\sum_{g\in G}\sum_{\psi\in\widehat{G}} \widehat{f}(\psi)\psi(g)\overline{\chi}(g)\\ &= \frac{1}{|G|}\sum_{\psi\in\widehat{G}} \widehat{f}(\psi)\sum_{g\in G}\psi(g)\overline{\chi}(g)\\ &= \widehat{f}(\chi).\end{aligned}$$

Esto motiva la siguiente definición de transformada de Fourier.

Definición (Transformada de Fourier)**.** Sea $f : G \to \mathbb{C}$, definimos su transformada de Fourier como la función $\widehat{f} : \widehat{G} \to \mathbb{C}$ dada por

$$\widehat{f}(\chi) = \sum_{g\in G} f(g)\overline{\chi}(g).$$

Teorema 2.9 (Representación de Fourier)**.** Dada $f : G \to \mathbb{C}$, tenemos la representación en "serie" de Fourier

$$f(g) = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \widehat{f}(\chi)\chi(g).$$

Demostración. Dada $f : G \to \mathbb{C}$, por la ortogonalidad

$$\begin{aligned} \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \widehat{f}(\chi)\chi(g) &= \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \sum_{h \in G} f(h)\overline{\chi}(h)\chi(g) \\ &= \frac{1}{|G|} \sum_{h \in G} f(h) \sum_{\chi \in \widehat{G}} \overline{\chi}(h)\chi(g) \\ &= f(g). \end{aligned}$$

□

2.1.2. Caracteres de Dirichlet

En esta sección nos restringimos a estudiar los caracteres del grupo $(\mathbb{Z}/q\mathbb{Z})^\times$, sabemos que este es un grupo multiplicativo formado por las unidades del grupo aditivo $(\mathbb{Z}/q\mathbb{Z})$, las clases de residuos módulo q, conocemos un par de cosas de este grupo, por ejemplo, su orden está determinado por el número de elementos de $(\mathbb{Z}/q\mathbb{Z})$ que son primos relativos con q, esto es $\varphi(q)$.

Un carácter de Dirichlet es un carácter del grupo $(\mathbb{Z}/q\mathbb{Z})^\times$, sin embargo es útil presentarlo como una función aritmética que extiende un carácter de $(\mathbb{Z}/q\mathbb{Z})^\times$ a todos los naturales.

Definición (Carácter de Dirichlet)**.** Un carácter de Dirichlet es un carácter del grupo $(\mathbb{Z}/m\mathbb{Z})^\times$. Más generalmente, un carácter de Dirichlet módulo m es una función aritmética que extiende un carácter de $(\mathbb{Z}/m\mathbb{Z})^\times$mediante

$$f(n) = \begin{cases} \chi(a) & \text{si } n \equiv a (\text{mód} m), 1 \leqslant a \leqslant m, \text{ y } (a, m) = 1, \\ 0 & \text{si } (n, m) > 1. \end{cases}$$

Por ejemplo el carácter trivial se extiende de la siguiente manera:

$$\chi_0(n) = \begin{cases} 1, & \text{si } (n, m) = 1 \\ 0, & \text{si } (n, m) > 1 \end{cases}.$$

En algunos casos simplificaremos la notación como antes, solo escribiremos si $(n, m) = 1$ y no la condición $1 \leqslant a \leqslant m$ y $n \equiv a($ mód $m)$ ya que estas condiciones se pueden asumir que están presentes dado que estamos trabajando sobre clases de residuos, de todas manera es importante aclarar que solo es una herramienta de notación.

Teorema 2.10. Sea χ un carácter de Dirichlet módulo q, entonces $\chi(mn) = \chi(m)\chi(n)$ y $\chi(n+q) = \chi(n)$ para todo $m, n \in \mathbb{N}$.

Esto es que en efecto cuando se extiende el carácter sobre todos los naturales, obtenemos una función periódica de periodo q y además esta sigue siendo completamente multiplicativa.

Demostración. Sea χ un carácter de Dirichlet módulo q y sean $m, n \in \mathbb{N}$. Sea f el carácter de $(\mathbb{Z}/q\mathbb{Z})^\times$ que induce χ, esto es, para todo $n \in \mathbb{N}$,

$$\chi(n) = \begin{cases} f(n), & \text{si } (n,q) = 1 \\ 0, & \text{si } (n,q) > 1 \end{cases}$$

si $(m,q) > 1$ o $(n,q) > 1$, entonces $(mn,q) > 1$ y por tanto $\chi(mn) = 0 = \chi(m)\chi(n)$. Si $(m,q) = (n,q) = 1$, entonces $(mn,q) = 1$ y

$$\chi(m)\chi(n) = f(m)f(n) = f(mn) = \chi(mn).$$

Esto prueba que χ es completamente multiplicativa, tenemos que $(n+q,q) = (n,q)$ se sigue que si $(n+q,q) = 1$, entonces $(n,q) = 1$ y por tanto $\chi(n+q) = f(n+q) = f(n) = \chi(n)$. Análogamente, si $(n+q,q) > 1$, entonces $(n,q) > 1$ y $\chi(n+q) = 0 = \chi(n)$, entonces χ es periódica de periodo q.

□

Dada $f(n)$ la función característica de la progresión aritmética, es decir

$$f(n) = \begin{cases} 1, & n \equiv a \pmod{m} \\ 0, & n \not\equiv a \pmod{m}, \end{cases}$$

entonces

$$\hat{f}(\chi) = \sum_{g \in (\mathbb{Z}/m\mathbb{Z})^\times} f(g)\overline{\chi}(g) = \overline{\chi}(a) = \chi\left(a^{-1}\right),$$

ya que el término $f(g)$ es 0 excepto cuando $g = a$ donde $f(g)$ vale 1, por lo tanto tenemos la representación de Fourier

$$\boxed{f(n) = \frac{1}{\varphi(m)} \sum_{\chi} \chi\left(a^{-1}\right)\chi(n).}$$

Con este resultado hemos resuelto nuestro primer inconveniente para construir una prueba de teorema de Dirichlet. Observemos lo siguiente:

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{f(n)}{n^s} &= \frac{1}{\varphi(m)} \sum_{n=1}^{\infty} \left(\frac{\sum_{\chi} \chi(a^{-1})\chi(n)}{n^s} \right) \\ &= \frac{1}{\varphi(m)} \sum_{\chi} \chi(a^{-1}) \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}, \end{aligned}$$

en donde la serie de Dirichlet

$$\sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = L(s,\chi)$$

si resulta tener producto de Euler por lo que obtendremos la prueba de estudiar la expresión

$$\frac{1}{\varphi(m)} \sum_{\chi} \chi(a^{-1}) \log(L(1,\chi)). \tag{2.2}$$

En la siguiente sección veremos que (2.2) es divergente y que en efecto es igual a la suma sobre los primos congruentes a $a \pmod m$.

2.2 La L-serie asociada a un carácter de Dirichlet

En esta sección estudiaremos el comportamiento de las sumas parciales de $\chi(n)$, en particular nuestro objetivo es completar los ingredientes para la prueba del teorema de Dirichlet, por lo que en particular estudiaremos la no nulidad de

$$\sum_{n=1}^{\infty} \frac{\chi(n)}{n}.$$

La razón de estudiar estas propiedades se esclarecerá con la prueba del teorema de Dirichlet, allí entenderemos la importancia de cada una de estas piezas que hemos ido construyendo.

Definición. Sea χ un carácter de Dirichlet módulo q, la L-serie de Dirichlet asociada a χ (también llamada L-función) es la serie (la función)

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

Teorema 2.11. $L(\chi, s)$ converge absolutamente si $\Re(s) > 1$ y en este semiplano tenemos el producto de Euler

$$\sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \prod_{p} \frac{1}{1-\chi(p)p^{-s}}.$$

Demostración. Tenemos que $|\chi(n)| = 1$, luego por la convergencia absoluta de $\zeta(s)$ en el semiplano $\Re(s) > 1$ obtenemos la convergencia deseada, además como χ es completamente multiplicativa el producto de Euler se sigue del corolario 1.35. □

El teorema anterior nos da la siguiente expresión:

$$L(s,\chi_0) = \prod_p \left(1 - \frac{\chi_0(p)}{p^s}\right)^{-1}$$
$$= \prod_{p\nmid m} \left(1 - \frac{1}{p^s}\right)^{-1},$$

para obtener una expresión precisa que podamos controlar sobre este producto una idea sería completar el producto sobre todos los primos y quitar los términos $p \mid m$ en otro producto, esto es:

$$\begin{aligned} L(s,\chi_0) &= \prod_p \left(1 - \frac{\chi_0(p)}{p^s}\right)^{-1} \\ &= \prod_{p\nmid m} \left(1 - \frac{1}{p^s}\right)^{-1} \\ &= \prod_{p\mid m} \left(1 - \frac{1}{p^s}\right) \prod_p \left(1 - \frac{1}{p^s}\right)^{-1} \\ &= \prod_{p\mid m} \left(1 - \frac{1}{p^s}\right) \zeta(s). \end{aligned}$$

La expresión $\prod_{p\mid m} \left(1 - \frac{1}{p^s}\right)$ es continua cuando tomamos el límite $s \to 1$, a saber

$$\prod_{p\mid m} \left(1 - \frac{1}{p}\right) = \frac{\varphi(m)}{m},$$

ya obtuvimos propiedades analíticas sobre la función $\zeta(s)$, sabemos que esta es analítica en el semiplano $\sigma > 0$ excepto por un polo simple en $s = 1$ con residuo 1 (teorema 1.40), esto nos da el siguiente corolario.

Corolario 2.12. $L(s,\chi_0)$ es una función analítica en $\Re(s) > 0$, excepto por un polo simple en $s = 1$ con residuo $\dfrac{\varphi(m)}{m}$.

Debemos obtener ahora propiedades analíticas sobre $L(s,\chi)$ con $\chi \neq \chi_0$, dado χ un carácter de Dirichlet módulo q, sea $A(x) = \sum_{n\leqslant x} \chi(n)$, como $\chi \neq \chi_0$, obtenemos que $\sum_{n=1}^{q} \chi(n) = 0$, en general, $\sum_{n=1}^{kq} \chi(n) = 0$ para todo $k \in \mathbb{N}$. Entonces

$$|A(x)| \leqslant \sum_{n=1}^{q} |\chi(n)| = \sum_{\substack{n=1 \\ (n,q)=1}}^{q} 1 = \varphi(q) = O(1),$$

y aplicando sumación parcial se sigue que

$$\begin{aligned} L(s,\chi) &= \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \int_{1^-}^{\infty} \frac{1}{x^s} dA(x) \\ &= \frac{A(x)}{x^s}\Big|_{1^-}^{\infty} + s\int_1^{\infty} \frac{A(x)}{x^s}\frac{dx}{x} \\ &= s\int_1^{\infty} \frac{A(x)}{x^{s+1}} dx, \end{aligned}$$

donde el término de borde se anula ya que $A(x)$ es $O(1)$ y en el límite inferior las sumas parciales son 0. Note que la expresión anterior es válida para $\Re(s) > 0$ dado que $A(x) = O(1)$, además ya vimos que la expresión

$$\int_1^{\infty} \frac{1}{x^{s+1}} dx$$

es analítica en el semiplano $\sigma > 0$ (Teorema 1.40), obtenemos el siguiente corolario

Corolario 2.13. Si $\chi \neq \chi_0$, $L(s,\chi)$ es una función analítica en $\Re(s) > 0$. (no tiene polos)

2.2.1. Prueba del teorema de Dirichlet

Teorema 2.14 (Dirichlet)**.** Sea $\chi \neq \chi_0$, entonces

$$\log L(1,\chi) \neq 0.$$

Omitiremos la prueba de este teorema hasta completar la prueba del teorema de Dirichlet.

Teorema 2.15 (Dirichlet 1837)**.** Sean $(a,m) = 1$, entonces existen infinitos primos en la progresión aritmética: $a, a+m, a+2m, a+3m, a+4m, \ldots$

Demostración. Imitando las ideas de Euler tenemos para $\Re(s) > 1$

$$\log L(s,\chi) = \sum_p \sum_{n=1}^{\infty} \frac{\chi(p)}{np^{ns}} = \sum_p \frac{\chi(p)}{p^s} + \sum_{n=2}^{\infty}\sum_p \frac{\chi(p)}{np^{ns}} = \sum_p \frac{\chi(p)}{p^s} + O(1).$$

Ahora usando que

$$f(n) = \frac{1}{\varphi(m)} \sum_{\chi(a^{-1})} \chi(a^{-1})\chi(n) = \begin{cases} 1, & n \equiv a \pmod m \\ 0, & n \not\equiv a \pmod m \end{cases}$$

tenemos lo siguiente:

$$\begin{aligned}\frac{1}{\varphi(m)}\sum_{\chi}\chi(a^{-1})\log L(s,\chi) &= \frac{1}{\varphi(m)}\sum_{\chi}\chi(a^{-1})\left(\sum_{p}\frac{\chi(p)}{p^s}+O(1)\right)\\ &= \sum_{p}\frac{1}{p^s}\left(\frac{1}{\varphi(m)}\sum_{\chi}\chi(a^{-1})\chi(p)\right)+O(1)\\ &= \sum_{p}\frac{f(p)}{p^s}+O(1)\\ &= \sum_{p\equiv a(m)}\frac{1}{p^s}+O(1).\end{aligned}$$

Así para mostrar que la serie diverge, y por lo tanto que existen infinitos primos en la progresión aritmética, resta ver que

$$\lim_{s\to 1^+}\frac{1}{\varphi(m)}\sum_{\chi(a^{-1})}\log L(s,\chi)=+\infty,$$

ya sabemos que

$$\lim_{s\to 1^+}\log L(s,\chi_0)=+\infty,$$

por el teorema anterior tenemos que para $\chi\neq\chi_0$

$$\lim_{s\to 1^+}\log L(s,\chi)$$

es finito ya que $L(1,\chi)\neq 0$ y $L(s,\chi)$ no tiene polos en $\Re(s)>0$, entonces el logaritmo no explota, lo anterior nos da que

$$\begin{aligned}+\infty &= \lim_{s\to 1^+}\frac{1}{\varphi(m)}\sum_{\chi}\chi(a^{-1})\log L(s,\chi)\\ &= \sum_{p\equiv a(m)}\frac{1}{p}+O(1).\end{aligned}$$

□

La idea que tuvo Dirichlet de usar caracteres para escribir la función característica como combinación lineal de funciones completamente multiplicativas quizás venga de que Dirichlet no estudió en Alemania, estudió en Francia bajo supervisión de varios matemáticos, entre ellos Fourier, por lo que venía con algunas ideas de representación de funciones, ideas que estaban surgiendo del estudio de la ecuación del calor, esta es la razón por la que se considera a este resultado como el nacimiento de la teoría analítica de números, conecta dos áreas de la matemática aparentemente alejadas.

La primera prueba que presentó Dirichlet estaba incompleta, solo abordaba el caso en el que m era un número primo, para el caso general tuvo que asumir su fórmula del número de clases que demostró en un artículo en 1839-1840, varios años luego de su primera prueba,

al final de su artículo menciona que la primera prueba que se le ocurrió del resultado vital que necesitaba (la no nulidad de $L(1,\chi)$, $\chi \neq \chi_0$) venía de argumentos más indirectos y complicados, dice "no creo que haya un indicio en algún lugar de su naturaleza".

Nota. La prueba original de $L(1,\chi) \neq 0$ para un carácter real χ viene de asociar este carácter al símbolo de Legendre y es una simple consecuencia de su fórmula del número de clases. Sea $K = \mathbb{Q}(\sqrt{D})$ un extensión cuadrática con $D \equiv 0,1 \pmod 4$, existe un carácter primitivo real asociado χ tal que

$$\zeta_K(s) = \sum_{\mathfrak{a}} (N\mathfrak{a})^{-s} = \zeta(s)L(s,\chi),$$

y cada carácter real primitivo módulo $q \neq 1$ se obtiene de esta forma a partir de $K = \mathbb{Q}(\sqrt{\chi(-1)q})$. Dirichlet demostró que

$$L(1,\chi) = \begin{cases} \dfrac{2\pi h}{w\sqrt{q}} & \text{si } \chi(-1) = -1, \\ \dfrac{2h\log\varepsilon}{\sqrt{q}} & \text{si } \chi(-1) = 1, \end{cases}$$

donde h es el número de clases de K, w es el número de unidades en el anillo de enteros de K si $\chi(-1) = -1$, y ε es la unidad fundamental de K si $\chi(-1) = 1$. Por lo tanto $L(1,\chi) > 1/\sqrt{q}$. [11]

2.3 La no nulidad de $L(1,\chi)$

En esta sección dividiremos la prueba en dos partes, primero veremos la no nulidad de $L(1,\chi)$ para $\chi \neq \chi_0$ un carácter real, luego la no nulidad para un carácter a valor complejo, haremos esto al igual que Dirichlet puesto que es más sencillo trabajar los dos casos por separado, toda nuestra atención estará al inicio sobre el caso del carácter real dado que es el caso más complicado, sin embargo la prueba que presentaremos no es la de Dirichlet, es posible obtener este resultado con el método de Dirichlet de la hipérbola que estudiamos en el capítulo 1. Dado un carácter real $\chi \neq \chi_0$ módulo n considere la función

$$A(n) = \sum_{j|n} \chi(j),$$

la pregunta natural es ¿cuál es la serie de Dirichlet asociada a $A(n)$?, note que $A(n) = (\chi * 1)(n)$, entonces

$$\sum_{n=1}^{\infty} \frac{A(n)}{n^s} = \sum_{n=1}^{\infty} \frac{(\chi*1)(n)}{n^s} = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} \sum_{n=1}^{\infty} \frac{1}{n^s} = L(s,\chi)\zeta(s),$$

es posible ver que esta expresión es justamente $\zeta_K(s)$ donde $\zeta_K(s)$ es la función zeta de Dedekind en la extensión cuadrática $K = \mathbb{Q}(\sqrt{D})$ con D el entero positivo mayor que 1 más pequeño que divide a n.

En 1895 Mertens presentó una prueba de este teorema haciendo estimaciones sobre la función

$$B(x) = \sum_{n\leqslant x} \frac{A(n)}{\sqrt{n}} = \sum_{n\leqslant x} \sum_{d|n} \frac{\chi(d)}{\sqrt{n}},$$

demostró que $B(x) = 2\sqrt{x}L(1,\chi) + O(1)$ y que $B(x)$ diverge cuando tomamos $x \to \infty$, por lo tanto $L(1,\chi) \neq 0$, para ver esto utiliza el método de Dirichlet de la hipérbola dado $B(x)$ son las sumas parciales de una convolución. La prueba que presentaremos es la de Landau salvo por unas pequeñas modificaciones, las ventaja es que esta demostración es más sencilla y autocontenida que la de Dirichlet, no requerimos herramientas de la teoría algebraica de números, aunque como vimos la idea subyacente proviene de allí.

Teorema 2.16. Sea $\chi = \chi_0$ un carácter módulo k y sea f una función no negativa con derivada continua tal que $f'(x) < 0$ para $x \geqslant x_0$. Entonces si $y \geqslant x \geqslant x_0$, tenemos que

$$\sum_{y<n\leqslant x} \chi(n)f(n) = O(f(x)),$$

si adicionalmente $f(x) \to 0$ cuando $x \to \infty$ entonces la serie

$$\sum_{n=1}^{\infty} \chi(n)f(n)$$

converge y tenemos que para $x \geqslant x_0$

$$\sum_{n\leqslant x} \chi(n)f(n) = \sum_{n=1}^{\infty} \chi(n)f(n) + O(f(x))$$

Demostración. Tenemos que $A(x) = \sum_{n\leqslant x} \chi(n) = O(1)$, por lo tanto

$$\begin{aligned}
\sum_{x<n\leqslant y} f(n)f(n) &= \int_x^y f(t)\,d(A(t)) \\
&= f(y)A(y) - f(x)A(x) - \int_x^y A(t)f'(t)\,dt \\
&\ll f(y) + f(x) + \int_x^y (-f'(t))\,dt \\
&\ll f(x),
\end{aligned}$$

si adicionalmente $f(x) \to 0$ cuando $x \to \infty$, $\sum_{x\leqslant n\leqslant y} x(n)f(n) \ll f(x) \to 0$ cuando $x, y \to \infty$ esto es $\sum_{n=1}^{\infty} x(n)f(n)$ converge por el criterio de Cauchy. Finalmente note que

$$\sum_{n=1}^{\infty} \chi(n)f(n) = \sum_{n\leqslant x} \chi(n)f(n) + \lim_{y\to\infty} \sum_{x<n\leqslant y} \chi(n)f(n)$$
$$= \sum_{n\leqslant x} \chi(n)f(n) + O(f(x))$$

□

Aplicando este teorema para $f(x) = \frac{1}{x}$, $f(x) = \frac{\log x}{x}$ y $f(x) = \frac{1}{\sqrt{x}}$ obtenemos el siguiente corolario

Corolario 2.17. Si $\chi \neq \chi_0$ es un carácter módulo k y si $x \geqslant 1$, tenemos que

$$\sum_{n\leqslant x} \frac{\chi(n)}{n} = L(1,\chi) + O\left(\frac{1}{x}\right)$$
$$\sum_{n\leqslant x} \frac{\chi(n)\log n}{n} = \sum_{n=1}^{\infty} \frac{\chi(n)\log n}{n} + O\left(\frac{\log x}{x}\right)$$
$$\sum_{n\leqslant x} \frac{\chi(n)}{\sqrt{n}} = L\left(\frac{1}{2},\chi\right) + O\left(\frac{1}{\sqrt{x}}\right).$$

2.3.1. No nulidad para el carácter real $\chi \neq \chi_0$

Teorema 2.18. Sea χ un carácter real módulo k y considere la función

$$A(n) = \sum_{d|n} \chi(n).$$

Entonces $A(n) \geqslant 0$ para todo n y $A(n) \geqslant 1$ si n es un cuadrado perfecto.

Demostración. Para las potencias de primos tenemos que

$$A(p^a) = \sum_{t=0}^{a} \chi(p^t) = 1 + \sum_{t=1}^{a} \chi(p)^t,$$

como χ es a valor real, entonces los único valores de $\chi(p)$ son $0, 1$ y -1, si $\chi(p) = 0$, entonces $A(p^a) = 1$, si $\chi(p) = 1$ entonces $A(p^a) = a + 1$, si $\chi(p) = -1$ entonces

$$A(p^a) = \begin{cases} 0 & \text{si } a \text{ es impar,} \\ 1 & \text{si } a \text{ es par.} \end{cases}$$

En cualquier caso, $A(p^a) \geqslant 1$ si a es par. Ahora, si $n = p_1^{a_1} \cdots p_r^{a_r}$ entonces $A(n) = A(p_1^{a_1}) \ldots A(p_r^{a_r})$ ya que A es multiplicativa, dado que es convolución de dos funciones multiplicativas. Cada factor $A(p_i^{a_i}) \geqslant 0$ de esto se sigue que $A(n) \geqslant 0$, además si n es un cuadrado perfecto, entonces cada exponente a_i es par, esto es que cada factor $A(p_i^{a_i}) \geqslant 1$ por lo cual $A(n) \geqslant 1$. □

Lema 2.19 (Sumas parciales de $\zeta(s)$)**.** Dado $x \geqslant 1$, $s > 0$ y $s \neq 1$ tenemos que

$$\sum_{n \leqslant x} \frac{1}{n^s} = \frac{x^{1-s}}{1-s} + \zeta(s) + O\left(x^{-s}\right)$$

Demostración. Note que

$$\begin{aligned}
\sum_{n \leqslant x} \frac{1}{n^s} &= 1 + \int_1^x \frac{1}{t^s} d[t] \\
&= 1 + \int_1^x \frac{1}{t^s} dt - \int_1^x \frac{1}{t^s} d\{t\} \\
&= 1 + \frac{x^{1-s}}{1-s} - \frac{1}{1-s} - \int_1^x \frac{1}{t^s} d\{t\} \\
&= 1 + \frac{x^{1-s}}{1-s} - \frac{1}{1-s} - \left(\frac{\{x\}}{x^s} + s \int_1^x t^{-s-1} \{t\} dt \right).
\end{aligned}$$

Como $\{t\} = O(1)$, entonces

$$\begin{aligned}
\sum_{n \leqslant x} \frac{1}{n^s} &= \frac{x^{1-s}}{1-s} + 1 + \frac{1}{s-1} - s \int_1^x \frac{\{t\}}{t^{s+1}} dt + O(x^{-s}) \\
&= \frac{x^{1-s}}{1-s} + \frac{s}{s-1} - s \int_1^x \frac{\{t\}}{t^{s+1}} dt + O(x^{-s}) \\
&= \frac{x^{1-s}}{1-s} + \zeta(s) + O(x^{-s}),
\end{aligned}$$

por la representación integral de $\zeta(s)$, esta representación integral tiene sentido para $s > 0$ con $s \neq 1$, lo que concluye el resultado. □

Teorema 2.20. Dado $\chi \neq \chi_0$ un carácter real módulo k, sea

$$A(n) = \sum_{d|n} \chi(d) \quad \text{y} \quad B(x) = \sum_{n \leqslant x} \frac{A(n)}{\sqrt{n}}.$$

Tenemos que

i) $B(x) \to \infty$ cuando $x \to \infty$

ii) $B(x) = 2\sqrt{x} L(1,\chi) + O(1)$

Demostración. Tenemos que

$$B(x) \geqslant \sum_{\substack{n \leqslant x \\ n = m^2}} \frac{1}{\sqrt{n}} = \sum_{n \leqslant x} \frac{1}{m}$$

y por lo tanto cuando $x \to \infty$ la suma diverge por la divergencia de la serie armónica, esto prueba i). Note que cambiando el orden de sumación

$$\begin{aligned} B(x) &= \sum_{n\leqslant x}\sum_{d|n}\frac{\chi(d)}{\sqrt{n}} = \sum_{d\leqslant x}\sum_{\substack{n\leqslant x\\ d|n}}\frac{\chi(d)}{\sqrt{n}} \\ &= \sum_{d\leqslant x}\chi(d)\sum_{\substack{n\leqslant x\\ d|n}}\frac{1}{\sqrt{n}} \\ &= \sum_{d\leqslant x}\chi(d)\sum_{k\leqslant \frac{x}{d}}\frac{1}{\sqrt{kd}} \\ &= \sum_{dk\leqslant x}\frac{\chi(d)}{\sqrt{kd}}. \end{aligned}$$

Dado que la última expresión para $B(x)$ son sumas parciales de una convolución, podemos aplicar el método de Dirichlet de hipérbola como sigue

$$\begin{aligned} B(x) &= \sum_{d\leqslant\sqrt{x}}\frac{\chi(d)}{\sqrt{d}}\sum_{k\leqslant\frac{x}{d}}\frac{1}{\sqrt{k}} + \sum_{k\leqslant\sqrt{x}}\frac{1}{\sqrt{k}}\sum_{d\leqslant x/k}\frac{\chi(d)}{\sqrt{d}} - \sum_{k\leqslant\sqrt{x}}\frac{1}{\sqrt{k}}\sum_{d\leqslant\sqrt{k}}\frac{\chi(d)}{\sqrt{d}} \\ &= S_1 + S_2 - S_3, \end{aligned}$$

Vamos a trabajar por separado las expresiones S_1, S_2 y S_3, note que podemos estimar $\sum_{k\leqslant\frac{x}{d}}\frac{1}{\sqrt{k}}$ aplicando el teorema anterior y tomaremos $a = \zeta(s)$ dado que para este valor de s la suma converge, así obtenemos

$$\begin{aligned} S_1 &= \sum_{d\leqslant\sqrt{x}}\frac{\chi(d)}{d}\left(2\left(\frac{x}{d}\right)^{1/2} + a + O\left(\left(\frac{d}{x}\right)^{1/2}\right)\right) \\ &= 2\sqrt{x}\sum_{d\leqslant\sqrt{x}}\frac{\chi(d)}{d} + a\sum_{d\leqslant\sqrt{x}}\frac{\chi(d)}{\sqrt{d}} + O\left(\frac{1}{\sqrt{x}}\sum_{d\leqslant\sqrt{x}}|\chi(d)|\right) \\ &= 2\sqrt{x}\left(L(1,\chi) + O(x^{1/2})\right) + aL\left(\frac{1}{2},\chi\right) + O\left(x^{-1/4}\right) + O(1) \\ &= 2\sqrt{x}L(1,\chi) + aL\left(\frac{1}{2},\chi\right) + O(1) \\ &= 2\sqrt{x}L(1,\chi) + O(1). \end{aligned}$$

Análogamente estimamos la suma $\sum_{k\leqslant\sqrt{x}}\frac{1}{\sqrt{k}}$

$$\begin{aligned} S_2 &= \sum_{k\leqslant\sqrt{x}} \frac{1}{\sqrt{k}} \left(L(1/2,\chi) + O\left(\frac{\sqrt{k}}{\sqrt{x}}\right)\right) \\ &= L\left(\frac{1}{2},\chi\right)\left(2x^{1/4}+O(1)\right) + O\left(\sum_{k\leqslant\sqrt{x}} \frac{1}{\sqrt{x}}\right) \\ &= L\left(\frac{1}{2},\chi\right) 2x^{1/4} + O(1) \end{aligned}$$

y

$$\begin{aligned} S_3 &= \left(L(1/2,\chi) + O\left(\frac{1}{x^{1/4}}\right)\right)\left(2x^{1/4}+O(1)\right) \\ &= L\left(\frac{1}{2},\chi\right) 2x^{1/4} + O(1), \end{aligned}$$

por lo tanto $B(x) = S_1 + S_2 - S_3 = 2\sqrt{x}L(1,\chi) + O(1)$ ya que los términos $L(\frac{1}{2},\chi)2x^{\frac{1}{4}}$ se cancelan.

□

Como $B(x)$ diverge cuando $x \to \infty$ entonces $L(1,\chi) \neq 0$

2.3.2. No nulidad del carácter complejo $\chi \neq \chi_0$

Considere la función

$$P(s) = \prod_{\chi} L(s,\chi),$$

donde χ es un carácter de Dirichlet módulo q, note que

$$P(s) = L(\chi_0,s) \prod_{\chi\neq\chi_0} L(s,\chi),$$

sabemos que este último producto está acotado en $\Re(s) > 0$ por el corolario 2.13 y que la función $L(\chi,s)$ tiene un polo simple en $s = 1$ dado por $\zeta(s)$, el teorema 1.40 nos dice más aún, nos da que cuando $s \to 1^+$, $L(s,\chi_0) = O\left(\frac{1}{s-1}\right)$, Supongamos que $L(1,\chi_1) = 0$ para algún carácter $x_1 \neq \chi_0$, como χ_1 es un carácter complejo, entonces $\chi_1 \neq \overline{\chi_1}$, además

$$L(s,\overline{\chi_1}) = \sum_{n=1}^{\infty} \frac{\overline{\chi_1}(n)}{n^s} = \overline{\sum_{n=1}^{\infty} \frac{\chi_1(n)}{n^s}} = \overline{L(s,\chi_1)}$$

de esto se sigue que si $L(1,\chi_1) = 0$, entonces $L(1,\overline{\chi_1}) = 0$, estas consideraciones nos permiten presentar una prueba sencilla de la no nulidad de $L(1,\chi)$, note que fue necesario separar el caso del carácter real dado que si χ_1 es un carácter real entonces $\chi_1 = \overline{\chi_1}$ y esto daña el argumento.

Lema 2.21. Sea $P(s) = \prod_{\chi} L(s,\chi)$, si $s > 1$ entonces $P(s) \geqslant 1$.

Demostración. Tenemos que

$$\begin{aligned}\log(P(s)) &= \sum_{\chi} \log L(s,\chi) = \sum_{\chi \bmod q} \sum_{p} \log\left(1 - \frac{\chi(p)}{p^s}\right)^{-1} \\ &= \sum_{\chi} \sum_{p} \sum_{m=1}^{\infty} \frac{\chi(p)^m}{m p^{ms}} \\ &= \sum_{p} \sum_{m=1}^{\infty} \frac{1}{m p^{ms}} \sum_{\chi} \chi(p)^m.\end{aligned} \quad ,$$

por la ortogonalidad de los caracteres tenemos que

$$\sum_{\chi} \chi(p)^m = \sum_{\chi} \overline{\chi}(1)\chi(p^m) = \begin{cases} \varphi(q) & \text{si } p^m \equiv 1 (\text{mód} q), \\ 0 & \text{e.o.c,} \end{cases}$$

esto nos da que $\sum_{p} \sum_{m=1}^{\infty} \frac{1}{m p^{ms}} \sum_{\chi} \chi(p)^m$ es un suma de términos no negativos, de esto se sigue que

$$P(s) = \exp\left(\sum_{p} \sum_{m=1}^{\infty} \frac{1}{m p^{ms}} \sum_{\chi} \chi(p)^m\right) \geqslant 1$$

□

Teorema 2.22. Sea $\chi \neq \chi_0$ un carácter complejo módulo q, entonces $L(1,\chi) \neq 0$.

Demostración. Supongamos que existe $\chi_1 \neq \chi_0$ tal que $L(1,\chi_1) = 0$, entonces $L(1,\overline{\chi_1}) = 0$, por lo tanto cuando $s \to 1^+$ tenemos que $L(s,\chi_1) = O(s-1) = L(s,\overline{\chi_1})$, en efecto

$$P(s) = L(s,\chi_0)L(s,\chi_1)L(s,\overline{\chi_1})Q(s),$$

con

$$Q(s) = \prod_{\chi \neq \chi_1, \overline{\chi_1}, \chi_0} L(s,\chi) = O(1), \quad \Re(s) > 0.$$

Al tomar el $\lim_{s \to 1^+} P(s)$ basta controlar los otros 3 factores, pero ya vimos que $L(s,\chi_1) = O(s-1) = L(s,\overline{\chi_1})$, por lo tanto $P(s) = O(s-1)$ cuando $s \to 1^+$, esto contradice el lema 2.21.

□

Sea $s > 0$, en esencia lo que ocurre en la prueba anterior es que los ceros de $L(s,\chi)$, con $\chi \neq \chi_0$ vienen en parejas por lo que si $L(s,\chi) = 0$, entonces $L(s,\overline{\chi}) = 0$, esto es conveniente para este caso, un cero nos anula el efecto del polo dado por el carácter trivial y el otro nos da la contradicción.

El teorema de los números primos

> *Es precipitado afirmar que un teorema matemático no puede demostrarse de una manera particular; pero algo parece bastante claro... tenemos ciertas ideas sobre la lógica de la teoría; creemos que algunos teoremas, como solemos decir, "yacen profundamente" y otros están más cerca de la superficie. Si alguien produce una demostración elemental del teorema de los números primos, mostrará que estas ideas son erróneas, que el tema no se sostiene de la manera que habíamos supuesto, y que es hora de desechar los libros y reescribir la teoría.*
>
> **— Godfrey Harold Hardy (1921)**
> *Conferencia a la Sociedad Matemática de Copenhague*

Obtener cotas, tanto superiores como inferiores, para la función contadora de primos $\pi(x)$ es un desafío notablemente difícil. En vista de estas dificultades, es excepcional que a mediados del siglo XIX, el matemático ruso P. L. Chebyshev lograra determinar el orden preciso de magnitud de $\pi(x)$. Chebyshev demostró que existen constantes positivas A y B tales que

$$A\frac{x}{\log x} \leqslant \pi(x) \leqslant B\frac{x}{\log x}$$

para valores de x suficientemente grandes. De hecho, obtuvo las cotas específicas de $A \approx 0,92129$ y $B = (6/5)A \approx 1,10555$, esto le permitió demostrar el Postulado de Bertrand, que establece que para $x > 2$, siempre existe un número primo p tal que $x < p < 2x$, además demostró que si $\pi(x)/(x \log x)$ tenía un límite cuando $x \to \infty$, entonces este límite tenía que ser 1.

El método de Chevyshev consistía en explotar el TFA para convoluciones

$$\sum_{j|n} \Lambda(j) = \log n,$$

para realizar estimaciones sobre las sumas parciales del logaritmo natural, esto es $\log([x]!)$, en retrospectiva la aproximación de Chebyshev se basa en la relación

$$[x] = \prod_{p} p^{\nu_p([x])}, \quad \text{donde } \nu_p([x]) \text{ es el orden } p\text{-ádico}$$

de esta identidad se puede comparar la valuación p-ádica con la arquimediana, muchos intentos de producir una prueba usando estos métodos fallaron, el teorema se resistió a una prueba elemental por lo siguientes cien años.

Años después, en 1859, aparece el famoso artículo de Riemann "Sobre la cantidad de primos menores que una magnitud dada", allí se encontraba el camino hacia una prueba del TNP. La idea revolucionaria de Riemann fue considerar a $\zeta(s)$ como una función de variable compleja y expresar a $\pi(x)$ en términos de una integral compleja que involucraba a $\zeta(s)$, más aún, Riemann obtiene la fórmula explícita

$$\pi(x) = \mathrm{Li}(x) - \sum_{\rho} \mathrm{Li}\,(x^{\rho}) - \log(2) + \int_{x}^{\infty} \frac{dt}{t\,(t^2-1)\log(t)},$$

en donde ρ denota un cero no trivial de la función $\zeta(s)$, sin embargo, no había suficiente análisis disponible para producir una prueba rigurosa. No fue hasta finales del siglo XIX cuando se proporcionó el ingrediente esencial que faltaba, este fue la teoría de las funciones enteras de orden finito, desarrollada por Hadamard para probar el TNP.

Riemann demostró que la función $\zeta(s)$ tiene una continuación analítica a $\mathbb{C}$ excepto por un polo simple en $s = 1$ con residuo 1 y además obtuvo su ecuación funcional

$$\zeta(s) = 2^s \pi^{s-1} \sin\left(\frac{\pi s}{2}\right) \Gamma(1-s)\zeta(1-s),$$

también reconoció el papel importante que juegan los ceros de $\zeta(s)$ en la teoría de números y conjetura varias propiedades de estos ceros, todas estas conjeturas excepto una fueron demostradas por Von Mangolth y Hadamard a finales del siglo XIX, la conjetura que se resistió es la ya muy famosa hipótesis de Riemann "la parte real de todos los ceros no triviales de $\zeta(s)$ es 1/2", al conjeturarla, Riemann dice "es muy probable que ..." (es ist sehr wahrscheinlich dass).[12]

En 1896 aparecen las pruebas de Hadamard y Charles-Jean de La Vallée Poussin, ambos prueban que $\zeta(1+it) \neq 0$ para todo $t \in \mathbb{R}$, esto es que la función $\zeta(s)$ no se anula en los complejos con parte real 1, esto les permite dar una prueba rigurosa siguiendo las ideas de Riemann, dicho esto, las pruebas de Hadamard y de la Vallée Poussin requieren análisis de la función zeta en una franja más grande que simplemente $\Re(s) \geqslant 1$ dado que involucran fórmulas explícitas en la prueba y en estas los ceros de $\zeta(s)$ juegan un papel importante.

En este capítulo veremos que el TNP implica que $\zeta(s) \neq 0$ para $\Re(s) = 1$, esto ya era conocido por los matemáticos de la época y nos dice que es necesaria la no nulidad en la recta vertical para obtener el teorema de los números primos, el primer matemático en dar una prueba que no dependía de la ecuación funcional de $\zeta(s)$ fue Landau, en su lugar trabaja con la extensión analítica en el semiplano $\Re(s) > 0$ que estudiamos previamente.

La pregunta natural que surge aquí es ¿puede el teorema de los números primos ser probado usando solo el hecho de que $\zeta(s) \neq 0$ en $\Re(s) = 1$?, la respuesta fue dada por N. Wiener en 1931 con su famoso teorema Tauberiano

Teorema 3.1. Sean $a_n \geqslant 0$ y $F(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ una serie absolutamente convergente. Supongamos que se cumplen las siguientes condiciones:

a) La función $F(s)$ se extiende a una función analítica en la región $\Re(s) \geqslant 1$ con un único polo simple en $s = 1$, cuyo residuo es 1.

b) $A(x) = \sum_{n \leqslant x} a_n = O(x)$.

Entonces, se tiene que

$$A(x) = x + o(x) \text{ cuando } x \to \infty.$$

De este teorema el TNP resulta ser un corolario, más aún, obtenemos una equivalencia entre el teorema de los números primos y $\zeta(1 + it) \neq 0$ ya que como podemos ver en las condiciones del teorema, no requerimos análisis por fuera de la franja $\Re(s) \geqslant 1$ y esta dirección será la que tomaremos en este trabajo.

3.1 El teorema de Wiener-Ikehara

Sea $\sum_{n=0}^{\infty} a_n x^n$, $x \in \mathbb{R}$ una serie de potencias centrada en 0 y con radio de convergencia 1, en el borde la región, la serie puede ser divergente o convergente. El teorema de Abel nos dice que si la serie es sumable en un punto del borde, entonces es continua en el punto, más precisamente, si

$$\sum_{n=0}^{\infty} a_n = A,$$

entonces

$$\lim_{x \to 1^-} \sum_{n=0}^{\infty} a_n x^n = A.$$

Uno podría preguntarse cuando esperar un recíproco de este teorema, bajo qué condiciones que este límite tenga cierto valor A, implica que la serie es sumable y converge a A, las primeras condiciones fueros establecidas por Tauber

Proposición 3.2 (Tauber, 1897)**.** Sea $f(x) = \sum_{n=0}^{\infty} a_n x^n$ una serie de potencias que converge absolutamente para $|x| < 1$. Si $\lim_{x \to 1^-} f(x) = A$ y se cumple la condición $a_n = o\left(\frac{1}{n}\right)$, entonces $f(1) = A$.

Los teoremas tauberianos son recíprocos condicionales del teorema de Abel, la condi-

ción $a_n = o\left(\frac{1}{n}\right)$ nos restringe el crecimiento de los coeficientes, estas condiciones fueron relajadas subsecuentemente por Hardy y Littlewood a $O\left(\frac{1}{n}\right)$ y posteriormente a $na_n > -K$.

Algunas de las aplicaciones más importantes de la teoría tauberiana son a la teoría analítica de números, en este contexto los resultados tauberianos se pueden considerar como estimaciones sobre las sumas parciales de los coeficientes de una serie de Dirichlet.
En 1980 el matemático Donald J. Newman dió una prueba corta del teorema de los números primos, utilizando el siguiente teorema tauberiano

Teorema 3.3 (Newman)**.** Supongamos que

$$|a_n| \leqslant 1 \text{ y } F(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}.$$

Entonces F es analítica en el semiplano $\sigma > 1$. Si F tiene extensión analítica en el semiplano $\sigma \geqslant 1$, entonces $\sum_{n=1}^{\infty} \frac{a_n}{n^s}$ converge en $\sigma \geqslant 1$.

La prueba de Newman inspiró dos grandes artículos que modificaron su prueba, los de Zagier [9] y Korevaar [13], en donde se aprovecha la relación que hay entre transformada de Laplace y series de Dirichlet, esto nos permite dar una prueba del TNP en la que solo se emplea el teorema de residuos de Cauchy.

En esta sección presentaremos una prueba del teorema de Wiener-Ikehara modificando la prueba de Newman, esto es, siguiendo el enfoque de Zagier y Korevaar de una manera distinta, este camino para la prueba fue encontrado por Korevaar [14], también presentaremos versiones más fuertes del teorema siguiendo a [15], estas versiones son las que necesitaremos para el TNP en progresiones aritmética del siguiente capítulo.

Teorema 3.4 (Korevaar y Zagier)**.** Para $t \geqslant 0$, sea $f(t)$ una función acotada y localmente integrable y sea

$$g(s) := \int_0^{\infty} f(t)e^{-st}dt,$$

para $\Re(s) > 0$. Si $g(s)$ tiene continuación analítica a $\Re(s) \geqslant 0$, entonces $\int_0^{\infty} f(t)dt$ existe y es igual a $g(0)$.

Demostración. Dado $T > 0$, sea

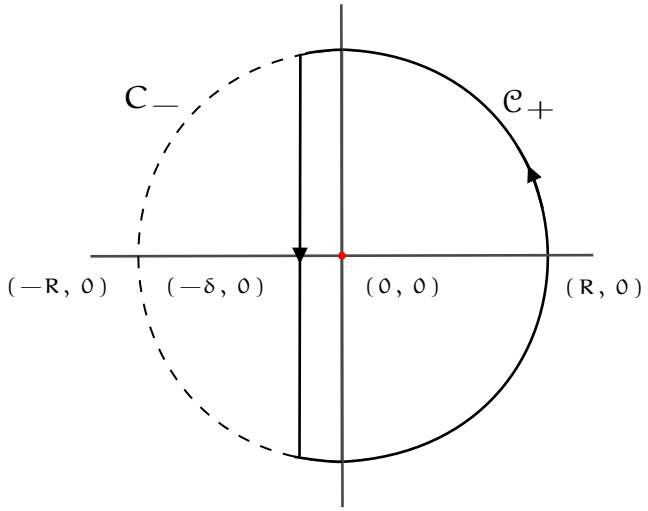
$$g_T(s) = \int_0^{T} f(t)e^{-st}dt,$$

por hipótesis $g_T(s)$ converge para todo $s \in \mathbb{C}$, y $g_T(s)$ es una función entera, queremos

ver que

$$\lim_{T\to\infty}\int_0^T f(t)dt = g(0).$$

Dado $R > 0$, considere el siguiente contorno positivamente orientado $\mathcal{C}$



donde $\delta > 0$ (dependiendo de R) se escoge de tal forma que $g(s)$ es analítica en $\mathcal{C}$, esto se puede hacer ya que el conjunto $A := \{(0, y) : y \in [-R, R]\}$ es compacto y como $g(s)$ es analítica, A se puede cubrir por bolas abiertas, en cada una de las cuales $g(s)$ es analítica, la compacidad permite escoger un subcubrimiento finito, del que obtenemos δ. Denotemos

$$\mathcal{C}_+ = \mathcal{C} \cap \{s : \sigma > 0\}, \qquad \mathcal{C}_- = \mathcal{C} \cap \{s : \sigma < 0\}$$

y como C_- el semicírculo de radio R a la izquierda de $\sigma = 0$, aplicando fórmula de la integral de Cauchy obtenemos que

$$I_{\mathcal{C}} = \frac{1}{2\pi i}\int_{\mathcal{C}} (g(s) - g_T(s))\, e^{sT}\left(1 + \frac{s^2}{R^2}\right)\frac{1}{s}ds = g(0) - g_T(0)$$

dado que todas las funciones en la integral $I_{\mathcal{C}}$ son analíticas, excepto por $1/s$ que tiene un polo simple en $s = 0$, de donde obtenemos el residuo $g(0) - g_T(0)$.

Sea $M = \sup_{t\geqslant 0}|f(t)|$. En $\mathcal{C}_+$, dado que $\sigma > 0$, tenemos

$$|g(s) - g_T(s)| = \left|\int_T^\infty f(t)e^{-st}dt\right| \leqslant M\int_T^\infty e^{-\sigma t}dt \ll \frac{e^{-\sigma T}}{\sigma}.$$

Tomando $s = Re^{i\theta}$, entonces $R\cos\theta = \sigma$ en $\mathcal{C}_+$, obtenemos la siguiente estimación:

$$\begin{aligned}\left|e^{sT}\frac{1}{s}\left(1 + \frac{s^2}{R^2}\right)\right| &= e^{\sigma T}\left|\frac{1}{Re^{i\theta}} + \frac{e^{i\theta}}{R}\right| \\ &= e^{\sigma T}\left|\frac{e^{-i\theta} + e^{i\theta}}{R}\right| \\ &= e^{\sigma T}\left|\frac{2\cos\theta}{R}\right| \ll e^{\sigma T}\frac{|\sigma|}{R^2},\end{aligned} \tag{3.1}$$

por lo tanto, acotando en la integral obtenemos que

$$|I_{\mathcal{C}_+}| \ll \frac{1}{R^2}\left|\int_{\mathcal{C}_+} ds\right| = \frac{1}{R^2}\pi R \ll \frac{1}{R}.$$

En $\mathcal{C}_-$, examinamos $g_T(s)$ y $g(s)$ por separado. Consideremos primero la integral

$$I_1 := \frac{1}{2\pi i}\int_{\mathcal{C}_-} g_T(s)e^{sT}\left(1+\frac{s^2}{R^2}\right)\frac{ds}{s}.$$

Como $g_T(s)$ es entera y el resto del integrando es analítico a la izquierda de $\sigma = 0$, tenemos que

$$I_1 = \frac{1}{2\pi i}\int_{C_-} g_T(s)e^{sT}\left(1+\frac{s^2}{R^2}\right)\frac{ds}{s}.$$

Es decir, podemos integrar sobre el semicírculo C_- en lugar de $\mathcal{C}_-$, con C_- orientado de la misma manera que $\mathcal{C}_-$. Luego, observando que $\sigma < 0$ en este caso, tenemos

$$|g_T(s)| = \left|\int_0^T f(t)e^{-st}dt\right| \leqslant M\int_0^T e^{-\sigma t}dt \ll \frac{e^{-\sigma T}}{|\sigma|}.$$

La estimación en (3.1) sigue siendo válida, por lo que de manera análoga obtenemos que

$$|I_1| \ll \frac{1}{R},$$

nos queda acotar la integral restante

$$I_2 = \frac{1}{2\pi i}\int_{\mathcal{C}_-} g(s)e^{st}\left(1+\frac{s^2}{R^2}\right)\frac{1}{s}ds.$$

Como $\mathcal{C}_-$ está contenido en un compacto en el que $g(s)$ es analítica, $|g(s)|$ se puede acotar por una constante que depende únicamente de R, digamos

$$|g(s)| \leqslant M_R,$$

para todo $s \in \mathcal{C}_-$, $s = e^{i\theta}$ o $s = -\delta + it$, esto es, s está en los arcos de circunferencia, o s está en la recta vertical $\Re(s) = -\delta$, debemos acotar la integral en estos dos caminos del contorno, para $s = e^{i\theta}$ ya tenemos una cota dada por (3.1), si $s = -\delta + it$ tenemos que

$$\begin{aligned}\left|\frac{e^{sT}}{s}\left(1+\frac{s^2}{R^2}\right)\right| &\leqslant \frac{e^{-\delta T}}{|s|}\left(1+\frac{|s|^2}{R^2}\right)\\ &\leqslant e^{-\delta T}\left(\frac{1}{|s|}+\frac{|s|}{R^2}\right)\\ &\leqslant e^{-\delta T}\left(\frac{1}{\delta}+\frac{1}{R}\right),\end{aligned}$$

ya que $|t| \leqslant R$, combinanto esta cota junto con (3.1) y sabiendo que $|g(s)| \leqslant M_R$, podemos acotar la integral como

$$|I_2| \ll \frac{M_R}{R^2}\left|\int_{\mathcal{C}_-} |\sigma| e^{\sigma T} ds\right| + \frac{R M_R e^{-\delta T}}{\delta} + M_R e^{-\delta T},$$

por lo tanto

$$\begin{aligned} |g(0) - g_T(0)| = |I_{\mathcal{C}}| &\leqslant |I_{\mathcal{C}_+}| + |I_1| + |I_2| \\ &\ll \frac{1}{R} + \frac{M_R}{R^2}\left|\int_{\mathcal{C}_-} |\sigma| e^{\sigma T} ds\right| + \frac{R M_R e^{-\delta T}}{\delta} + M_R e^{-\delta T}. \end{aligned} \tag{3.2}$$

En C_-, tenemos $\sigma < 0$ y $|\sigma| e^{\sigma T} \to 0$ uniformemente en C_- cuando $T \to \infty$. Por lo tanto, cuando $T \to \infty$, el lado derecho de (3.2) converge a $\frac{1}{R}$. Luego, tomando $R \to \infty$, obtenemos

$$g(0) = \lim_{T\to\infty} g_T(0).$$

□

Lema 3.5 (Korevaar y Zagier)**.** Sean $a_n \geqslant 0$ y $A(x) = \sum_{n\leqslant x} a_n$, si la integral

$$\int_1^\infty \frac{A(x) - x}{x^2} dx$$

converge, entonces $A(x) \sim x$

Demostración. Supongamos que $A(x) \not\sim x$, entonces existe una constante $\lambda > 1$ (o $\lambda < 1$) tal que $A(x) \geqslant \lambda x$ (o $A(x) \leqslant \lambda x$) para infinitos valores de $x \to \infty$.

Sin pérdida de generalidad consideramos el caso $\lambda > 1$. Dado que $A(x)$ es una función creciente, se sigue que para cualquiera de los valores de x con $A(x) \geqslant \lambda x$ y para $x \leqslant t \leqslant \lambda x$, tenemos $A(t) \geqslant A(x) \geqslant \lambda x$ y

$$\begin{aligned} \int_x^{\lambda x} \frac{A(t) - t}{t^2} dt \geqslant \int_x^{\lambda x} \frac{\lambda x - t}{t^2} dt &= \int_1^\lambda \frac{\lambda x - \nu x}{(\nu x)^2} x d\nu \\ &= \int_1^\lambda \frac{\lambda - \nu}{\nu^2} d\nu \\ &= c(\lambda) \end{aligned}$$

que es una constante positiva que depende únicamente de λ. Por lo tanto

$$\left|\int_x^\infty \frac{A(t) - t}{t^2} dt - \int_{\lambda x}^\infty \frac{A(t) - t}{t^2} dt\right| = c(\lambda),$$

sin embargo, las integrales anteriores son colas de una integral convergente, por lo que ambas convergen a cero cuando $x \to \infty$, esto contradice que $c(\lambda) > 0$.

□

Con esto ya podemos presentar una prueba del teorema de Wiener-Ikehara.

Demostración. *Teorema 3.1.* Tenemos, para $\Re(s) > 1$ que

$$F(s) = \sum_{n=1}^{\infty} a_n n^{-s} = s \int_1^{\infty} \frac{A(t)}{t^{s+1}} dt,$$

Esto implica que en el semiplano $\sigma > 1$, $\sum_{n=1}^{\infty} a_n n^{-s}$ converge absolutamente, $F(s)$ es analítica y

$$F(s) - \frac{s}{s-1} = s \int_1^{\infty} \frac{A(t) - t}{t^{s+1}} dt.$$

Cambiando s por $s + 1$ y $t = e^u$ en la integral, tenemos para $\Re(s) > 0$,

$$F(s+1) - 1 - \frac{1}{s} = (s+1) \int_0^{\infty} \frac{A\left(e^u\right) - e^u}{e^u} e^{-us} du.$$

Por hipótesis $A\left(e^u\right) = O\left(e^u\right)$, se sigue que la función f dada por

$$f(u) = \frac{A\left(e^u\right) - e^u}{e^u}$$

es acotada en $[0, \infty)$ e integrable en cada subintervalo cerrado y acotado de $[0, \infty)$, y

$$\frac{F(s+1) - 1 - \frac{1}{s}}{s+1} = \int_0^{\infty} f(u) e^{-su} du. \tag{3.3}$$

El lado izquierdo en (3.3) es analítico para $\sigma > 0$. Como $F(s+1)$ tiene un polo simple en $s = 0$, vemos que

$$F(s+1) - \frac{1}{s}$$

es analítica en $s = 0$, y para todo s con $\sigma \geqslant 0$. Por lo tanto, podemos aplicar el teorema 3.4 para obtener que la integral

$$\int_0^{\infty} \frac{A\left(e^u\right) - e^u}{e^u} du = \int_1^{\infty} \frac{A(t) - t}{t^2} dt$$

converge. Por el lema anterior, obtenemos que $A(x) \sim x$.

□

Nota. En esta prueba del teorema de Wiener-Ikehara, la idea principal fue la introducción de un kernel en la integral $I_{\mathcal{C}}$, a saber,

$$\left(1 + \frac{s^2}{R^2}\right) \frac{1}{s}.$$

Esta función tiene la característica de tener un polo en $s = 0$, se sigue que

$$I_{\mathcal{C}} = \frac{1}{2\pi i} \int_{\mathcal{C}} \left(g(s) - g_T(s)\right) e^{sT} \left(1 + \frac{s^2}{R^2}\right) \frac{1}{s} ds = g(0) - g_T(0),$$

esta expresión nos permite ver que $\lim_{T\to\infty} g_T(0) = g(0)$ acotando de manera adecuada los términos en la integral y tomando el límite.

Si revisamos pruebas clásicas del teorema, por ejemplo la que se encuentra en [16], notaremos que aquí esta función juega un papel similar al del kernel de Fejér, las ideas subyacentes de las pruebas son muy similares ya que si tomamos $s = \sigma + it$, la expresión (3.3) toma la forma de una trasformada de Fourier.

Con este teorema, podemos obtener una prueba del TNP, sin embargo, en progresiones aritmética la serie de Dirichlet

$$L(\chi, s) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$$

no cumple la hipótesis $\chi(n) \geqslant 0$ del teorema de Wiener-Ikehara, también sabemos que el teorema de los números primos nos dice que

$$\pi(a, q, x) \sim \frac{x}{\varphi(q)\log x},$$

por lo que esperamos aplicar el teorema de Wiener-Ikehara a una serie de Dirichlet con residuo $1/\varphi(q)$, necesitamos extender el resultado para un conjunto más grande de series de Dirichlet, como mencionamos antes, haremos esto siguiendo a [15], estas extensiones del teorema tauberiano son también tratadas en el texto de Ram Murty [16]

Corolario 3.6. Sean $a_n \geqslant 0$ y $F(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ una serie absolutamente convergente. Supongamos que se cumplen las siguientes condiciones:

a) La función $F(s)$ se extiende a una función analítica en la región $\Re(s) \geqslant 1$ con un único polo simple en $s = 1$, cuyo residuo es R.

b) $A(x) = \sum_{n\leqslant x} a_n = O(x)$.

Entonces, se tiene que

$$A(x) = Rx + o(x) \text{ cuando } x \to \infty.$$

Demostración. Note que basta considerar $R > 0$, ya que si $R \leqslant 0$ basta probar el teorema para

$$F(s) + m\zeta(s) = \sum_{n=1}^{\infty} \frac{a_n + m}{n^s}$$

con $m \in \mathbb{Z}$ tal que $m > |R|$. Para $R > 0$ note que cambiando a_n por $\frac{a_n}{R}$ y aplicando el teorema 3.1 obtenemos lo deseado.

□

Finalizamos esta sección con el siguiente corolario

Corolario 3.7. Sea $F(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ una serie de Dirichlet con coeficientes complejos. Sea $A(x)$ la suma parcial de los coeficientes. Supongamos que existe una serie de Dirichlet $G(s) = \sum_{n=1}^{\infty} \frac{b_n}{n^s}$ con coeficientes no negativos, tal que:

(a) $|a_n| \leqslant b_n$ para todo $n \in \mathbb{N}$.

(b) $G(s)$ es absolutamente convergente para $\Re(s) > 1$.

(c) La función $G(s)$ (respectivamente $F(s)$) tiene una extensión analítica en la región $\Re(s) \geqslant 1$, sin polos excepto por un polo simple en $s = 1$ con residuo R (respectivamente r).

(d) $B(x) := \sum_{n \leqslant x} b_n = O(x)$.

Entonces, cuando $x \to \infty$, $A(x) = rx + o(x)$.

Demostración. Si los coeficientes a_n son reales, consideramos la serie $G(s) - F(s)$, que tiene coeficientes no negativos y satisface las condiciones del corolario anterior, obteniendo

$$\sum_{n \leqslant x} (b_n - a_n) = (R - r)x + o(x)$$

cuando $x \to \infty$. Dado que $B(x) = Rx + o(x)$, esto prueba el resultado en el caso de coeficientes reales.
Si los coeficientes a_n no son reales, definimos

$$F^*(s) = \sum_{n=1}^{\infty} \frac{\overline{a_n}}{n^s},$$

de modo que

$$F = \frac{F + F^*}{2} + i\left(\frac{F - F^*}{2i}\right),$$

y aplicamos el resultado para coeficientes reales separadamente a la parte real e imaginaria de la expresión anterior. □

3.2 Prueba del teorema de los números primos

Para probar el TNP queremos ver que $\psi(x) \sim x$, en este punto, el camino a seguir es claro, necesitamos probar que la serie de Dirichlet de la función de Von Mangolth cumple las condiciones del teorema de Wiener-Ikehara.

Primero vamos a verificar que en efecto $\psi(x) = O(x)$, en el capítulo de premilinares vimos que $\psi(x) - \vartheta(x) = O(\sqrt{x}\log^2 x)$, por lo tanto, basta ver que en efecto $\vartheta(x) = O(x)$. Dado $n \in \mathbb{N}$, note que

$$2^{2n} = (1+1)^{2n} = \binom{2n}{0} + \ldots + \binom{2n}{2n} \geqslant \binom{2n}{n},$$

todo número primo $n < p < 2n$ divide a $(2n)!$ y no a $(n!)^2$, de lo que obtenemos

$$2^{2n} = (1+1)^{2n} = \binom{2n}{0} + \ldots + \binom{2n}{2n} \geqslant \binom{2n}{n} \geqslant \prod_{n<p<2n} p = e^{\vartheta(2n) - \vartheta(n)},$$

tomando el logaritmo en los extremos, se sigue que $2n\log 2 \geqslant \vartheta(2n) - \vartheta(n)$, podemos reescribrir esto como

$$\vartheta(x) - \vartheta\left(\frac{x}{2}\right) \leqslant x\log 2,$$

entonces

$$\begin{aligned} \vartheta(x) - \vartheta\left(\frac{x}{2}\right) &\leqslant x\log 2, \\ \vartheta\left(\frac{x}{2}\right) - \vartheta\left(\frac{x}{4}\right) &\leqslant \frac{x}{2}\log 2, \\ &\vdots \\ \vartheta\left(\frac{x}{2^r}\right) - 0 &\leqslant \frac{x}{2^r}\log 2, \end{aligned} \tag{3.4}$$

donde r es el mayor entero tal que $x > 2^r$, sumando las desigualdades en (3.4), obtenemos que

$$\vartheta(x) \leqslant x\log 2\left(\sum_{i=0}^{r}\frac{1}{2^i}\right) \leqslant x\log(4)\left(1 - \frac{1}{2^{r+1}}\right) \leqslant x\log 4,$$

con lo que hemos probado que $\vartheta(x) = O(x)$.

Teorema 3.8 (Principio del argumento)**.** Sea $f(z)$ una función analítica,

- si z_0 es un polo de $f(z)$ de orden m, entonces z_0 es un polo simple de $\dfrac{f'(z)}{f(z)}$ con residuo $-m$,
- si z_0 es un cero de $f(z)$ de orden m, entonces z_0 es un polo simple de $\dfrac{f'(z)}{f(z)}$ con residuo m.

Demostración. Sea z_0 un cero de f. Podemos escribir

$$f(z) = (z - z_0)^m\, g(z),$$

donde m es la multiplicidad del cero y por lo tanto $g(z_0) \neq 0$, tenemos que

$$f'(z) = m(z-z_0)^{m-1} g(z) + (z-z_0)^m g'(z),$$

y

$$\frac{f'(z)}{f(z)} = \frac{m}{z-z_0} + \frac{g'(z)}{g(z)}.$$

Dado que $g(z_0) \neq 0$, se sigue que $\dfrac{g'(z)}{g(z)}$ no tiene singularidades en z_0, por lo tanto es analítica en z_0. Esto implica que el residuo de $\dfrac{f'(z)}{f(z)}$ en z_0 es m.

Análogamente, si z_0 es un polo de f,

$$f(z) = (z-z_0)^{-m} h(z),$$

donde m es el orden del polo, y $h(z_0) \neq 0$. Entonces,

$$f'(z) = -m(z-z_0)^{-m-1} h(z) + (z-z_0)^{-m} h'(z),$$

y

$$\frac{f'(z)}{f(z)} = \frac{-m}{z-z_0} + \frac{h'(z)}{h(z)}.$$

□

Recordemos que

$$F(s) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)},$$

converge absolutamente en $\Re(s) > 1$ (corolario 1.30), entonces el principio del argumento nos dice que $F(s)$ tiene un polo simple en $s = 1$ con residuo 1, dado por el polo que tiene $\zeta(s)$ en $s = 1$.

Para garantizar que $F(s)$ tiene una extensión analítica a $\Re(s) \geqslant 1$, excepto por un **único** polo simple en $s = 1$ con residuo 1, debemos ver que $\zeta(s) \neq 0$ en la recta $\Re(s) = 1$, este hecho junto con el teorema de Wiener-Ikehara implican inmediatamente el teorema de los números primos.

Lema 3.9. Si $\theta \in \mathbb{R}$, entonces $3 + 4\cos\theta + \cos 2\theta \geqslant 0$

Demostración. En efecto

$$\begin{aligned} 0 &\leqslant 2(1+\cos\theta)^2 \\ &= 2(1 + 2\cos\theta + \cos^2\theta) \\ &= 2 + 4\cos\theta + 1 + \cos(2\theta). \end{aligned}$$

□

Teorema 3.10. $\zeta(1+it) \neq 0$ para todo $t \neq 0$.

Demostración. Note que

$$\begin{aligned} p^{-n\sigma}\cos(nt\log p) &= p^{-n\sigma}\cos(-nt\log p) \\ &= p^{-n\sigma}\Re(exp(i\log p^{-nt})) \\ &= \Re(p^{-n\sigma}p^{-int}) \\ &= \Re(p^{-ns}), \end{aligned}$$

por lo tanto

$$\begin{aligned} \Re(\log\zeta(s)) &= \sum_p \sum_{n=1}^{\infty} \frac{\Re(p^{-ns})}{n} \\ &= \sum_p \sum_{n=1}^{\infty} \frac{p^{-n\sigma}}{n}\cos(nt\log p). \end{aligned}$$

Ahora, como $\Re(\log z) = \log|z|$, entonces por el lema 3.9

$$\begin{aligned} 0 &\leqslant \sum_p \sum_{n=1}^{\infty} \frac{p^{-n\sigma}}{n}(3+4\cos(nt\log p)+\cos(2nt\log p)) \\ &= \Re(3\log\zeta(\sigma)+4\log\zeta(\sigma+it)+\log\zeta(\sigma+2it)) \\ &= 3\log|\zeta(\sigma)|+4\log|\zeta(\sigma+it)|+\log|\zeta(\sigma+2it)|, \end{aligned}$$

esto es

$$|\zeta(\sigma)|^3|\zeta(\sigma+it)|^4|\zeta(\sigma+2it)| \geqslant 1,$$

obtenemos la identidad

$$|(\sigma-1)\zeta(\sigma)|^3\left|\frac{\zeta(\sigma+it)}{\sigma-1}\right|^4|\zeta(\sigma+2it)| \geqslant \frac{1}{\sigma-1}. \tag{3.5}$$

Supongamos que $\zeta(1+it) = 0$ para algún $t \neq 0$, entonces

$$\lim_{\sigma\to 1^+}\frac{\zeta(\sigma+it)-\zeta(1+it)}{\sigma-1} = \lim_{\sigma\to 1^+}\frac{\zeta(\sigma+it)}{\sigma-1} = \zeta'(1+it),$$

como $\zeta(s)$ tiene un polo simple en $s = 1$, se sigue que

$$\lim_{\sigma\to 1^+}(\sigma-1)\zeta(\sigma) = 1,$$

esto muestra que

$$\lim_{\sigma\to 1^+}|(\sigma-1)\zeta(\sigma)|^3\left|\frac{\zeta(\sigma+it)}{\sigma-1}\right|^4|\zeta(\sigma+2it)|$$

existe y es igual a $|\zeta'(1+it)|^4|\zeta(1+2it)|$, lo que contradice (3.5). □

Corolario 3.11 (Teorema de los números primos)**.**

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

Demostración. Sea $F(s) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}$, como $F(s)$ tiene un único polo simple en $s = 1$ con residuo 1, aplicando el teorema de Wiener-Ikehara se sigue que

$$\psi(x) = \sum_{n \leqslant x} \Lambda(n) = x + o(x),$$

esto es

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

□

Nota. Una aplicación elemental de la regla de L´Hopital es la siguiente:

$$\lim_{x\to\infty} \frac{\int_2^x \frac{1}{\log t} dt}{\frac{x}{\log x}} = \lim_{x\to\infty} \frac{\frac{1}{\log x}}{\frac{\log x - x^{-1}}{\log^2 x}} = \lim_{x\to\infty} \frac{1}{1 - (x \log x)^{-1}} = 1.$$

Esta presentación del teorema de los números primos, a saber,

$$\pi(x) \sim \int_2^x \frac{1}{\log t} dt,$$

nos dice que integrar la densidad de los primos nos da el números de primos menores que x.

Terminaremos esta sección abordando el discurso de Hardy sobre si es posible una prueba elemental del teorema de los números primos

> *No se conoce ninguna prueba elemental del teorema de los números primos, y uno puede preguntarse si es razonable esperar una. Ahora sabemos que el teorema es, en términos generales, equivalente a un teorema sobre una función analítica, el teorema de que la función zeta de Riemann no tiene raíces en cierta línea. Una prueba de tal teorema, que no dependa fundamentalmente de la teoría de funciones, me parece extraordinariamente improbable*
>
> **— G. H. Hardy.**

Las razones de Hardy se esclarecen con la prueba que hemos visto en este capítulo, la no nulidad de $\zeta(1 + it)$ es una condición suficiente para el TNP, ver que es una condición necesaria no es una tarea complicada y era ya conocido por Hadamard y de La Vallée Poussin.

Tenemos que

$$\int_1^\infty \frac{\psi(x)}{x^{s+1}}dx = -\frac{1}{s}\frac{\zeta'(s)}{\zeta(s)},$$

podemos escribir a $\zeta(s)$ como $\zeta(s) = (s-1)^{-1}h(s)$, en efecto

$$\begin{aligned}\int_1^\infty \frac{\psi(x)-x}{x^{s+1}}dx &= -\frac{1}{s}\left(\frac{h'(s)(s-1)^{-1} - (s-1)^{-2}h(s)}{(s-1)^{-1}h(s)}\right) - \frac{1}{s-1}\\ &= -\frac{1}{s}\left(\frac{h'(s)}{h(s)} + 1\right),\end{aligned}$$

por el principio del argumento, dado $\tau \neq 0$, si $\zeta(1+i\tau) = 0$, entonces

$$\lim_{\sigma\to 1}(\sigma-1)\frac{h'(\sigma+i\tau)}{h(\sigma+i\tau)} = m,$$

donde m es la multiplicidad del 0. Se sigue que

$$\lim_{\sigma\to 1}(\sigma-1)\int_1^\infty \frac{\psi(x)-x}{x^{\sigma+1+i\tau}}dx = -\frac{m}{\sigma+\tau i} \neq 0, \tag{3.6}$$

por otro lado si $\psi(x) \sim x$,

$$\left|\int_1^\infty \frac{\psi(x)-x}{x^{s+1}}dx\right| \leqslant \int_1^\infty \frac{|\psi(x)-x|}{x^{\sigma+1}}dx \leqslant k.$$

Esto muestra que

$$\lim_{\sigma\to 1}(\sigma-1)\int_1^\infty \frac{\psi(x)-x}{x^{\sigma+1+i\tau}}dx = 0,$$

lo que contradice (3.6).

Corolario 3.12. $\zeta(1+it) \neq 0$ para todo $t \neq 0$ si y solamente si $\pi(x) \sim x$.

En 1950 el matemático Atle Selberg ganó la medalla fields por presentar una prueba del TNP en la que no se requiere variable compleja, esta fue la primera prueba elemental, cuya idea principal se basa en explotar la identidad

$$\vartheta(x)\log(x) + \sum_{p\leqslant x}\log(p)\vartheta\left(\frac{x}{p}\right) = 2x\log(x) + O(x),$$

no abordaremos esta prueba aquí, el argumento es bastante indirecto y extenso, con varias pasos intermedios suavizando funciones, para ver en detalle la prueba se puede consultar [17].

Nota. La desventaja de abordar la prueba del TNP desde la teoría tauberiana es que no controlamos el error, solo obtenemos un término o pequeña y sabemos que las estimaciones O grande son mejores, sin embargo obtuvimos algo muy fuerte, una equivalencia.

Algunas versiones del TNP con término de error son las siguientes.

Autores	**Cota para** $\psi(x) - x$	**Región de no nulidad** $t^* = \text{máx}(\lvert t \rvert, 3)$
Chebyshev (1851)	cx	
Hadamard, Poussin (1896)	$O\left(xe^{-c\sqrt{\log x}}\right)$	$\sigma \geqslant 1 - \dfrac{c}{\log t^*}$
Littlewood (1922)	$O\left(xe^{-c\sqrt{\log x \log\log x}}\right)$	$\sigma \geqslant 1 - \dfrac{c \log\log t^*}{\log t^*}$
Wiener-Ikehara (1931)	$o(x)$	$\sigma \geqslant 1$
Hipótesis de Riemann	$O_\varepsilon(x^{1/2+\varepsilon}), \varepsilon > 0$	$\sigma > 1/2$

Esto permite ver la relación que hay entre ceros de la función zeta de Riemann y el término de error del teorema.

Para finalizar este trabajo estudiaremos como se extiende este teorema a las progresiones aritmética en el siguiente capítulo.

Primos en progresiones aritméti-ca

> *Los encantos de esta ciencia sublime, las matemáticas, solo se le revelan a aquellos que tienen el valor de profundizar en ella.*
>
> **— Carl Friedrich Gauss**

El teorema de Dirichlet nos dice que hay infinitos primos de la forma $a, a+n, a+2n, \ldots$ siempre que $(a, n) = 1$, este es un resultado importante en teoría de números sin embargo podemos decir mucho más sobre primos en progresiones aritmética. Hemos probado el teorema de los números primos en el capítulo anterior y una pregunta natural es cómo se extiende este a progresiones aritmética, por ejemplo, si estudiamos los primos de la forma $4k+1$ y $4k+3$, ¿hay la misma cantidad de primos de la forma $4k+1$ que $4k+3$?, ¿podemos esperar una distribución uniforme de estos?, la respuesta nos la dará el TNP, este nos dice que

$$\pi(1,4,x) \sim \frac{x}{\varphi(4)\log x} = \frac{x}{2\log x}, \quad \pi(3,4,x) \sim \frac{x}{2\log x}.$$

Más precisamente $(\mathbb{Z}/4\mathbb{Z})^\times = \{\overline{1}, \overline{3}\}$, tenemos dos clases generadoras de primos y en cada una tenemos la mitad, este resultado no es un hecho trivial, por mucho tiempo los matemáticos pensaron que este no era el caso, la evidencia heurística apuntaba a que habían más primos de la forma $4k+3$ hasta que Littlewood demostró que $\pi(3,4,x) - \pi(1,4,x)$ tiene infinitos cambios de signo, estos temas se abordan de manera más detallada en [18].

4.1 Teorema de los números primos en progresiones aritmética

Como antes, la idea es aplicar el teorema de Wiener-Ikehara para probar el TNP, por lo que presentamos antes este teorema sobre no nulidad de L−funciones, análogo al que tenemos para $\zeta(s)$

Teorema 4.1. Sea χ un carácter de Dirichlet, entonces

$$L(\chi, s) \neq 0, \text{ si } \Re(s) > 1$$

Demostración. Como $\Re(s) > 1$, entonces

$$L(\chi, s) = \prod_p \frac{1}{1 - \chi(p)p^{-s}},$$

de manera análoga al teorema 1.38, el término $\frac{1}{1-\chi(p)p^{-s}} = \frac{p^s}{p^s-\chi(p)} \neq 0$ para todo p, esto nos da que el producto no tiene factores nulos y por tanto no converge a 0. □

Teorema 4.2 (Teorema de los números primos en progresiones aritmética)**.** Dados a y m primos relativos.

$$\pi(a, m, x) \sim \frac{x}{\varphi(m)\log x}$$

Observemos lo siguiente, si h es una función aritmética completamente multiplicativa

$$\begin{aligned}((f*g)h)(n) &= \sum_{j|n} f(j)g\left(\frac{n}{j}\right)h(n) = \sum_{j|n} f(j)g\left(\frac{n}{j}\right)h(j)\,h\left(\frac{n}{j}\right)\\ &= \sum_{j|n} f(j)h(j)\,g\left(\frac{n}{j}\right)h\left(\frac{n}{j}\right)\\ &= \sum_{j|n} fh(j)gh\left(\frac{n}{j}\right)\\ &= (fh*gh)(n).\end{aligned}$$

Podemos aplicar esto para obtener una fórmula para la derivada logarítmica de la función $L(\chi, s)$

$$\begin{aligned}-L'(\chi,s) &= \sum_{n=1}^{\infty}\frac{\chi(n)\log(n)}{n^s} = \sum_{n=1}^{\infty}\frac{\chi(n)(\Lambda*1)(n)}{n^s} = \sum_{n=1}^{\infty}\frac{(\chi\Lambda*1\chi)(n)}{n^s}\\ &= \sum_{n=1}^{\infty}\frac{\chi(n)\Lambda(n)}{n^s}\sum_{n=1}^{\infty}\frac{\chi(n)}{n^s},\end{aligned}$$

esto es

$$\boxed{-\frac{L'}{L}(\chi,s) = \sum_{n=1}^{\infty}\frac{\chi(n)\Lambda(n)}{n^s}}.$$

Obtendremos el TNP en progresiones aritmética de aplicar el teorema de Wiener Ikehara a la función

$$F(s) = \sum_{n\equiv a \bmod m}\frac{\Lambda(n)}{n^s},$$

pero para esto debemos primero ver qué forma tiene esta función, sea $f(n)$ la función característica de la progresión, usando su representación de Fourier obtenemos que

$$\begin{aligned} F(s) &= \sum_{n=1}^{\infty} \frac{\Lambda(n) f(n)}{n^s} = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \left(\frac{1}{\varphi(m)} \sum_{\chi} \chi(a^{-1}) \chi(n) \right) \\ &= \frac{1}{\varphi(m)} \sum_{\chi} \chi(a^{-1}) \left(\sum_{n=1}^{\infty} \frac{\chi(n)\Lambda(n)}{n^s} \right) \\ &= \frac{1}{\varphi(m)} \sum_{\chi} \chi(a^{-1}) \left(-\frac{L'}{L}(\chi, s) \right). \end{aligned}$$

Sabemos que todas estas L−funciones presentes en la suma son analíticas en $\Re(s) \geqslant 1$ excepto $L(\chi_0, s)$ que tiene un polo simple en $s = 1$ dado por el polo simple que tiene $\zeta(s)$, aplicando el principio del argumento y suponiendo que ninguna de estas L funciones se anula en la recta vertical $\Re(s) = 1$ sabremos que en la suma anterior el único carácter que contribuye un polo es el trivial, este polo tendrá residuo 1, luego la función $F(s)$ cumple todas las condiciones del teorema de Wiener Ikehara y además tiene residuo $\dfrac{1}{\varphi(m)}$, esto es

$$\psi(a, m, x) = \sum_{\substack{n \leqslant x \\ n \equiv a \text{ mód } m}} \Lambda(n) \sim \frac{x}{\varphi(m)}.$$

Repasando lo anterior, las derivadas logarítmicas son todas analíticas en $\Re(s) \geqslant 1$ por el principio del argumento, solo una tiene un polo en $s = 1$, la del carácter trivial y esta nos da el residuo que esperamos. El teorema de los números primos es una consecuencia inmediata de la siguiente afirmación

Teorema 4.3. Sea χ un carácter de Dirichlet, $L(\chi, 1 + it) \neq 0$ para todo t

La prueba de este teorema no es sencilla, recordemos que para $\zeta(s)$ no lo fue, requería de ciertas estimaciones que no son evidentes, en general estudiar las regiones de no nulidad de L−funciones es un problema profundo en teoría analítica de números que además permanece abierto. La prueba que veremos aquí es la propuesta en [16].

4.1.1. La no nulidad de $L(1 + it, \chi)$

Como en el capítulo anterior, probar esta no nulidad requiere de identidades trigonométricas, ya que al tomarle parte real a una serie de Dirichlet nos aparecen cosenos de manera natural, sin embargo, cuando vimos que $L(1, \chi) \neq 0$ para $\chi \neq \chi_0$ consideramos un producto de L−funciones y realizamos un argumento de conteo de ceros y polos. En este caso también vamos a considerar el producto de todas las L-funciones, veremos que

$$f(s) = \prod_{\chi} L(s, \chi)$$

no se anula en la recta vertical $\Re(s) = 1$ y por lo tanto ninguna L−función, para esto aplicaremos un argumento que proviene de Landau y que es modificado en [16], el estudio de

las series de Dirichlet con coeficientes no negativos, por lo que estudiaremos la expresión $\log(f(s))$.

Finalmente, como estamos trabajando con $\log(f(s))$, el logaritmo transforma este producto de L−funciones en una suma y vamos a necesitar identidades sobre suma de cosenos, particularmente la identidad de Lagrange para el núcleo de Dirichlet.

Teorema 4.4. Para $0 < \theta < 2\pi$.

$$\frac{1}{2} + \cos\theta + \cos(2\theta) + \cdots + \cos(n\theta) = \frac{\sin\left((n+\frac{1}{2})\theta\right)}{2\sin\left(\frac{\theta}{2}\right)},$$

Demostración. Note que si tomamos $|z| = 1$ con $z \neq 1$, entonces $z = exp(i\theta)$, de esto se sigue que

$$\Re\left(\sum_{k=0}^{n} z^k\right) = \sum_{k=0}^{n} \Re(z^k) = \sum_{k=0}^{n} \cos(k\theta).$$

En efecto

$$\begin{aligned}
\Re\left(\sum_{k=0}^{n} z^k\right) &= \Re\left(\frac{z^{n+1}-1}{z-1}\right) \\
&= \Re\left(\frac{exp\left(i(n+1)\theta\right)-1}{exp(i\theta)-1}\right) \\
&= \Re\left(\frac{exp(i(n+1)\theta/2)}{exp(i\theta/2)}\frac{exp(-i(n+1)\theta/2)-exp(i(n+1)\theta/2)}{exp(-i\theta/2)-exp(i\theta/2)}\right).
\end{aligned}$$

Note que el último término de la derecha de puede escribir como

$$\Re\left(exp(in\theta/2)\frac{\sin\left((n+1)\theta/2\right)}{\sin(\theta/2)}\right) = \cos(n\theta/2)\frac{\sin((n+1)\theta/2)}{\sin(\theta/2)},$$

además

$$\begin{aligned}
\cos\left(\frac{n\theta}{2}\right)\sin\left(\frac{(n+1)\theta}{2}\right) &= \frac{1}{2}\left(\sin\left(\frac{\theta n}{2}+\frac{\theta(n+1)}{2}\right)+\sin\left(\frac{\theta(n+1)}{2}-\frac{n\theta}{2}\right)\right) \\
&= \frac{1}{2}\left(\sin\left(\frac{\theta(2n+1)}{2}\right)+\sin\left(\frac{\theta}{2}\right)\right) \\
&= \frac{1}{2}\sin\left(\theta\left(n+\frac{1}{2}\right)\right)+\frac{1}{2}\sin\left(\frac{\theta}{2}\right),
\end{aligned}$$

dividiendo entre $\sin\left(\frac{\theta}{2}\right)$ obtenemos que

$$1 + \cos\theta + \cos(2\theta) + \cdots + \cos(n\theta) = \frac{1}{2} + \frac{\sin\left((n+\frac{1}{2})\theta\right)}{2\sin\left(\frac{\theta}{2}\right)},$$

de lo que se sigue el resultado. □

Lema 4.5.

$$1+\frac{\sin 3\theta}{\sin\theta}+\frac{\sin 5\theta}{\sin\theta}+\cdots+\frac{\sin(2n-1)\theta}{\sin\theta}=\left(\frac{\sin n\theta}{\sin\theta}\right)^2$$

Omitimos la prueba aquí ya que el resultado se sigue por inducción y se puede consultar en [16].

Teorema 4.6. Para todo entero $m \geqslant 0$,

$$(2m+1)+2\sum_{j=0}^{2m-1}(j+1)\cos(2m-j)\theta=\left(\frac{\sin\left(m+\frac{1}{2}\right)\theta}{\sin\frac{\theta}{2}}\right)^2.$$

Demostración. Cambiando el orden de sumación, podemos reescribir la identidad como sigue

$$2m+1+2\sum_{j=1}^{2m}(2m-j+1)\cos j\theta=\left(\frac{\sin\left(m+\frac{1}{2}\right)\theta}{\sin\frac{\theta}{2}}\right)^2.$$

Tomando $\theta = 2\varphi$, debemos probar que

$$2m+1+2\sum_{j=1}^{2m}(2m-j+1)\cos 2j\varphi=\left(\frac{\sin(2m+1)\varphi}{\sin\varphi}\right)^2,$$

por el teorema 4.4

$$1+2\sum_{j=1}^{n}\cos 2j\varphi=\frac{\sin(2n+1)\varphi}{\sin\varphi}.$$

Ahora note que

$$\begin{aligned}\sum_{n=0}^{2m}(1+2\sum_{j=1}^{n}\cos 2j\varphi) &= (2m+1)+2\sum_{n=0}^{2m}\sum_{j=1}^{n}\cos 2j\varphi\\ &= \sum_{n=0}^{2m}\frac{\sin(2n+1)\varphi}{\sin\varphi}.\end{aligned}$$

De esto se sigue que,

$$
\begin{aligned}
(2m+1)+2\sum_{n=0}^{2m}\sum_{j=1}^{n}\cos 2j\varphi &= (2m+1)+2\sum_{j=1}^{2m}\cos 2j\varphi \sum_{j\leqslant n\leqslant 2m} 1 \\
&= (2m+1)+2\sum_{j=1}^{2m}(2m-j+1)\cos 2j\varphi \\
&= \sum_{n=0}^{2m}\frac{\sin(2n+1)\varphi}{\sin\varphi} \\
&= \left(\frac{\sin(2m+1)\varphi}{\sin\varphi}\right)^2 .
\end{aligned}
$$

□

Teorema 4.7 (K. Murty)**.** Sea $f(s)$ una función a valor complejo que satisface las siguientes condiciones

- f es holomorfica y no nula en $\Re(s) > 1$,
- $\log f(s)$ se puede escribir como una serie de Dirichlet

$$\sum_{n=1}^{\infty}\frac{b_n}{n^s}$$

con $b_n \geqslant 0$ para $\Re(s) > 1$;

- en la recta $\Re(s) = 1$, f es holomorfica excepto por un polo de orden $e \geqslant 0$ en $s = 1$.

Si f tiene un cero en $\Re(s) = 1$, entonces el orden del cero está acotado por $e/2$.

Demostración. Supongamos que f tiene un cero en $1 + it_0$ de orden $k > e/2$,entonces $e \leqslant 2k-1$, considere la función

$$
\begin{aligned}
g(s) &= f(s)^{2k+1}\prod_{j=1}^{2k} f\left(s+ijt_0\right)^{2(2k+1-j)} \\
&= f(s)^{2k+1} f\left(s+it_0\right)^{4k} f\left(s+2it_0\right)^{4k-2}\cdots f\left(s+2kit_0\right)^2 ,
\end{aligned}
$$

entonces $g(s)$ es holomorfica en $\Re(s) > 1$ y tiene un cero de al menos orden 1 en $s = 1$ ya que

$$4k^2-(2k+1)e \geqslant 4k^2-(2k+1)(2k-1) = 1$$

donde k es el orden del cero y e el orden del polo. Por otro lado, para $\Re(s) > 1$,

$$\log g(s) = \sum_{n=1}^{\infty}\frac{b_n}{n^s}\left(2k+1+2\sum_{j=1}^{2n}2(2k+1-j)n^{-ijt_0}\right).$$

Sea $\theta = t_0 \log n$. Entonces para $s = \sigma > 1$,

$$\Re(\log g(\sigma)) = \log|g(\sigma)| = \sum_{n=1}^{\infty} \frac{b_n}{n^{\sigma}} \left(2k + 1 + \sum_{j=1}^{2k} 2(2k+1-j)\cos j\theta \right),$$

por el teorema anterior, el término del paréntesis es no negativo, de esto se sigue que $|g(\sigma)| \geqslant 1$, una contradicción ya que cuando $\sigma \to 1$, $g(s) = 0$. □

Teorema 4.8. $L(s,\chi) \neq 0$ para todo s con $\Re(s) = 1$.

Demostración. Sea $f(s) = \prod_{\chi} L(s,\chi)$ con χ un carácter de Dirichlet módulo q, claramente $f(s)$ es holomorfica y no nula en $\Re(s) > 1$, además

$$\log f(s) = \sum_{\chi} \log L(s,\chi) = \sum_{n,p} \frac{1}{np^{ns}} \sum_{\chi} \chi\left(p^n\right).$$

Note que

$$f(p^n) = \begin{cases} 1 & \text{si } p^n \equiv 1 \text{ mód } q \\ 0 & \text{e.o.c.} \end{cases} = \frac{1}{\varphi(q)} \sum_{\chi} \chi(p^n)\overline{\chi(1)}$$
$$= \frac{1}{\varphi(q)} \sum_{\chi} \chi(p^n),$$

de esto se sigue que

$$\log(f(s)) = \varphi(q) \sum_{\substack{n,p \\ p^n \equiv 1 (\text{ mód } q)}} \frac{1}{np^{ns}},$$

luego $\log(f(s))$ es una serie de Dirichlet con coeficientes no negativos. En la recta $\Re(s) = 1$, f es holomorfica excepto por un polo simple en $s = 1$, a saber $L(1,\chi_0)$. Por el teorema 4.7, si $f(s)$ tiene un cero en $\Re(s) = 1$ entonces el orden del cero está acotado por $1/2$, esto es que $f(s)$ no se anula en $\Re(s) = 1$ y por lo tanto $L(s,\chi)$. □

Corolario 4.9. Dados a y q primos relativos,

$$\psi(a,q,x) \sim \frac{x}{\varphi(q)}.$$

Como en el teorema de los números primos, es posible obtener mejores cotas para el error estudiando regiones donde $L(s,\chi) \neq 0$

Teorema 4.10 (Siegel–Walfisz)**.** Dados $N \in \mathbb{R}$ y $(a,q) = 1$, existe C_N tal que

$$\psi(a,q,x) = \frac{x}{\varphi(q)} + O\left(x \exp\left(-C_N (\log x)^{\frac{1}{2}}\right)\right)$$

si $q \leqslant (\log x)^N$.

Sin embargo es algo que va más allá de los límites de este, ya muy largo trabajo.

Bibliografía

[1] Bernhard Riemann. On the number of primes less than a given magnitude. *Complete Works. Kendrick Press*, 2004.

[2] Adolf J Hildebrand. Introduction to analytic number theory math 531 lecture notes, fall 2005. *URL: http://www. math. uiuc. edu/hildebr/ant. Version*, 1, 2006.

[3] Norman Levinson. A motivated account of an elementary proof of the prime number theorem. *The American Mathematical Monthly*, 76(3):225–245, 1969.

[4] Tom M Apostol. *Mathematical analysis; 2nd ed.* Addison-Wesley series in mathematics. Addison-Wesley, Reading, MA, 1974.

[5] Prapanpong Pongsriiam. *Analytic Number Theory for Beginners*, volume 103. American Mathematical Society, 2023.

[6] E.M. Stein and R. Shakarchi. *Complex Analysis*. Princeton lectures in analysis. Princeton University Press, 2010.

[7] Hugh L Montgomery and Robert C Vaughan. *Multiplicative number theory I: Classical theory*. Number 97. Cambridge university press, 2007.

[8] T.M. Apostol. *Introduction to Analytic Number Theory*. Undergraduate Texts in Mathematics. Springer New York, 1998.

[9] Don Zagier. Newman's short proof of the prime number theorem. *The American mathematical monthly*, 104(8):705–708, 1997.

[10] Donald J Newman. Simple analytic proof of the prime number theorem. *The American Mathematical Monthly*, 87(9):693–696, 1980.

[11] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53. American Mathematical Soc., 2021.

[12] Paul T Bateman and Harold G Diamond. A hundred years of prime numbers. *The American mathematical monthly*, 103(9):729–741, 1996.

[13] Jacob Korevaar. On newman's quick way to the prime number theorem. *The Mathematical Intelligencer*, 4(3):108–115, 1982.

[14] Jaap Korevaar. The wiener–ikehara theorem by complex analysis. *Proceedings of the American Mathematical Society*, 134(4):1107–1116, 2006.

[15] Akshaa Vatwani. A simple proof of the wiener–ikehara tauberian theorem. *Math. Student*, 84(3-4):127–134, 2015.

[16] USR Murty. *Problems in analytic number theory*, volume 206. Springer Science & Business Media, 2007.

[17] Graham James Oscar Jameson. *The prime number theorem*. Cambridge University Press, 2003.

[18] Andrew Granville and Greg Martin. Prime number races. *The American Mathematical Monthly*, 113(1):1–33, 2006.

[19] Samuel J Patterson. *An introduction to the theory of the Riemann zeta-function*. Cambridge University Press, 1995.

[20] Harold Davenport. *Multiplicative number theory*, volume 74. Springer Science & Business Media, 2013.

[21] M Ram Murty and V Kumar Murty. *Non-vanishing of L-functions and applications*. Springer Science & Business Media, 2012.