La Función de Moebius y un Análisis Probabilístico de la Hipótesis de Riemann

Jhon Jaiver Rodríguez Herrera
Matemático

Universidad Nacional de Colombia
Facultad de Ciencias
Departamento de Matemáticas
Bogotá, D.C.
Junio de 2011

La Función de Moebius y un Análisis Probabilístico de la Hipótesis de Riemann

Jhon Jaiver Rodríguez Herrera
Matemático

Director
Jean Carlos Cortissoz Iriarte, Ph.D.
Profesor Asociado Universidad de los Andes

Codirector
Víctor Samuel Álbis Gonzaléz, Ph.D.
Profesor Titular Universidad Nacional

Universidad Nacional de Colombia
Facultad de Ciencias
Departamento de Matemáticas
Bogotá, D.C.
Junio de 2011

Título en español

La Función de Moebius y un Análisis Porbabilístico de la Hipótesis de Riemann

Title in English

The Moebius Function and a Probabilistic Analysis of the Riemann Hypothesis

Resumen: En este trabajo estudiamos las principales características de la función Zeta de Riemann; como lo son el producto de Euler, la continuación analítica y la ecuación funcional, además damos un breve panorama de la Hipótesis de Riemann mediante dos importantes equivalencias que la relacionan con la función de Moebius y el Teorema de los Números Primos. El objetivo principal del trabajo es estudiar la validez de la Hipótesis de Riemann mediante su relacion con la función de Moebius pero desde un punto de vista probabilístico y experimental. por último damos una breve explicación sobre el trabajo de Sarnak acerca de la aleatoriedad de la función μ de Moebius.

Abstract: We study the main features of the Riemann zeta function, such as are its Euler product, analytic continuation and functional equation. Also we give a brief overview of the Riemann Hypothesis by two important equivalences related to the Moebius function and the Prime Number Theorem. The main objective is to study the validity of the Riemann Hypothesis and its relationship with the Moebius function but from a probabilistic and experimental viewpoint. Finally we give a brief explanation about the work of Sarnak about the randomness of the Moebius function μ.

Palabras clave: Función zeta, función de Moebius, argumento de Denjoy, índice de Hurst.

Keywords: Zeta function, Moebius function, Denjoy argument, Hurst.

Dedicado a

A mi familia y a mis compañeros.

Agradecimientos

Agradezco principalmente a Dios por darme la fortaleza para seguir adelante cada día.

Agradezco tambíen a Elkin Quintero y Leonardo Chacón por dedicar parte de su tiempo a la discusión sobre temas particulares que permitieron la realización de este trabajo, a Daniel Bojacá y Arley Fernando Torres por acompañarme de una u otra manera en este proceso que llevo a la culminación de este trabajo.

Al profesor Víctor Albis por su tiempo dedicado a leer este trabajo.

Finalmente expreso todo mi agradecimiento al profesor Jean Carlos Cortissoz, no sólo por todo el conocimiento que me trasmitió, sino por motivarme cada vez a seguir adelante, por enseñarme el gran valor del quehacer matemático y por permitirme el honor de estudiar a su lado uno de los problemas más grandes de la historia matemática.

Índice general

Índice general **I**

Introducción **II**

1. La Función Zeta de Riemann **1**

1.1. Definición y el producto de Euler 1

1.2. Extensión analítica a todo el plano 2

1.3. Ecuación funcional 5

1.4. Acerca de las Raíces de $\zeta(s)$ 7

2. Equivalencias de la Hipótesis de Riemann **10**

3. Análisis probabilístico de la Hipótesis de Riemann **19**

3.1. Argumento de Denjoy 19

3.2. El índice de Hurst 23

3.3. Determinación de índice de Hurst 26

3.3.1. El índice de Hurst para $M(x)$ 28

4. Aleatoriedad de la función μ de Moebius **29**

A. Algoritmos **35**

A.1. Algoritmo para Moebius 35

A.2. Algoritmo de Hurst 38

Conclusiones **45**

Bibliografía **46**

Introducción

La hipótesis de Riemann ha sido, y hasta el momento es, uno de los problemas abiertos más importantes de las Matemáticas. Dicha hipótesis sugiere que todos los ceros no triviales de la función Zeta de Riemann $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ tienen parte real $\frac{1}{2}$. Además $\zeta(s)$ está fuertemente relacionada con los números primos mediante la fórmula conocida como el producto de Euler

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ primo}} \left(1 - \frac{1}{p^s}\right)^{-1}$$

A través de la historia se ha intentado mejorar el término de error en la fórmula asintótica del Teorema de los Números Primos y la mejor posible se ha obtenido asuminedo la Hipótesis de Riemann y este error es $O(\sqrt{x}\log x)$.

La función Zeta de Riemann también está relacionada con varias funciones aritméticas. Una de ellas entre otras de importancia es la función de Moebius y dicha relación es descrita por la fórmula

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

Es un hecho conocido que la hipótesis de Riemann es equivalente a que para cada $\epsilon > 0$ la función acumulativa de Möbius $M(x) = \sum_{n \leq x} \mu(n)$ decrece menos rápidamente que $x^{1/2+\epsilon}$. Este hecho sugiere investigar el comportamiento de la función $M(x)$.

Si cambiamos el punto de vista podemos ver esta función como una caminata aleatoria simple, pues $\mu(n)$ solo toma valores en el conjunto $\{0, -1, 1\}$, y nos preguntamos que tan "aleatoria", es en si esta serie, es decir si se comporta como un movimiento browniano usual (independencia en los términos de la serie). Para ello nos valdremos de un proceso más general llamado Movimiento Browniano Fraccionario el cual involucra el *Índice de Hurst*, y mediante este medimos el grado de independencia de esta serie en intervalos de tiempo disyuntos. Si este índice resulta ser 1/2 el movimiento fraccionario se reduce al movimiento browniano usual y tendríamos una evidencia númerica de la veracidad de la hipótesis de Riemann.

Por otro lado, Denjoy dio una interpretación probabilística de la Hipótesis de Riemann suponiendo que la función μ de Moebius se comporta aleatoriamente, considerando un entero n suficientemente grande sin factores cudráticos, es decir, $\mu(n) \neq 0$, y lanzando una moneda por cada uno de estos números. De esta manera el trabajó los valores de μ, y haciendo unos cálculos probabilísticos Denjoy da un argumento el cual muestra que con

probabilidad muy cercana a uno la Hipótesis de Riemann es cierta. Nosotros además de exponer este argumento lo refinaremos incluyendo las correlaciones obvias de la función μ, y obtendremos el mismo resultado pero utilizando la desigualdad de Chebyshev. Esta es una razón más para estudiar el comportamiento asíntotico de la función de Möbius, además de su “grado de aleatoriedad", para lo cual incluiremos también el punto de vista dinámico de Peter Sarnak.

CAPÍTULO 1

La Función Zeta de Riemann

Comenzamos nuestro estudio describiendo y demostrando algunas propiedades fundamentales, de lo que será la principal herramienta y a la vez el objetivo primordial de este trabajo: La función Zeta de Riemann.

La principales ideas que iniciaron el estudio de la función Zeta, fueron desarrolladas, al menos de manera intuitiva por Riemann en su artículo, escrito originalmente en alemán y traducido al inglés; *On the Number of Primes Less a Given Magnitude*, el cual fue publicado en 1859.

1.1. Definición y el producto de Euler

Tenemos en principio dos representaciones del la función zeta; una como serie y otra como producto. Ambas definiciones son equivalentes, y el punto de partida de Riemann fue la equivalencia de estas representaciones para la función $\zeta(s)$, la cual definiremos de la siguente manera,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \tag{1.1}$$

Esta suma representa una función analítica de $s = \sigma + it$, para $\sigma > 1$.

En efecto, para todo $\epsilon > 0$ y todo s tal que $\sigma \geq 1 + \epsilon$ se tiene la desigualdad

$$|\zeta(s)| \leq \sum_{n=1}^{\infty} n^{-\sigma} \leq \sum_{n=1}^{\infty} n^{-(1+\epsilon)},$$

por lo tanto la serie para $\zeta(s)$ es uniformemente convergente.

Ahora mostramos la segunda representacion de la función zeta, mediante la cual Riemann comenzó su trabajo, y que muestra la relación entre la función $\zeta(s)$ y los números primos.

Teorema 1 (Producto de Euler)**.** *Para $s = \sigma + it, \sigma > 1$, se tiene:*

$$\zeta(s) = \prod_{p \text{ primo}} \frac{1}{(1 - \frac{1}{p^s})}. \tag{1.2}$$

Demostración. Consideremos la suma geométrica $\sum_{k=1}^{\infty} \left(\frac{1}{p^s}\right)^k = \frac{1}{1-\frac{1}{p^s}}$, notemos que cada uno de los factores en (1.2) es una suma de este tipo.

Como $x \geq 2$ y como $\sigma > 1$, entonces $|\frac{1}{p^s}| < 1$, por lo que cada una de estas sumas es absolutamente convergente y así podemos multiplicar término a término cada una de ellas para obtener:

$$\prod_{p \leq x} \frac{1}{1 - \frac{1}{p^s}} = \prod_{p \leq x} \sum_{k=0}^{\infty} \frac{1}{p^{ks}} = \sum_{k_1=0}^{\infty} \cdots \sum_{k_j=0}^{\infty} \frac{1}{(p_1^{k_1} \cdots p_j^{k_j})^s}, \tag{1.3}$$

Con $2 = p_1 < p_2 < \cdots < p_j$ y $p_i < x$ para $i = 1, \ldots, j$, ninguno de estos términos se repite debido a la unicidad en la descomposición en factores primos (Teorema Fundamental de la Aritmética), por este mismo cada entero $n \leq x$ se puede descomponer usando los p_i, $i = 1, \ldots, j$ pues todos sus divisores serán menores que x también; así mismo habrá otros números mayores que x para los que sus divisores serán todos menores que x y por lo tanto la expresión (1.3) toma la forma

$$\sum_{n \leq x} \frac{1}{n^s} + \overline{\sum_{n > x}} \frac{1}{n^s},$$

donde ambas sumas son tomadas sobre los enteros para los cuales sus divisores primos son menores o iguales a x, donde la suma con barra la hacemos sobre los $n > x$ cuyos divisores primos son menores o iguales que x.

Ahora acotaremos la segunda suma de la siguiente manera teniendo en cuenta que $\sigma > 1$:

$$\left| \overline{\sum_{n>x}} \frac{1}{n^s} \right| \leq \left| \sum_{n>x} \frac{1}{n^s} \right| \leq \sum_{n>x} \frac{1}{n^\sigma} \leq \int_{x-1}^{\infty} t^{-\sigma} dt = \frac{(x-1)^{1-\sigma}}{\sigma - 1} \xrightarrow[x \to \infty]{} 0.$$

Y así, haciendo $x \to \infty$, (1.3) se convierte en

$$\lim_{x \to \infty} \prod_{p \leq x} \frac{1}{1 - \frac{1}{p^s}} = \prod_{p \text{ primo}} \frac{1}{(1 - \frac{1}{p^s})} = \sum_{n=1}^{\infty} \frac{1}{n^s} = \zeta(s).$$

□

1.2. Extensión analítica a todo el plano

Como vimos en (1.1) la función zeta es analítica en el semiplano $Re(s) > 1$, pero Riemann encontró una fórmula válida para todo el plano a excepción del polo simple en $s = 1$, haciendo uso de la función gama

$$\Gamma(s) = \int_0^{\infty} t^{s-1} e^{-t} \, dt.$$

Veamos cómo haciendo la sustitución $t = nx$ obtenemos

$$\frac{\Gamma(s)}{n^s} = \int_0^\infty x^{s-1} e^{-nx}\, dx,$$

y sumando con respecto a n obtenemos, notando que lo que tenemos es una suma geométrica de razón e^{-x}, que

$$\zeta(s)\Gamma(s) = \int_0^\infty \frac{x^{s-1}}{e^x - 1}\, dx,$$

cuando $Re(s) > 1$.

Valiéndonos de esta identidad obtendremos la continuación analítica para $\zeta(s)$.

Para lograrlo, vamos a considerar los siguientes caminos C y C_n, como se muestran en la siguiente figura.

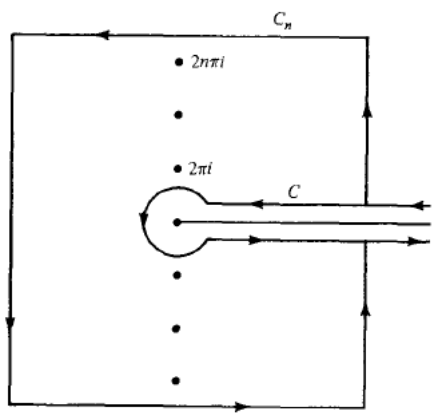

Figura 1.1:

Tenemos entonces el siguiente

Lema 1. *La función Gamma satisface:*

a. $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}$.

b. $\Gamma\left(\frac{1-s}{2}\right)\Gamma\left(\frac{1+s}{2}\right) = \frac{\pi}{\cos \frac{\pi}{2}}$.

c. $\Gamma(s) = 2^{s-1}\pi^{-1/2}\Gamma(\frac{s}{2})\Gamma(\frac{s+1}{2})$ *(Fórmula de duplicación de Legendre en* $\frac{s}{2}$*).*

Demostración. Ver [Ahl79]. □

Podemos ahora concluir el siguiente

Teorema 2 (Extensión analítica de $\zeta(s)$ a todo el plano)**.** *La función* $\zeta(s)$ *se puede extender analíticamente a todo el plano con excepción del polo simple* $s = 1$

$$\zeta(s) = -\frac{\Gamma(1-s)}{2\pi i}\int_C \frac{(-z)^{s-1}}{e^z - 1}\, dz \tag{1.4}$$

Donde $(-z)^{s-1} = e^{(s-1)\log(-z)}$*, y* $\log(-z)$ *es definido cuando* $0 < \arg(z) < 2\pi$*, es decir que* $-\pi < \arg(-z) < \pi$.

Demostración. Las únicas singularidades del integrando son los múltiplos de $2\pi i$, y como e^x crece mucho más rapido que x^s entonces la integral es convergente sobre C. Ahora usando el teorema de Cauchy vemos que la integral no depende de la forma de C, ya que C, como en 5.1 no encierra múltiplos de $2\pi i$, y podemos formar el siguiente contorno, sobre el cual la integral vale cero. (ver figura (1.2))

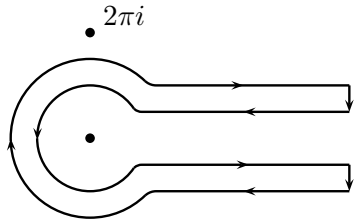

Figura 1.2:

Por lo tanto, podemos en particular hacer tender el radio r a cero, y la integral no cambia.

Por otro lado, tenemos que el módulo de la integral sobre el segmento circular satisface,

$$\int_{|z|=r} \frac{|(-z)^{s-1}|}{|e^z-1|}\,dz \leq 2\pi r \cdot \frac{r^{Re(s)-1}}{r} \xrightarrow[r\to 0]{} 0.$$

Por lo tanto la integral sobre el segmento circular tiende a cero. Ahora tenemos una integral sobre dos caminos que se acercan al eje real positivo, uno por arriba y otro por abajo. En la parte superior $(-z)^{s-1} = e^{(s-1)(\log|z|-i\pi)} = |z|^{s-1}e^{-(s-1)\pi i}$, pues $(-z)$ se estaría acercando por debajo al real real negativo y por lo tanto el argumento de $(-z)$ es $-\pi$. Análogamente en parte inferior $(-z)^{s-1} = e^{(s-1)(\log|z|+i\pi)} = |z|^{s-1}e^{(s-1)\pi i}$. Teniendo cuenta lo anterior tenemos que

$$\begin{aligned}
\int_C \frac{(-z)^{s-1}}{e^z-1}\,dz &= -\int_0^\infty \frac{x^{s-1}e^{-(s-1)\pi i}}{e^x-1}\,dx + \int_0^\infty \frac{x^{s-1}e^{(s-1)\pi i}}{e^x-1}\,dx \\
&= \left(\int_0^\infty \frac{x^{s-1}}{e^{x-1}}\,dx\right)\left(e^{(s-1)\pi i} - e^{-(s-1)\pi i}\right) \\
&= \zeta(s)\Gamma(s)2i\sin[(s-1)\pi] = -2i\zeta(s)\Gamma(s)\sin(s\pi) \\
&= \zeta(s)\frac{-2\pi i}{\Gamma(1-s)},
\end{aligned} \tag{1.5}$$

esto usando la identidad (a) del Lema 1. Así obtenemos la fórmula (1.4).

Ahora bien, la función $\Gamma(1-s)$ es meromorfa y tiene polos en $s = 1, 2, 3, \ldots$, pero sabemos que en estos valores exepto en $s = 1$, $\zeta(s)$ es analítica, según vimos en (1.1), además la integral en (1.4) es una función entera de s, por lo tanto estos polos se deben cancelar con ceros de la integral, así esta fórmula posee únicamente el polo simple en $s = 1$, el cual es polo de la función $\Gamma(1-s)$.

De esta manera concluimos que la fórmula (1.4) coincide con $\zeta(s)$, para $Re(s) > 1$, y la extiende de manera analítica a todo el plano complejo (excepto en $s = 1$). $\square$

1.3. Ecuación funcional

Inicialmente definimos la función $\zeta(s)$, para $Re(s) = \sigma > 1$, pero Riemann encontró que había una relación importante entre $\zeta(s)$ y $\zeta(1-s)$, por la cual podemos tener cierto control sobre los valores de $\zeta(s)$, con $\sigma < 0$.

Teorema 3.

$$\zeta(s) = 2^s \pi^{s-1} \sin \frac{\pi s}{2} \Gamma(1-s)\zeta(1-s)$$

Demostración. Ahora haremos también uso del camino C_n, como en la figura (1.1). Asumiendo que el cuadrado por donde pasa el camino tiene sus lados sobre las rectas $y = \pm(2n+1)\pi i$ y $x = \pm(2n+1)\pi$, formamos el camino cerrado $C_n - C$, el cual encierra los puntos $\pm 2k\pi i$, $m = 1, 2, \ldots$ en los cuales la función $\frac{(-z)^{s-1}}{e^z-1}$ tiene polos simple. En efecto:

$$\lim_{z \to 2k\pi i} (z - 2k\pi i) \cdot \frac{(-z)^{s-1}}{e^z - 1} = \lim_{z \to 2k\pi i} (-z)^{s-1} \cdot \frac{(z - 2k\pi i)}{e^z - 1} = (-2k\pi i)^{s-1} \frac{1}{(e^z)'|_{z=2k\pi i}}.$$

Así tenemos polos simples en $(\pm 2k\pi i)$ con residuos $(\mp 2k\pi i)^{s-1}$. Y por lo tanto por el teorema de residuos de Cauchy:

$$\begin{aligned}
\frac{1}{2\pi i} \int_{C_n - C} \frac{(-z)^{s-1}}{e^z - 1}\, dz &= \sum_{k=1}^{n} \left((2k\pi i)^{s-1} + (-2k\pi i)^{s-1} \right) \\
&= \sum_{k=1}^{n} (2k\pi)^{s-1} \{ (-i)^{s-1} + (i)^{s-1} \} \\
&= \sum_{k=1}^{n} \left((2k\pi)^{s-1} \{ e^{(s-1)i\frac{\pi}{2}} + e^{-(s-1)i\frac{\pi}{2}} \} \right) \\
&= \sum_{k=1}^{n} \left((2k\pi)^{s-1} 2 \cos \left\{ (s-1)\frac{\pi}{2} \right\} \right) \\
&= 2 \sum_{k=1}^{n} (2k\pi)^{s-1} \sin \frac{\pi s}{2}.
\end{aligned} \tag{1.6}$$

Ahora dividimos el camino C_n en dos partes: la parte que pasa por el cuadrado la notaremos C_n^1, y el resto del camino lo notamos por C_n^2.

Como C_n no pasa por los puntos donde se anula $e^z - 1$, es decir en $z = 2k\pi i$, se tiene que $|e^z - 1| \geq N$ y por lo tanto $\frac{1}{e^z-1}$ es acotado por una constante independiente de n sobre C_n^1, y $|(-z)^{s-1}| = |e^{(s-1)\log|z| + i\arg(-z)}| = e^{(Re(s)-1)\log|z|} \leq e^{(\sigma-1)\log\frac{\sqrt{2}}{2}(2n+1)} = (\sqrt{2}e^{\frac{\sqrt{2}}{2}})n^{\sigma-1} = N' n^{\sigma-1}$ y por lo tanto para $\sigma < 1$

$$\left| \int_{C_n^1} \frac{(-z)^{s-1}}{e^z - 1}\, dz \right| \leq \frac{N'}{N} n^{\sigma-1} n = M n^{\sigma} \xrightarrow[n \to \infty]{} 0.$$

La misma afirmación se tiene para el camino C_n^2. Así tenemos que la integral sobre el camino $C_n - C$ tiende a la integral sobre $-C$, por un lado tenemos por (1.5) que

$$\frac{1}{2\pi i}\int_{C_n - C} \frac{(-z)^{s-1}}{e^z - 1}\, dz \longrightarrow \frac{\zeta(s)}{\Gamma(1-s)}.$$

De la misma manera tenemos que si $\sigma < 0$ entonces $Re(1-s) > 1$, y $\sum_{k=1}^{\infty} \frac{1}{k^{1-s}} \longrightarrow \zeta(1-s)$, de tal forma que el límite en la parte derecha de (1.6) es

$$2\sum_{k=1}^{n}(2k\pi)^{s-1}\sin\frac{\pi s}{2} \xrightarrow[n\to\infty]{} 2^s\pi^{1-s}\sin\frac{\pi s}{2}\zeta(1-s)$$

Así de las 2 ecuaciones anteriores obtenemos la conclusión buscada para $\sigma < 0$. Tenemos 2 funciones meromorfas que coinciden en un conjunto no vacío, por lo tanto estas funciones son la misma. Y concluimos así la ecuación funcional para todo $s \neq 1$. □

Cambiando s por $1-s$, en la ecuación funcional obtenemos:

$$\zeta(1-s) = 2^{1-s}\pi^{-s}\sin\frac{\pi(1-s)}{2}\Gamma(s)\zeta(s)$$

y como $\sin(\frac{\pi(1-s)}{2}) = \cos\frac{-\pi s}{2}$, obtenemos una forma equivalente de la ecuación funcional

$$\zeta(1-s) = 2^{1-s}\pi^{-s}\cos\frac{\pi s}{2}\Gamma(s)\zeta(s).$$

El problema fundamental en el estudio de la función zeta de Riemann es la distribución de sus raíces y más especificamente el estudio de sus ceros no triviales, para esto se define la función $\xi(s)$.

Corolario 1. *La función*

$$\xi(s) = \frac{s}{2}(1-s)\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s)$$

es entera y satisface $\xi(s) = \xi(1-s)$.

Demostración. Para ver que es analítica, solo fijémonos que el polo simple de $\zeta(s)$ se cancela con el cero del factor $1-s$, y los polos de $\Gamma(\frac{s}{2})$, se cancelan con los ceros triviales de $\zeta(s)$.

Usando la ecuación funcional de $\zeta(s)$ y el Lema 1, obtenemos

$$\begin{aligned}
\xi(1-s) &= \frac{(1-s)}{2}s\pi^{-(1-s)/2}\Gamma\left(\frac{1-s}{2}\right)\zeta(1-s)\\
&= \frac{(1-s)}{2}s\pi^{-1/2}\pi^{s/2}\Gamma\left(\frac{1-s}{2}\right)2^{1-s}\pi^{-s}\cos\frac{\pi s}{2}\Gamma(s)\zeta(s)\\
&= \frac{s}{2}(1-s)\pi^{-s/2}\pi^{-1/2}\frac{\pi}{\Gamma\left(\frac{1+s}{2}\right)}2^{1-s}\left(2^{s-1}\pi^{-1/2}\Gamma\left(\frac{s}{2}\right)\Gamma\left(\frac{s+1}{2}\right)\right)\zeta(s)\\
&= \frac{s}{2}(1-s)\pi^{-s/2}\Gamma(\frac{s}{2})\zeta(s)\\
&= \xi(s)
\end{aligned}$$

Vemos entonces que las raíces de $\xi(s)$, son precisamente los ceros no triviales de $\zeta(s)$ (el factor s cancela el polo en cero de $\Gamma(\frac{s}{2})$). □

1.4. Acerca de las Raíces de $\zeta(s)$

Como ya dijimos antes una de las grandes preguntas que involucran la función $\zeta(s)$ es como están distribuidas sus raíces.

Riemann en su artículo *On the Number of primes less than a given magnitude* de 1859, planteó una conjetura, hoy en día conocida como *La Hipótesis de Riemann*, la cual afirma que todos los ceros no triviales de la función $\zeta(s)$, están ubicados en la franja $Re(s) = \frac{1}{2}$, y la cual es aún hoy considerada como uno de los mayores problemas de la matemática.

Debido a la fórmula del producto de Eluer

$$\zeta(s) = \prod_{\rho}(1 - p^{-s})^{-1} \qquad (Re\ s > 1)$$

y puesto que ninguno de estos factores es cero vemos que $\zeta(s)$ no tiene raíces para $Re(s) > 1$. De la ecuación funcional (3) vemos además que los números de la forma $s = -2k$, k entero positivo, son raíces de $\zeta(s)$, puesto que el factor $\sin(\pi s/2)$ se anula, cuando k es positivo, los ceros de este factor cancelan los polos del factor $\Gamma(s/2)$. Estas raíces en los enteros pares negativos son llamados los *ceros triviales de la función Zeta*. Ahora también de (3) se concluye que si s es raíz, lo es también $1-s$, es decir que sus raíces son simétricas con respecto al eje $Re(s) = \frac{1}{2}$, y por lo tanto zeta tampoco puede tener raíces en $Re(s) < 0$, así los ceros no triviales de $\zeta(s)$, están ubicados en lo que se conoce como la *franja crítica*, $0 < Re(s) < 1$.

Como paso siguiente veremos un importante resultado acerca de la densidad vertical de los ceros no triviales de $\zeta(s)$ que necesitaremos más adelante, y para ello nos valdremos de los siguientes lemas, que daremos sin demostración:

Lema 2. *Si ρ es tal que $\xi(\rho) = 0$, la siguiente suma converge:*

$$\sum_{\rho} \frac{1}{|p - \frac{1}{2}|^2}$$

Lema 3. *Si $\xi(\rho) = 0$*

$$\xi(s) = \xi(0) \prod_{\rho} \left(1 - \frac{s}{\rho}\right)$$

Para la prueba de estos lemas ver [Edw74].

Lema 4. *La densidad vertical de las raíces ρ de $\varepsilon(\rho) = 0$ es menor que $2\log T$, para T, lo suficientemente grande, es decir existe un H, tal que $T \geq H$ y el número de raíces ρ con parte imaginaria en el rango $T \leq Im(\rho) \leq T + 1$ es menor que $2\log T$*

Demostración. De acuerdo con la fórmula del producto para $\xi(s)$, se tiene que

$$\begin{aligned} \log \xi(s) &= \log \xi(0) + \sum_{\rho} \log\left(1 - \frac{s}{\rho}\right) \\ &= \log \xi(0) + \sum_{\rho} \left(\log(\rho - s) - \log \rho\right). \end{aligned}$$

Y así

$$\frac{\xi'(s)}{\xi(s)} = \sum_{\rho} \frac{1}{s-\rho} \tag{1.7}$$

la diferenciación bajo la suma se justifica por el siguiente cálculo

$$\begin{aligned}\left|\frac{1}{s-\rho} + \frac{1}{s-(1-\rho)}\right| &= \left|\frac{1}{s-\frac{1}{2}-(\rho-\frac{1}{2})} + \frac{1}{s-\frac{1}{2}+(\rho-\frac{1}{2})}\right| \\ &= \left|\frac{2(s-\frac{1}{2})}{(s-\frac{1}{2})^2-(\rho-\frac{1}{2})^2}\right| \leq cte\frac{1}{|\rho-\frac{1}{2}|^2}.\end{aligned}$$

Por el lema anterior y el criterio M-Weierstrass podemos derivar término a término $\log \xi(s)$ y además integrar término a término (1.7) sobre un segmento finito. Luego para cualquier T

$$\int_{2+iT}^{2+i(T+1)} \frac{\xi'(s)}{\xi(s)}\, ds = \sum_{\rho} \int_{2+iT}^{2+i(T+1)} \frac{ds}{s-\rho}. \tag{1.8}$$

Ahora fijando ρ

$$\begin{aligned}Im\left(\int_{2+iT}^{2+i(T+1)} \frac{ds}{s-\rho}\right) &= Im(\log(s-\rho)|_{2+iT}^{2+i(T+1)}) \\ &= \arg(2+i(T+1)-\rho) - \arg(2+iT-\rho) > 0.\end{aligned}$$

Por lo tanto la parte imaginaria de esta integral resulta ser el ángulo entre los segmentos $[\,\rho, 2+iT\,]$ y $[\,\rho, 2i(T+1)\,]$ el cual es positivo pues ya vimos que $Re(\rho) < 1 < 2$, y además estamos suponiendo que $T \leq Im(\rho) \leq T+1$, además este angulo es por lo menos el ángulo conformado por los segmentos $[\,iT, 2+iT\,]$ y $[\,iT, 2+i(T+1)\,]$, el cual vale $\arctan(1/2)$.

Y si n es el número de raíces en la franja $T \leq Im(\rho) \leq T+1$, entonces encontramos una cota inferior para la integral en (1.8).

$$n \cdot \arctan\frac{1}{2} \leq Im\left(\int_{2+iT}^{2+i(T+1)} \frac{\xi'(s)}{\xi(s)}\, ds\right). \tag{1.9}$$

Por otro lado la integral en (1.9) la podemos evaluar directamente con el Teorema fundamental del cálculo evaluando la función

$$\log \xi(s) = \log\left(\frac{s}{2}\right) + \log(1-s) - \frac{s}{2}\log\pi + \log\Gamma\left(\frac{s}{2}\right) + \log\zeta(s) \tag{1.10}$$

entre $2+iT$ y $2+i(T+1)$.

Usaremos el siguiente estimativo para el logaritmo de la función Gama

$$\log\Gamma(s) \approx \left(s-\frac{1}{2}\right)\log s - s + \frac{1}{2}\log 2\pi.$$

Luego (1.10) se comporta como

$$\begin{aligned}\log\xi(s) &\backsim \log\left(\frac{s}{2}\right) + \log(1-s) - \frac{s}{2}\log\pi + \left(\frac{s}{2}-\frac{1}{2}\right)\log\left(\frac{s}{2}\right) - \frac{s}{2} + \frac{1}{2}\log\pi + \log\zeta(s) \\ &= \frac{(s+1)}{2}\log s - \frac{s}{2}\log 2\pi - \frac{s}{2} + \log(1-s) + \log\zeta(s) + \text{ constante}\end{aligned}$$

Puesto que $\log(1-x) = -\sum_n \frac{x^n}{n} \quad |1-x|<1$, entonces

$$\begin{aligned}|\log(\sigma+it)| &= \left|\sum_\rho -\log(1-p^{\sigma+it})\right| = \left|\sum_\rho \sum_n \frac{p^{-n(\sigma+it)}}{n}\right| \\ &\leq \sum_\rho \sum_n \frac{|p^{-n(\sigma+it)}|}{n} = \sum_\rho \sum_n \frac{|p^{-n\sigma}|}{n} \\ &= \log\zeta(\sigma).\end{aligned}$$

Así $|\log(s)| \leq \log\zeta(2) = \log(\pi^2/6) < 1$, sobre el eje $Re(s)=2$, y por esto el cambio de $\log\xi(s)$ entre $2+i(T+1)$ y $2+iT$ es aproximadamente

$$\frac{1+i(T+1)}{2}\log[2+i(T+1)] - \frac{1+iT}{2}\log(2+iT) - \frac{i}{2}\log 2\pi - \frac{i}{2} + \log\left(\frac{1+i(T+1)}{1+iT}\right),$$

o lo que es lo mismo, para M constante,

$$\frac{1+iT}{2}\log[2+i(T+1)] + \frac{i}{2}\log[2+i(T+1)] - \frac{1+iT}{2}\log(2+iT) + \log\left(\frac{1+i(T+1)}{1+iT}\right) + M$$

$$\begin{aligned}&= \frac{i}{2}\log[\,2+i(T+1)\,] + \frac{1+iT}{2}\log\left(\frac{2+i(T+1)}{2+iT}\right) + \log\left(1+\frac{i}{1+iT}\right) + M \\ &= \frac{i}{2}\log[\,2+i(T+1)\,] + \frac{1+iT}{2}\log\left(1+\frac{i}{2+iT}\right) + \log\left(1+\frac{i}{1+iT}\right) + M.\end{aligned}$$

Sin tener en cuenta los términos que son acotados para T, suficientemente grande, el estimativo queda en la forma

$$\frac{i}{2}\log(iT) + \frac{1+iT}{2}\cdot\frac{i}{1+iT} \backsim \frac{i}{2}\log T,$$

y así

$$\int_{2+iT}^{2+i(T+1)} \frac{\xi'(s)}{\xi(s)}\,ds \backsim \frac{i}{2}\log T,$$

con un error acotado por una constante K, cuanto $T\to\infty$. Por lo anterior y de la cota en (1.9)

$$\begin{aligned}n\cdot\arctan\frac{1}{2} &\leq \frac{1}{2}\log T + K, \\ n &\leq \frac{\frac{1}{2}\log T + K}{\pi/8} < 2\log T\end{aligned}$$

para T suficientemente grande, así queda mostrado el resultado.

□

CAPÍTULO 2

Equivalencias de la Hipótesis de Riemann

Muchos matemáticos han intentado resolver la Hipótesis de Riemann indirectamente, demostrando equivalencias de ésta, desafortunadamente ninguna de ellas ha sido suficientemente útil para lograr este fin. Sin embargo, esto no quita valor a estas equivalencias, y muchas son importantes por su relación con diversos resultados, como lo es el Teorema de los Números Primos.

Demostraremos en este capítulo, dos importantes equivalencias de la Hipótesis de Riemann.

Consideremos ahora las funciones

$$\psi(x) = \sum_{p^n \leq x} \log p = \sum_{n \leq x} \Lambda(n) \text{ y } \Theta(x) = \sum_{p \leq x} \log p,$$

donde $\Lambda(n)$ es la función de Von Mangoldt, definida como sigue,

$$\Lambda(n) = \begin{cases} \log p & \text{si } n = p^m \text{ para algún primo } p \text{ y algún } m \geq 1 \\ 0 & \text{en otro caso.} \end{cases}$$

De la definición se tiene la siguiente relación entre $\psi(x)$ y $\Theta(x)$

$$\psi(x) = \Theta(x) + \Theta(x^{1/2}) + \Theta(x^{1/3}) + \cdots + \Theta(x^{1/m}),$$

como $x^{1/m} \geq 2$ entonces $m \leq \frac{\log x}{\log 2}$.

Lema 5. *para* $x > 1$

$$\psi(x) = x - \sum \frac{x^\rho}{\rho} + \sum \frac{x^{-2n}}{2n} - \frac{\zeta'(0)}{\zeta(0)}.$$

Demostración. Ver [Edw74] □

Volviendo a la función zeta, una buena pregunta es que tan rápido crece ésta, desafortunadamente no sabemos mucho acerca de esto, pero lo que si sabemos es lo siguiente:

$$|\zeta(\sigma + it)| \leq \sum_{n=1}^{\infty} \frac{1}{|n^s|} = \sum_{n=1}^{\infty} \frac{1}{n^\sigma} = \zeta(\sigma) \tag{2.1}$$

Por lo tanto $\zeta(s)$, es controlada en cierto modo cuando $t \to \infty$, siempre que $\sigma > 1$.

Existe una conjetura similar, sobre $\sigma = \frac{1}{2}$, la cual es llamada la Hipótesis de Lindelöf:

Hipótesis de Lindelöf: $\forall \epsilon > 0 \;\; \zeta(\frac{1}{2} + it) = O(|t|^{\epsilon})$.

esta conejetura es más débil que la Hipótesis de Riemann, y aunque sería un camino más para descartar la hipótesis de Riemann, su prueba no parece fácil, y hasta el momento no ha sido probada ni refutada. Como es de esperarse si suponemos la Hipótesis de Riemann, se obtienen mejores resultados, y como vimos en (2.1), $\zeta(s)$ se acota, mediante su parte real para valores grandes de su parte imaginaria, pero esto no nos dice mucho pues estamos considerando el creciemiento de la parte imaginaria y no de la parte real, por lo tanto sería deseable estimar su crecimiento en términos de su parte imaginaria. Al respecto, se tiene el siguiente resultado.

Teorema 4. *La Hipótesis de Riemann implica que si* $\frac{1}{2} \leq \sigma \leq 1$ *y* $t \geq 2$ *entonces para todo* $\epsilon > 0$

$$\log \zeta(s) = O((\log t)^{2-2\sigma+\epsilon}).$$

La prueba se puede encontrar en [Ívi85], página 46.

En realidad no vamos a usar directamente este resultado sino la siguiente consecuencia: observemos que como $1/2 < \sigma < 1$ entonces $0 < 2 - 2\sigma < 1$, y si ϵ es suficientemente pequeño, podemos hacer este exponente menor a uno, y por lo tanto:

$$-\log t^{\epsilon} < \log |\zeta(s)| < \log t^{\epsilon} \quad t \geq t_0(\epsilon)$$

y de aquí obtenemos

$$\zeta(s) = O(t^{\epsilon}) \quad \frac{1}{\zeta(s)} = O(t^{\epsilon}) \quad t \geq t_0 \tag{2.2}$$

para todo $\sigma > 1/2$.

Con el propósito de obtener nuestra primera equivalencia, nos centraremos en el orden de la función $\psi(x)$,

Lema 6. *La Hipótesis de Riemann implica que*

$$\psi(x) = x + O(\sqrt{(}x) \log^2 x).$$

Demostración. Siendo $\psi(x)$ una función positiva y creciente, entonces el área del rectángulo de base 1 y altura $\psi(x)$ es menor que el área bajo $\psi(x)$ desde x hasta $x+1$, luego

$$\begin{aligned}\psi(x) &\leq \int_x^{x+1} \psi(t)\, dt = \int_0^{x+1} \psi(t)\, dt - \int_0^x \psi(t)\, dt \\ &= \left| \frac{(x+1)^2}{2} - \frac{x^2}{2} + \sum_{\rho} \frac{(x+1)^{\rho+1} - x^{\rho+1}}{\rho(\rho+1)} + K_1 \right|,\end{aligned}$$

con K_1 una constante. Asumiendo la Hipótesis de Riemann, tenemos que las raíces de $\zeta(s)$ son de la forma $\rho = \frac{1}{2} + i\gamma$, y así tenemos los siguientes estimativos para acotar la suma anterior sobre las raíces ρ:

- Si $|\gamma| \leq x$ usaremos

$$\left|\frac{(x+1)^{\rho+1}-x^{\rho+1}}{\rho(\rho+1)}\right| = \frac{1}{|\rho|}\left|\int_x^{x+1} t^\rho\, dt\right| \leq \frac{(x+1)^{1/2}}{|\rho|} \leq \frac{(2x)^{1/2}}{|\gamma|}$$

- Si $|\gamma| > x$ usaremos

$$\left|\frac{(x+1)^{\rho+1}-x^{\rho+1}}{\rho(\rho+1)}\right| \leq \frac{2|x+1|^{Re(\rho)+1}}{|\rho(\rho+1)|} \leq \frac{2|2x|^{1/2+1}}{\gamma^2}$$

Con H, como en el Lema 2. El número de raíces $\rho = \frac{1}{2}+i\gamma$ en el intervalo $t \leq \gamma \leq t+1$ es menor que $2\log t$, $t \geq H$, así para $H < x$, dividimos la recta en los siguientes intervalos: $\gamma < -x, -x < \gamma < -H, |\gamma| < H, H < \gamma < x$ y $\gamma > x$. la suma de los estimativos de la sumatoria sobre $H < \gamma < x$ y sobre $-x < \gamma < -H$, es lo mismo que el de dos veces la sumatoria sobre $H < \gamma < x$, pues estamos acotando en norma. Análogamente se tiene esto mismo para la suma sobre $\gamma > x$ y $\gamma < -x$, por lo tanto tenemos que:

$$\begin{aligned}
\psi(x) &\leq x + x^{1/2}\sum_{|\gamma|<H}\frac{\sqrt{2}}{|\gamma|} + 2x^{1/2}\sum_{H<\gamma<x}\frac{\sqrt{2}}{\gamma} + 2x^{3/2}\sum_{x<\gamma}\frac{2^{5/2}}{\gamma^2}\\
&\leq x + \text{cte}\cdot x^{1/2} + \text{cte}\cdot x^{1/2}\sum_{t=H}^{x}\frac{\log t}{t} + \text{cte}\cdot x^{3/2}\sum_{t=x}^{\infty}\frac{\log t}{t^2}\\
&\leq x + \text{cte}\cdot x^{1/2} + \text{cte}\cdot x^{1/2}\int_H^x \frac{\log t}{t} + \text{cte}\cdot x^{3/2}\int_x^\infty \frac{\log t}{t^2}\, dt\\
&\leq x + \text{cte}\cdot x^{1/2} + \text{cte}\cdot x^{1/2}\frac{(\log t)^2}{2}\bigg|_H^x + \text{cte}\cdot x^{3/2}\left\{-\frac{\log t}{t}|_x^\infty + \int_x^\infty \frac{dt}{t^2}\right\}\\
&= x + \text{cte}\cdot x^{1/2} + \text{cte}\cdot x^{1/2}(\log x)^2 + \text{cte}\cdot x^{3/2}\left(\frac{logx}{x}+\frac{1}{x}\right)\\
&\leq x + \text{cte}\cdot x^{1/2}(\log x)^2,
\end{aligned}$$

y

$$\psi(x) \geq \int_{x-1}^x = \left|\frac{x^2}{2} - \frac{(x-1)^2}{2} + \sum_\rho \frac{x^{\rho+1}-x^{\rho+1}}{\rho(\rho+1)} + K_2\right|.$$

Como antes vamos a estimar la suma sobre ρ:

- Si $|\gamma| \leq x$ usaremos

$$\left|\frac{x^{\rho+1}-(x-1)^{\rho+1}}{\rho(\rho+1)}\right| = \frac{1}{|\rho|}\left|\int_{x-1}^x t^\rho\, dt\right| \leq \frac{|x|^{1/2}}{|\rho|}.$$

- Si $|\gamma| > x$ usaremos

$$\left|\frac{x^{\rho+1}-(x-1)^{\rho+1}}{\rho(\rho+1)}\right| \leq \frac{2|x|^{Re(\rho)+1}}{|\rho(\rho+1)|} = \frac{2|x|^{3/2}}{\gamma^2}.$$

Luego

$$\begin{aligned}\psi(x) &\geq x - \sum_{|\gamma|<H} \frac{x^{1/2}}{|\gamma|} - 2\sum_{H<\gamma<x} \frac{x^{1/2}}{\gamma} - 2\sum_{x<\gamma} \frac{2x^{3/2}}{\gamma^2} \\ &\geq x - \text{cte}\cdot x^{1/2} - \text{cte}\cdot x^{1/2}\int_H^x \frac{\log t}{t}\,dt - \text{cte}\cdot x^{3/2}\int_x^\infty \frac{\log t}{t^2}\,dt \\ &\geq x - \text{cte}\cdot x^{1/2}(\log)^2,\end{aligned}$$

y por lo tanto se ha probado que $|\psi(x) - x| \leq \text{cte}\cdot x^{1/2}(\log x)^2$. □

Lema 7. $\Theta(x) \leq (\log 4)x$

Demostración. Ver [Jam03]. □

Lema 8. $\psi(x) - \Theta(x) \leq 6x^{1/2}$

Demostración. Como $\Theta(x)$ es una función creciente se tiene que

$$\begin{aligned}\psi(x) - \Theta(x) &\leq \Theta(x^{1/2}) + m\Theta(x^{1/3}), \text{ con } m \leq \frac{\log x}{\log 2} \\ &\leq x^{1/2}\log 4 + \frac{\log x}{\log 2}x^{1/3}\log 4 = x^{1/2}\log 4 + 2x^{1/3}\log x,\end{aligned}$$

y derivando obtenemos la desigualdad $\frac{\log x}{x^\alpha} \leq \frac{1}{\alpha e}$, entonces

$$x^{1/3}\log x = \frac{x^{1/2}x^{1/3}\log x}{x^{1/2}} = x^{1/2}\frac{\log x}{x^{1/6}} \leq \frac{6}{e}x^{1/2}$$

y así obtenemos

$$\psi(x) - \Theta(x) \leq \left(\log 4 + \frac{12}{e}\right)x^{1/2} < 6x^{1/2}$$

□

Como $\psi(x) - \Theta(x) < 6x^{1/2}$, de la definición de ψ y θ, también tenemos que $\Theta(x) \leq \psi(x)$ y por el Lema 3 la Hipótesis de Riemann implica que $\psi(x) = x + O(\sqrt{x}\log^2 x)$, por lo tanto tenemos que la Hipótesis de Riemann también implica que $\Theta(x) = x + O(\sqrt{x}\log^2 x)$.

En lo que sigue haremos uso de la siguiente identidad en repetidas ocasiones.

Lema 9 (Sumación de Abel)**.**

$$\sum_{1\leq r\leq x} a(r)f(r) = A(x)f(x) - \int_1^x A(t)f'(t)\,dt,$$

en donde $A(x) = \sum_{r\leq x} a(r)$

Para la prueba de este lema ver [Apo76].

En este punto ya tenemos todas las herramientas, para ver una importante equivalencia, la cual relaciona la función Zeta con el Teorema de los Números Primos, y más allá de esto, nos da la magnitud del error en dicho Teorema, que es la mejor posible, dada la Hipótesis de Riemann.

Teorema 5. *La Hipótesis de Riemann es equivalente a la siguiente afirmcación: existe x_1 tal que si $x > x_1$ entonces*

$$\pi(x) = \frac{x}{\log x} + O(\sqrt{x}\log x),$$

es decir que el error en el Teorema de los Números Primos es del orden de $\sqrt{x}\log x$.

Demostración. $\Rightarrow$) rescribimos $\Theta(x)$ de la siguiente manera:

$$\Theta(x) = \sum_{n\leq x} b(n) \qquad b(n) = \begin{cases} \log n & n \text{ es primo} \\ 0 & \text{en otro caso} \end{cases}$$

Como $\frac{b(n)}{\log n} = 1$, para n primo entonces $\pi(x) = \sum_{n\leq x} \frac{b(n)}{\log n}$, entonces:

$$\pi(x) = \left(\sum_{n\leq x} b(n)\right)\frac{1}{\log x} + \int_2^x \left(\sum_{n\leq t} b(n)\right)\frac{1}{\log^2} \cdot \frac{1}{t}\, dt,$$

luego

$$\pi(x) = \frac{\Theta(x)}{\log x} + \int_2^x \frac{\Theta(x)\, dt}{t\log^2 t}, \quad x \geq 2, \tag{2.3}$$

e integrando por partes obtenemos:

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

y así

$$Li(x) = \frac{x}{\log x} - \alpha + \int_2^x \frac{1}{\log^2 t}\, dt \tag{2.4}$$

con $\alpha = \frac{2}{\log 2}$, ahora restando las ecuaciones en (2.3) y (2.4), obtenemos:

$$\begin{aligned} \pi(x) - Li(x) &= \alpha + \frac{\Theta(x) - x}{\log x} + \int_2^x \frac{\Theta(t) - t}{t\log^2 t}\, dt \\ &= \alpha + I_1 + I_2. \end{aligned}$$

Ahora acotamos I_1, I_2, y usaremos la cota para $\Theta(x)$, asumiendo la Hipótesis de Riemann.

$$|I_1| \leq K\sqrt{x}\log x, \text{ si } x > x_1, \qquad I_2 = \underbrace{\int_2^{x_1} \frac{\Theta(t) - t}{t\log^2 t}\, dt}_{A} + \int_{x_1}^x \frac{\Theta(t) - t}{t\log^2 t}\, dt,$$

y si $x_1 \leq t \leq x$ entonces $\frac{|\Theta(t)-t|}{\log^2 t} \leq K\sqrt{t} \leq K\sqrt{x}$, y por lo tanto

$$|I_2| \leq A + K\sqrt{x} \cdot \int_{x_1}^x \frac{1}{t}\, dt \leq A + K\sqrt{x}\log x,$$

y por lo cual se tiene que

$$\begin{aligned} |\pi(x) - Li(x)| &\leq \alpha + K\sqrt{x}\log x + A + K\sqrt{x}\log x \\ &\leq K'\sqrt{x}\log x. \end{aligned}$$

Entonces

$$\pi(x) = Li(x) + O(\sqrt{x}\log x),$$

y usando el hecho de que $Li(x) \backsim \frac{x}{\log x}$ obtenemos el resultado buscado.

$\Leftarrow$) la otra implicación necesita un poco más de cuidado, por eso referimos el lector al primer capítulo de [Edw74], y daremos una breve explicación.

Tomando logaritmo en la fórmula del producto de Euler (1.2), y usando la serie de Taylor de logaritmo centrada en 0 $\log(1-x) = -\sum_{n=1}^{\infty} \frac{x^n}{n}$, tenemos que

$$\log\zeta(s) = -\sum_{p \text{ primo}} \log(1-p^{-s}) = \sum_{p \text{ primo}} \sum_{n=1}^{\infty} \frac{p^{-sn}}{n} \qquad Re(s) > 1,$$

y escribimos esta suma como una integral de Riemann-Stieltjes

$$\log\zeta(s) = \int_0^{\infty} x^{-s}\, dJ(x) \quad Re(s) > 1$$

donde $J(x)$, es la función escalonada que da un salto de $1/n$ en potencias n-ésimas de p.

$$J(x) = \frac{1}{2}\left(\sum_{p^n<x} \frac{1}{n} + \sum_{p^n \le x} \frac{1}{n}\right),$$

la cual en el mismo de contexto de Riemann-Stieltjes, puede escribirse como:

$$\frac{\log\zeta(s)}{s} = \int_0^{\infty} J(x)x^{-s-1}\, dx \quad Re(s) > 1.$$

Necesitaremos el siguiente lema

Lema 10.

$$\Phi(s) = \int_0^{\infty} u^{-s}\frac{\phi(u)}{u}\, du \Leftrightarrow \phi(u) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \Phi(s)u^s\, ds$$

Usando el resultado anterior despejamos la función $J(x)$

$$J(x) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \log\zeta(s)x^s\frac{ds}{s}$$

después usando la ecuación que define $\xi(s)$, para rexpresar a $\log\zeta(s)$, se obtiene la fórmula principal

$$J(x) = Li(x) - \sum_{\rho} Li(x^{\rho}) - \log 2 + \int_o^{\infty} \frac{dt}{t(t^2-1)\log t}.$$

se muestra además también que

$$Li(x^{\rho}) = O(\frac{x^{\rho}}{\rho\log x}).$$

Además $\pi(x)$ y $J(x)$ están relacionados por la fórmula

$$J(x) = \pi(x) + \frac{1}{2}\pi(x^{1/2}) + \frac{1}{3}\pi(x^{1/3}) + \cdots + \frac{1}{n}\pi(x^{1/n}),$$

la cual es una suma finita pues en algún punto $x^{1/n} < 2$, y $\pi(x^{1/n}) = 0$. Usando la fórmula de inversión de Moebius tenemos que:

$$\pi(x) = J(x) - \frac{1}{2}J(x^{1/2}) - \frac{1}{3}J(x^{1/3}) - \frac{1}{5}J(x^{1/5}) + \frac{1}{6}J(x^{1/6}) + \cdots + \frac{\mu(n)}{n}J(x^{1/n}) + \cdots .$$

Ahora, suponemos que la Hipótesis de Riemann es falsa, es decir que existe alguna raíz no trivial ρ con $Re(\rho) > \frac{1}{2}$, y por lo tanto un término en la parte de $Li(x^\rho)$ crece más rápidamente que $x^{1/2}$, y por lo tanto el error en el T. de los Números Primos $\pi(x) \sim Li(x)$, no puede crecer menos rápido que $x^{1/2+\epsilon}$. □

Ahora veremos la razón fundamental para interesarnos en la función de Moebius $\mu(n)$. Recordemos la definición de la función μ de Moebius para $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$ con $n > 1$:

$$\begin{aligned}
&\mu(1) = 1\\
&\mu(n) = (-1)^k \text{si } a_1 = a_2 = \cdots = a_k = 1,\\
&\mu(n) = 0 \text{ en otro caso}
\end{aligned}$$

de la definición vemos que $\mu(n) = 0$ si y sólo si n es divisible por un factor cuadrático mayor a uno. La función acumulativa $M(x)$ de Moebius se define como $M(x) = \sum_{n \le x} \mu(n)$.

La siguiente integral es muy útil para la demostración del lema 5, y la usaremos también para el próximo teorema

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{Y^s}{s}\, ds = \begin{cases} 0 & \text{si } 0 \le Y < 1 \\ \frac{1}{2} & \text{si } Y = 1 \\ 1 & \text{si } Y > 1 \end{cases} \tag{2.5}$$

la deducción de esta integral puede ser consultada en [Edw74].

Teorema 6. *La Hipótesis de Riemann es equivalente a que $\forall \varepsilon > 0$ se tiene que $M(x) = O(x^{1/2+\varepsilon})$.*

Demostración. $\Leftarrow$)

Del producto de Euler (1.2) se ve que

$$\prod_{p \text{ primo}} \left(1 - \frac{1}{p^s}\right) = \frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} \quad (Re\ s > 1)$$

Calculamos esta última, sumando por partes como en el Lema 7,

$$\sum_{n \le x} \frac{\mu(n)}{n^s} = \frac{M(x)}{x^s} + s\int_1^x \frac{M(t)}{t^{s+1}}\, dt, \tag{2.6}$$

donde $M(x) = \sum_{n \le x} \mu(n)$, la cual cumple que $|M(x)| \le x$ pues $|\mu(n)| \le 1$ y por lo tanto $\frac{M(x)}{x^{s+1}} \xrightarrow[x\to\infty]{} 0$. Entonces, tomando el límite cuando $x \to \infty$, obtenemos

$$\frac{1}{\zeta(s)} = s\int_1^\infty \frac{M(x)}{x^{s+1}}\, dx \quad Re\ s > 1.$$

Ahora si $|M(x)| < Cx^{1/2}$ entonces se cumple que,

$$\left|\frac{M(x)}{x^{s+1}}\right| \leq \frac{Cx^{1/2}}{|x^{s+1}|} = \frac{C}{|x^s|x^{1/2}} = \frac{C}{x^{Re(s)+1/2}},$$

y la integral converge si $Re(s) + \frac{1}{2} > 1$, es decir si $Re(s) > \frac{1}{2}$. Por lo tanto la ecuación en (2.6) define una función analítica en el semiplano $Re(s) > \frac{1}{2}$, de tal manera que $\zeta(s)$ no tendría raíces en esta región, pero por la simetría de las raíces que se deduce de la ecuación funcional tampoco las habría en el semiplano $Re(s) < \frac{1}{2}$, así todas las raíces deberían estar en la franja $Re(s) = \frac{1}{2}$, y la Hipótesis de Riemann sería cierta.

$\Rightarrow$) Observemos lo siguiente. Recordando que $\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}$,

es razonable pensar en la aproximación siguiente para $M(x)$,

$$M(x) = \sum_{n<x} \mu(n) \sim \frac{1}{2\pi i} \int_{2-iT}^{2+iT} \left[\sum_{n=1}^{\infty} \mu(n) \left(\frac{x}{n}\right)^s\right] \frac{ds}{s},$$

pues hablando en términos no formales, si sacamos la suma de la integral, tomamos límite cuando $T \to \infty$ y usamos (2.5), sobreviven los término de la suma en los que $n < x$, y así $\frac{x}{n} > 1$, por lo que la integral valdría 1. Seguido de esto, vamos a calcular el orden de error en esta aproximación, que depende de T y de x.

$$\begin{aligned}
&\sum_{n<x} \mu(n) - \frac{1}{2\pi i} \int_{2-iT}^{2+iT} \left[\sum_{n=1}^{\infty} \mu(n) \left(\frac{x}{n}\right)^s\right] \frac{ds}{s} \\
&= \sum_{n<x} \mu(n) \left[1 - \frac{1}{2\pi i} \int_{2-iT}^{2+iT} \left(\frac{x}{n}\right)^s \frac{ds}{s}\right] + \sum_{n>x} \mu(n) \frac{1}{2\pi i} \int_{2-iT}^{2+iT} \left(\frac{x}{n}\right)^s \frac{ds}{s}.
\end{aligned}$$

Hemos sacado la suma de la integral, esto se justifica pues la sumatoria es absolutamente convergente, y el intervalo de integración es finito. Usaremos los siguientes estimativos, que pueden ser encontrados en [Edw74], páginas 54-55,

$$\begin{aligned}
\left|\int_{a-ih}^{a+ih} \frac{x^s}{s}\, ds\right| &\leq \frac{x^a}{\pi h |\log x|} \quad (a > 0,\ 0 < x < 1) \\
\left|\int_{a-ih}^{a+ih} \frac{x^s}{s}\, ds - 1\right| &\leq \frac{x^a}{\pi h \log x} \quad (a > 0,\ x > 1).
\end{aligned}$$

Con esto el estimativo de arriba queda en la forma

$$\begin{aligned}
&\sum_{n<x} \frac{|\mu(n)|(x/n)^2}{nT\log(x/n)} + \sum_{n>x} \frac{|\mu(n)|(x/n)^2}{nT\log(n/x)} \\
&\leq \frac{x^2}{\pi T} \left[\sum_{n<x} \frac{1}{n^2 \log(x/n)} + \sum_{n>x} \frac{1}{n^2 \log(n/x)}\right]
\end{aligned}$$

y la idea es ahora ver que las dos sumas son acotadas por una constante, lo haremos con la primera suma la otra es similar. Para la primera suma tenemos que

$$\sum_{n<x} \frac{1}{n^2 \log(x/n)} = \sum_{n \leq x/2x} \frac{1}{n^2 \log(x/n)} + \sum_{x/2<n<x} \frac{1}{n^2 \log(x/n)}$$

si x no es entero y como $\log(1+y) \geq \frac{y}{2}$, si $0 \leq y \leq 1$, entonces la suma anterior queda acotada por

$$\frac{1}{\log 2}\sum_{n\leq x/2}\frac{1}{n^2}+\sum_{x/2<n<x}\frac{1}{n^2}\frac{2n}{x-n} \leq \frac{\zeta(2)}{\log 2}+\frac{4}{x}\sum_{0<x-n<x/2}\frac{1}{x-n}$$
$$\frac{\zeta(2)}{\log 2}+\frac{4}{x}\cdot\frac{1}{x-[x]}+\frac{4}{x}\sum_{0<j<x/2}\frac{1}{j}.$$

Y así vemos que la primera suma es acotada, pues $x-[x]$, no va a ser pequeño bajo la suposición de que x no es entero. y por lo tanto el error en la aproximación

$$M(x) \sim \frac{1}{2\pi i}\int_{2-iT}^{2+iT}\frac{x^s}{\zeta(s)}\frac{ds}{s} \tag{2.7}$$

es menor que $B\cdot x^2/T$, con B constante, cuando $x\to\infty$.

Dividimos ahora la integral en (2.7) en 3 integrales, mediante el cuadrado con vértices en $2-iT, 1/2+\epsilon-iT, 1/2+\epsilon+iT, 2+iT$, así por el Teorema de Cauchy dicha integral es igual a:

$$\frac{1}{2\pi i}\int_{2-iT}^{(1/2)+\epsilon-iT}\frac{x^s}{\zeta(s)}\frac{ds}{s}+\frac{1}{2\pi i}\int_{(1/2)+\epsilon-iT}^{(1/2)+\epsilon+iT}\frac{x^s}{\zeta(s)}\frac{ds}{s}+\frac{1}{2\pi i}\int_{(1/2)+\epsilon+iT}^{2+iT}\frac{x^s}{\zeta(s)}\frac{ds}{s},$$

pues $\zeta(s)\neq 0$, en dicha región si suponemos la Hipótesis de Riemann. Ahora acotamos la primera y la tercera integral, usando (2.2), con lo que $|\frac{1}{\zeta(s)}| \leq Kt^{\delta}$:

$$\left|\frac{1}{2\pi i}\int_{(1/2)+\epsilon+iT}^{2+iT}\frac{x^s}{\zeta(s)}\frac{ds}{s}\right| \leq \frac{2-(1/2)-\epsilon}{2\pi}\frac{x^{\sigma}}{Im(s)}\cdot KT^{\delta}$$

por lo que estas integrales están acotadas por una constante veces $x^2\cdot KT^{\delta-1}$. En cuanto a la integral de la mitad usamos la siguiente cota, para un T_0 fijo:

$$\left|\frac{1}{2\pi i}\int_{(1/2)+\epsilon-iT_0}^{(1/2)+\epsilon+iT_0}\frac{x^s}{\zeta(s)}\frac{ds}{s}\right|+\frac{2}{2\pi}\int_{T_0}^{T}\frac{x^{(1/2)+\epsilon}Kt^{\delta}}{t}\,dt$$
$$\leq x^{(1/2)+\epsilon}\cdot const+\frac{x^{(1/2)+\epsilon}KT^{\delta}}{\pi\delta},$$

estos pues T_0 es constante. Ahora hagamos $T=x^2$, y así tenemos que $M(x)$ es menor que una constante veces $x^{(1/2)+\epsilon+2\delta}$, para x suficientemente grande. Por lo tanto, siendo $\epsilon>0$ y $\delta>0$ arbitrarios, y como el cambio de $M(x)$ entre enteros es a lo más ± 1, lo mismo es cierto para todos los valores de x, luego hemos obtenido, asumiendo la Hipótesis de Riemann, que $M(x)$ crece menos rápido que $x^{1/2+\epsilon}$, para todo $\epsilon>0$. □

CAPÍTULO 3

Análisis probabilístico de la Hipótesis de Riemann

3.1. Argumento de Denjoy

Ahora cambiamos el modo de ver la Hipótesis de Riemann, a una manera un poco más intuitiva. Aplicaremos argumentos probabilísticos y haremos experimentos numéricos sobre la función μ de Moebius para estudiar la posible veracidad de la Hipótesis de Riemann.

Comenzaremos con un argumento probabilístico debido a Denjoy.

El siguiente lema es un resultado clásico en Teoría de probabilidad, y será nuestra principal herramienta en este capítulo.

Lema 11 (Desigualdad de Chebyshev)**.** *Sea X una variable aleatoria con media μ y varianza σ^2 entonces para todo número real $k > 0$ se tiene*

$$P\{|x-\mu| \geq k\sigma\} \leq \frac{1}{k^2}.$$

Supongamos que lanzamos una moneda justa un número dado de veces N, y que N es un valor grande, y consideremos la variable aleatoria X como el número de caras obtenidas en N lanzamientos de la moneda. Puesto que la probabilidad de obtener una cara en un lanzamiento es $\frac{1}{2}$, X tiene una distribución binomial con media $\mu = \frac{N}{2}$ y varianza $\sigma^2 = \frac{N}{4}$.

Usando la desigualdad de Chebyshev tenemos:

$$P\left\{\left|X - \frac{N}{2}\right| < KN^{1/2}\right\} = 1 - P\left\{\left|X - \frac{N}{2}\right| \geq 2K\frac{N^{1/2}}{2}\right\} \geq 1 - \frac{1}{4k^2}$$

Ahora sea y el número de caras menos el número de sellos, entonces
$y = X - (N - X) = 2X - N = 2(X - \frac{N}{2})$ y por lo tanto,

$$P\left\{|y| < 2KN^{1/2}\right\} = P\left\{2\left|X - \frac{N}{2}\right| < 2KN^{1/2}\right\} \geq 1 - \frac{1}{4K^2}.$$

Entonces tomando $k = \frac{N^\epsilon}{2}$ tenemos que

$$P\{|y| < N^{1/2+\epsilon}\} \geq 1 - \frac{1}{4(N^\epsilon/2)^2} = 1 - \frac{1}{N^{2\epsilon}} \underset{N\to\infty}{\longrightarrow} 1,$$

es decir que con probabilidad cercana a uno el número de caras menos el número de sellos crece menos rápido que $N^{1/2+\epsilon}$.

Ahora tomamos un entero n suficientemenete grande que no sea divisible por un cuadrado, es decir que $\mu(n) \neq 0$. Entonces tenemos que la probabilidad de que un entero m sea múltiplo de x es

$$P\{m = kx \text{ algún k} \in \mathbb{Z}\} = \frac{1}{x},$$

luego la probabilidad de que n sea libre de cuadrados es la probabilidad de que n no sea múltiplo de 4, ni de 9, ni de 25 , etc., lo cual nos lleva a que

$$\begin{aligned} P\{\mu(n) \neq 0\} &= \left(1 - \frac{1}{2^2}\right)\left(1 - \frac{1}{3^2}\right)\left(1 - \frac{1}{5^2}\right)\cdots \\ &= \prod_{p \text{ primo}} \left(1 - \frac{1}{p^2}\right) = \frac{1}{\zeta(2)} \\ &= \frac{1}{\sum_{n=1}^{\infty} \frac{1}{n^2}} = \frac{6}{\pi^2}. \end{aligned}$$

Así obtenemos que

$$P\{\mu(n) \neq 0\} = \frac{6}{\pi^2}, \quad \text{y } P\{\mu(n) = 0\} = 1 - \frac{6}{\pi^2}.$$

pero no conocemos exactamente los valores de $P\{\mu(n) = 1\}$ ni con $P\{\mu(n) = -1\}$.

Pero si suponemos que

$$P\{\mu(n) = 1\} = P\{\mu(n) = -1\} = \frac{3}{\pi^2},$$

entonces podemos pensar en $M(x) = \sum_{n \leq x} \mu(n)$ como el experimento de lanzar una moneda por cada entero libre de cuadrados menor que x restando el número de caras del número de sellos obtenidos. Por lo tanto, dado $\epsilon > 0$, el resultado de este experimento es con probabilidad cercana a 1, menor que el número de lanzamientos elevado a la $1/2 + \epsilon$. Así concluimos que $M(x) = O(x^{1/2+\epsilon})$, y por el Teorema 6 se concluye que la Hipótesis de Riemann es cierta con probabilidad uno, claro está bajo la suposición de que la probabilidad de obtener un número par de primos en la descomposición de n es igual a la probabilidad de obtener un número impar de factores primos para este número libre de cuadrados es igual.

Realmente en el anterior argumento no es importante la suposición $P\{\mu(n) = 1\} = P\{\mu(n) = -1\} = \frac{3}{\pi^2}$, pero podemos dar un argumento que concluye lo mismo teniendo encuenta este valor para las probabilidades, y utilizando como antes la deisgualdad de chebyshev.

Consideramos la variable aleatoria $\sum_{n \leq x} \mu(n)$, y olvidemos por un momento la definición para los valores de $\mu(n)$, y solamente asumamos que

$$\mu(n) = \begin{cases} -1 & \text{con probabilidad } \frac{3}{\pi^2} \\ 0 & \text{con probabilidad } 1 - \frac{6}{\pi^2} \\ -1 & \text{con probabilidad } \frac{3}{\pi^2}. \end{cases}$$

Debemos calcular su media y su varianza

$$\begin{aligned} E(\mu(n)) &= -1 \cdot P\{\mu(n) = -1\} + 0 \cdot P\{\mu(n) = 0\} + 1 \cdot P\{\mu(n) = 1\} \\ &= -\frac{3}{\pi^2} + 0 + \frac{3}{\pi^2} \\ &= 0. \end{aligned}$$

Además tenemos que

$$[\mu(n)]^2 = \begin{cases} 1 & \text{con probabilidad } \frac{6}{\pi^2} \\ 0 & \text{con probabilidad } 1 - \frac{6}{\pi^2}, \end{cases}$$

y por lo tanto como $E[\mu(n)] = 0$, y $E[(\mu(n))^2] = \frac{6}{\pi^2}$, entonces

$$V[(\mu(n))] = \frac{6}{\pi^2} - 0^2 = \frac{6}{\pi^2}.$$

Al igual que en el experimento de la moneda supondremos además que las evaluaciones de $\mu(n)$ son independientes, y así viendo a nuestra variable aleatoria como una suma de variables de aleatorias, podremos decir lo siguiente:

$$E\left[\sum_{n \leq x} \mu(n)\right] = \sum_{n \leq x} E[\mu(n)] = 0 \quad \text{y} \quad Var\left[\sum_{n \leq x} \mu(n)\right] = \sum_{n \leq x} Var[\mu(n)] = x\frac{6}{\pi^2}.$$

Con esto estamos listos para usar la desigualdad de Chebyshev con $\mu = 0$, $\sigma = \sqrt{x} \cdot \frac{\sqrt{6}}{\pi}$ y $k = \frac{\pi}{6}x^{\epsilon}$, obteniendo:

$$\begin{aligned} P\left\{\left|\sum_{n \leq x} \mu(n)\right| \leq x^{1/2+\epsilon}\right\} &= P\{|\sum_{n \leq x} \mu(n)| \leq \sqrt{x}\frac{\sqrt{6}}{\pi} \cdot \frac{\pi}{\sqrt{6}}x^{\epsilon}\} \\ &\geq 1 - \frac{6}{\pi^2 x^{2\epsilon}} \end{aligned}$$

haciendo tender x a infinito obtenemos la conclusion deseada, y nuevamente con probabilidad muy cercana a 1, la Hipótesis de Riemann vale.

Este argumento muestra una vez más la fuerte relación que existe entre la aleatoridad de $M(x)$ y la Hipótesis de Riemann, y la importancia de estudiar la aleatoriedad de la función de Moebius.

El argumento de Denjoy está sujeto a suposiciones de independencia, que no se cumplen para la función μ de Mobius, fue entonces cuando el profesor *Francois Leyvraz*, se preguntó, que pasaría al incluir las correlaciones obvias entre los valores de la función $\mu(n)$, como lo es por ejemplo la relación que hay entre el valor de $\mu(n)$ y $\mu(2n)$, pues el valor de la función en $2n$, esta ligado directamente con el de n, ya que por ejemplo si $\mu(n) = 0$, entonces a $\mu(2n)$ no le queda otra opción que valer cero. Este fenómeno pasa entonces para cada valor primo de p, en el que existe esta correlación entre $\mu(n)$ y $\mu(pn)$. Bajo esta sugerencia, hemos modificado el argumento de Denjoy, y es sorprendente que a pesar de descartar la independencia entre los valores nombrados anteriormente (aunque suponemos independencia en el resto), el resultado es el mismo!.

Igual que antes nuestra variable aletoria es $\sum_{n\leq x}\mu(n)$, el agregar la dependencia entre los valores citados no afecta el cálculo de la esperanza, por lo que $E\left[\sum_{n\leq x}\mu(n)\right]=0$, y debemos centrarnos en el valor de la varianza.

$$V\left[\sum_{n=1}^{x}\mu(n)\right]=\sum_{n=1}^{x}V(\mu(n))+2\sum_{i<j\leq x}\sum Cov(\mu(i),\mu(j)),$$

donde $Cov(X,Y)$ representa la covarianza entre las variables aleatorias X e Y. Recordemos que

$$\begin{aligned}Cov(\mu(n),\mu(pn))&=E(\mu(n)\cdot\mu(pn))-E(\mu(n))\cdot E(\mu(pn))\\&=E(\mu(n)\cdot\mu(pn)).\end{aligned}$$

y para el resto de valores la covarianza es cero, entonces necesitamos calcular la esperanza de este producto, para lo cual observemos que

$$\mu(n)\cdot\mu(pn)=\begin{cases}0 & \mu(n)=0 \text{ o } p\mid n\\ -1 & \mu(n)\neq 0 \text{ y } p\nmid n.\end{cases}$$

Supondremos por simplicidad que cada uno de estos valores se obtiene con probabilidad $1/2$, esta suposición no afecta el resultado pues el verdadero valor de la probabilidad es menor a uno, y al fin y al cabo es una constante que no influye en el comportamiento asintótico de la probabilidad que vamos a calular al hacer x grande. Siendo así, obtenemos

$$Cov(\mu(n),\mu(pn))=-\frac{1}{2}$$

Seguido a esto debemos observar cuantos pares de la forma k,kp existen en la suma de la covarianza, dichos pares son del orden de $\frac{x}{p}$, entonces la varianza es del orden de

$$V\left[\sum_{n=1}^{x}\mu(n)\right]=O\left(\frac{6}{\pi^2}\cdot x+2\sum_{p\text{ primo}}\frac{x}{p}\left(\frac{-1}{2}\right)\right),$$

y haciendo uso del conocido Teorema de Mertens (1874), el cual afirma que, existe una constante A, tal que

$$\sum_{p\leq x}\frac{1}{p}=\log\log x+A+O\left(\frac{1}{\log x}\right)\quad\forall x\geq 2,$$

(En realidad nos interesa que esta estimación es del orden de $\log\log x$. ver [Apo76], [Jam03].) obtenemos

$$V\left[\sum_{n=1}^{x}\mu(n)\right]=O(x\log\log x)$$

El siguiente paso es utilizar Chebyshev, con $\sigma=x^{1/2}\sqrt{\log\log x}$ y $k=\frac{x^{\epsilon}}{\sqrt{\log\log x}}$, entonces:

$$\begin{aligned}P\left\{\left|\sum_{n\leq x}\mu(n)\right|\leq x^{1/2+\epsilon}\right\}&=P\left\{\left|\sum_{n\leq x}\mu(n)\right|\leq x^{1/2}\sqrt{\log\log x}\frac{x^{\epsilon}}{\sqrt{\log\log x}}\right\}\\&\geq 1-\frac{1}{\frac{x^{2\epsilon}}{\log\log x}}=1-\frac{\log\log x}{x^{2\epsilon}}\xrightarrow[x\to\infty]{}1\end{aligned}$$

Y nuevamente con probabilidad muy cercana a uno, la Hipótesis de Riemann vale.

Como dijimos antes, es interesante que el resultado se tenga a pesar de introducir las nuevas correlaciones que aproximan el problema cada vez más al problema real.

3.2. El índice de Hurst

Por lo descrito es la sección anterior seria conveniente medir de alguna manera que tan aleatoria es la función $\mu(n)$, y aunque hasta el momento no se sabe con exactitud la respuesta a esta pregunta, vamos a calcular experimentalmente una medida de la aleatoriedad de la función de Moebius mediante la metodología que se describe en esta sección.

Primero damos unas deficiones.

Definición 1 (Proceso estocástico)**.** *Un proceso estocástico X, es una colección de variables aleatorias $X = \{x(t) \mid t \in T\}$, es decir para cada t en el conjunto de índices T, $X(t)$ es una variable aleatoria.*

Definición 2. *Para un proceso estocástico $\{X(t), t \in T\}$:*

1. *Se dice que tiene* incrementos independientes *si para cualquier $t_o < t_1 < \cdots, t_n$, las variables aleatorias $X(t_1) - X(t_0), X(t_2) - X(t_1), \cdots, X(t_n) - X(t_{n-1})$ son independientes.*

2. *Se dice que tiene* incrementos estacionarios *si $X(t+s) - X(t)$ tiene la misma distribución para todo t.*

Definición 3. *Un proceso estocástico $\{X(t), t \in T\}$ se dice* proceso Gaussiano *si $X(t_1), X(t_2), \ldots, X(t_n)$, tiene una disribucion conjunta normal, para cualquier $t_1, t_2, \ldots, t_n$.*

Comenzamos considerando una caminata aleatoria simple, es decir una caminata en la que en cada unidad de tiempo nos despalzamos una unidad hacia la izquierda o hacia la derecha con igual probabilidad. Ahora supongamos que hacemos este proceso, en unidades de tiempo y de desplazamiento cada vez más pequeñas, al llevar este proceso al límite obtenemos lo que se conoce como *Movimiento Browniano*. Siendo un poco más precisos, suponemos que caminamos un paso de tamaño Δx a la izquierda o a la derecha con igual probabilidad, cada Δt unidades de tiempo. Sea $X(t)$ la variable de posición en el tiempo t, entonces habremos avanzado $\frac{t}{\Delta t}$ pasos de tamaño Δx cada uno, y así

$$X(t) = \Delta x \left(X_1 + X_2 + \cdots + X_{\frac{t}{\Delta t}} \right),$$

donde

$$X_i = \begin{cases} +1 & \text{el paso } i, \text{ es hacia la derecha} \\ -1 & \text{si es hacia la izquierda} \end{cases}$$

Además el valor de cada X_i es independiente del valor de los otros, con

$$P\{X_i = 1\} = P\{X_i = -1\} = \frac{1}{2}.$$

Como $E(X_i) = 0$, $Var(X_i) = E[X_i^2] = 1$ y como $X(t)$ es una suma de variables independientes, entonces

$$E[\,X(t)\,] = 0 \qquad Var(X(t)) = (\Delta x)^2 \left[\frac{t}{\Delta t}\right]. \tag{3.1}$$

ahora debemos hacer tender tanto a Δx como a Δt a cero de manera adecuada, pero si no lo hacemos conevenientemente podría resultarnos en un proceso trivial, como al hacer $\Delta x = \Delta t$ y hacer $\Delta t \to 0$. Escogemos $\Delta x = c\sqrt{\Delta t}$, para alguna constante positiva c, y de (3.1) vemos que si $\Delta t \to 0$

$$E[\,X(t)\,] = 0 \qquad Var(X(t)) \to c^2 t.$$

Ahora damos una descripcion de este proceso, teniendo en cuenta la definición de $X(t)$ y el Teorema del límite central:

(i) $X(t)$ es normal con media 0 y varianza $c^2 t$

Puesto que los cambios en el valor de la caminata aleatoria, en intervalos de tiempo disyuntos son independientes, entonces

(ii) $\{X(t), t \geq 0\}$ tiene incrementos independientes.

Finalmente, como la distribución del cambio en la posición de la caminata sobre cualquier intervalo de tiempo depende solamente de la longitud del intervalo, entonces

(iii) $\{X(t), t \geq 0\}$ tiene incrementos estacionarios.

Resumiendo, tenemos la siguiente definición,

Definición 4. *Un proceso estocástico* $\{X(t), t \geq 0\}$ *es un* movimiento browniano *si:*

(i) X(0) = 0;

(ii) $\{X(t), t \geq 0\}$ *tiene incrementos estacionarios independientes;*

(iii) Para todo $t > 0$, $X(t)$ *tiene distribución normal con media 0 y varianza* $c^2 t$.

Enistein en 1905 dio una justificación matemática del movimiento browniano. Él mostró que este proceso explicaba el movimiento de una partícula sumergida en un fluido, la cual es sometida al bombardeo continuo de las partículas alrededor de esta, en la cual la distancia recorrida por la partícula es proporcional a la raíz cuadrada del tiempo.

Cuando $c = 1$ el proceso es llamado movimiento browniano estándar. Como $X(t)$ es normal con media 0 y varianza t, su función de densidad es

$$f_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}.$$

Como $X(t)$ tiene incrementos independientes, se sigue que la función de densidad conjunta de $X(t_1), X(t_2), \ldots, X(t_n)$ es dada por

$$f(x_1, x_2, \ldots, x_n) = f_{t_1}(x_1) f_{t_2 - t_1}(x_2 - x_1) \cdots f_{t_n - t_{n-1}}(x_n - x_{n-1}),$$

la cual es una distrubución normal multivariada, y por lo tanto es un proceso gaussiano, los cuales están determinados por los valores de su media y de su covarianza.

Luego el movimeinto browniano (estándar) se puede definir como un proceso gaussiano con $E[X(t) = 0]$ y para $s \leq t$,

$$\begin{aligned} Cov(X(s), X(t)) &= Cov(X(s), X(s) + X(t) - X(s)) \\ &= Cov(X(s), X(s)) + Cov(X(s), X(t) - X(s)) \\ &= E(X(s)^2) - (E[X(s)])^2 + 0 \\ &= Var(X(s)) = s, \end{aligned}$$

donde hemos usado para la tercera igualdad los incrementos independientes.

Así $Cov(X(s), X(t)) = \text{mín}(s, t)$.

Así definido el movimiento browniano define un proceso completamente aleatorio. Pero desde la naturaleza misma hasta la economía poseen procesos que no son completamente aleatorios y en los que el valor del proceso depende de lo ocurrido en el pasado, como es el caso de la variación en las tasas de cambio de las diferentes monedas.

Por esto se ha generalizado el movimiento browniano, para incluir posibles correlaciones, por un proceso llamado *movimiento browniano fraccionario*, que tiene en cuenta la posible correlación de las variables aleatorias $X(t)$ para distintos valores de t, correlación que es medida mediante el *índice de Hurst*. El índice de Hurst fue introducido por el hidrólogo británico H.E Hurst en su arículo (1951) “The Long-Term Storage Capacity of Reservoirs", en el cual propone un proceso, llamado el proceso de Hurst, mediante el cual, usando registros de más de 800 años, concluyó que el flujo que entraba al rio Nilo proveniente de lluvias y otras vertientes entre otros, no era un proceso completamente aleatorio. Hurst uso este proceso y más adelante Mandelbrot lo formalizó en la siguiente definición

Definición 5 (Movimiento Browniano Fraccionario)**.** *Sea $H \in (0, 1)$ constante. Un* movimiento browniano fraccionario (mBf) $(B^H(t))_{t\geq 0}$ *con índice de Hurst H, es un proceso gaussiano continuo y centrado, con función de covarianza:*

$$E[B^H(t)B^H(s)] = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H}).$$

Observemos que en la parte de la función de covarianza, correspondiente al producto de las esperanzas, es cero, debido a la suposición de ser un proceso centrado de media 0.

El valor de H determina el tipo de proceso:

1. Si $H = 1/2$, entonces el proceso es un movimiento browniano usual, y por lo tanto un proceso aleatorio.
2. Si $H > 1/2$, el proceso es persistente , es decir un proceso con memoria a largo plazo.
3. Si $H < 1/2$, el proceso es antipersistente.

El tipo de proceso más comúnmente encontrado en la naturaleza así como en los mercados y en la economía, son los procesos persistentes. De la función de covarianza vemos que si $H = 1/2$, entonces un mBf, es un movimiento browniano usual.

De la definición 5 también observamos que un mBf estándar tiene la siguientes propiedades:

1. $B^H(0) = 0$ y $E[B^H(t)] = 0$ para todo $t \geq 0$.
2. B^H tiene incrementos estacionarios, es decir $B^H(t+s) - B^H(s)$ tiene la misma distribución que $B^H(t)$ para $s, t \geq 0$.
3. B^H es un proceso gaussiano y $Var(B^H(t)) = E[B^H(t)^2] = t^{2H}$, $t \geq 0$, para todo $H \in (0,1)$.
4. B^H es un proceso continuo.

3.3. Determinación de índice de Hurst

Como hemos dicho anteriormente, Hurst desarrollo una metodología para evaluar el índice, denomiada también análisis R/S y de esta decidir el nivel de independencia de una serie de datos.

Hurst se basó en el trabajo sobre movimiento browniano de Einstein, el cual encontró que la distancia recorrida por una partícula aleatoria, es proporcional a la raíz cuadrada del tiempo, es decir:

$$R = T^{0,5}, \tag{3.2}$$

donde $R =$ la distancia cubierta y T es el índice para el tiempo.

Para calcular el índice de Hurst comenzamos con una serie de tiempo, $x = x_1, x_2, \ldots, x_n$, y calculamos su valor medio m y su varianza s

$$m = \frac{x_1 + x_2 + \cdots + x_n}{n} \quad S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - m)^2}{n}}.$$

Ahora calcularemos lo que se conoce como el rango rescalado, para lo cual primero debemos normalizar los datos de la siguiente manera

$$z_i = x_i - m, \quad i = 1, 2, \ldots, n.$$

Esta nueva serie z, tiene entonces media 0. Después de esto creamos una nueva serie de tiempo acumulativa y

$$y_i = z_1 + z_2 + \cdots + z_i, \quad i = 1, \ldots, n.$$

Como z tiene media cero, el último valor de y, es decir y_n, es siempre cero.

El rango ajustado, R_n se define como

$$R_n = \text{máx}(y_1, y_2, \ldots, y_n) - \text{mín}(y_1, y_2, \ldots, y_n).$$

Además, habiendo definido y, de la forma en que lo hicimos y tiene media 0, y por lo tanto el máximo valor de y es siempre no negativo y el mínimo es siempre menor o igual a cero, luego R_n será siempre mayor o igual a cero.

Intuitivamente R_n es la distancia cubierta por el sistema através del tiempo. Si nuestra serie de tiempo fuera un movimiento browniano podríamos usar la ecuación de Einstein (3.2), haciendo $n = T$. Pero Hurst encontró una forma más general de esta ecuación la cual puede ser aplicada a cualquier serie de tiempo

$$(R/S)_n = c \cdot n^H, \tag{3.3}$$

donde c es una constante, H es el índice de Hurst y el subíndice n hace referencia al tamaño de la serie $x = x_1, \ldots, x_n$.

El valor de R/S, es conocido como el rango rescalado. El rescalar los datos para que tengan media cero y varianza uno, nos permite comparar varios fenómenos en diversos períodos de tiempo.

Si realizamos este proceso para diferentes valores de n, el índice de Hurst se obtiene tomando logaritmo a ambos lados de (3.3):

$$\log(R/S)_n = \log c + H \cdot \log n, \tag{3.4}$$

la cual es la ecuación de una recta, con variable dependiente $\log(R/S)_n$, variable independiente $\log n$ y con pendiente H; así H puede ser estimado usando una regresión lineal por mínimos cuadrados, teniendo varios valores de $\log(R/S_n)$ y de $\log n$.

Para obtener los valores de las parejas $(\log n, \log(R/S)_n)$, subdividimos de manera adecuada la serie, como se describe en el siguiente algoritmo:

1. Si nuestra serie inicial w, es de tamaño M, la convertimos en una nueva serie x de tamaño $N = M - 1$:

$$x_i = \log\left(\frac{x_{i+1}}{x_i}\right), \qquad i = 1, 2, \ldots, N.$$

2. Dividimos nuestra serie en A subperiodos de tamaño d, para cada divisor de N, es decir para cada d, tal que $A \cdot d = N$. Denotamos cada subperiodo por I_a, $a = 1, 2, \cdots, A$. Y cada elemento del subperiodo lo notamos por $X_{k,a}$, para $k = 1, 2, \cdots, d$. Y para cada subperiodo I_a de tamaño d, calculamos su media:

$$m_a = \frac{1}{n}\sum_{k=1}^{d} X_{k,a}.$$

3. Definimos la serie acumulativa $Y_{k,a}$, para cada subperiodo I_a, como sigue:

$$Y_{k,a} = \sum_{i=1}^{k}(X_{i,a} - m_a) \qquad k = 1, 2, \ldots, d.$$

4. Definimos el rango, para cada subperiodo I_a:

$$R_{I_a} = \text{máx}(Y_{k,a}) - \text{mín}(Y_{k,a}),$$

donde $1 \leq k \leq d$.

5. Calculamos la desviación estandar de cada subperiodo I_a:

$$S_{I_a} = \sqrt{\frac{\sum_{k=1}^{d}(Y_{k,a} - m_a)^2}{d}}.$$

6. Normalizamos el rango de cada subperiodo R_{I_a}, dividiendo por la desviación estandar S_{I_a}, y así calculamos el rango rescalado para cada I_a R_{I_a}/S_{I_a}. Así tenemos A subperiodos de tamaño d, y el rango rescalado es definido como el promedio de los rangos de cada subperiodo de tamaño d:

$$(R/S)_d = \frac{1}{A}\sum_{a=1}^{A}\frac{R_{I_a}}{S_{I_a}}.$$

7. La longitud d se incrementa al siguiente divisior de N. Para cada d, los subperiodos comprenden el comienzo y el final de cada subserie. Luego se repiten los pasos 1 al 6, hasta que $d = \frac{N}{2} = \frac{M-1}{2}$. Y con esto estamos listos para realizar el calculo estimado de H, mediante la ecuación (3.4), usando una regresion lineal, en donde la pendiente de esta ecuación es el índice de Hurst, y el intercepto es el valor de la constante $\log c$

NOTA: Cuando las muestras son pequeñas, se pueden producir resultados inestables usando valores pequeños para d, por lo que tomaremos valores de $d \geq 10$.

3.3.1. El índice de Hurst para $M(x)$

Según la metodología anterior hemos calculado el índice de Hurst, a la función acumulativa de Moebius $M(x)$, y al aplicar el algoritmo obtuvimos los siguientes resultados. La primera fila corresponde al número de datos que empleamos, y la segunda es el valor del índice de Hurst obtenido:

Datos	10000	30000	40000	50000	100000	200000	400000
H	0,5398	0,53688	0,52731	0,542194	0,53854	0,52571	0,51838

Debemos observar que la estimación del índice de Hurst va a ser más precisa en cuanto tomemos más datos, pero según el algoritmo descrito anteriormente vemos que los valores que más aportan son aquellas longitudes de la serie que tengan muchos divisores, por eso es mejor darle importancia a valores como 40.000 que a 30.000, pues este primero tiene más divisores y aporta más a la estimación, vemos que para el último valor hallado 400000, (que tiene un buen número de divisores), el valor del índice de Hurst fue de aproximadamente 0.518, con tendencia a bajar. Este fue calculado en un tiempo de 14 minutos, con un computador de triple núcleo de 2.53GHz y 2Gb de Ram. Esta estimación nos da una vez más evidencia, en este caso númerica de lo que podría ser la veracidad de la Hipótesis de Riemann!. Para más detalles del algoritmo implementado y programado por el autor ver el apéndice.

CAPÍTULO 4

Aleatoriedad de la función μ de Moebius

Sabemos que la función de Moebius es una sucesión de 1's, 0's y -1's:

$$1, -1, -1, 0, -1, 1, -1, 0, 0, 1, \cdots ,$$

la pregunta es como antes: ¿ es esta sucesión una sucesión aleatoria?

Daremos un breve panorama del las distintas maneras de aleatotiedad de la función de Moebius que dio Peter Sarnak, en su conferencia titulada "Moebius Randomness and Dynamics", en el IAS. En este capítulo sólo queremos mostrar ideas intuitivas de lo que Sarnak quiere dar a entender como aletoriedad de la función μ, por lo que no daremos demostraciones formales. Este capítulo esta basado en sus manuscritos y el video de su conferencia.

Ya vimos en el capítulo 2, mediante el teorema (6), que afirma:

$$\text{Hipótesis de Riemann es equivalente a } \sum_{n\leq x} \mu(n) = O(x^{1/2+\epsilon}) \quad \forall \epsilon > 0$$

y la fórmula:

$$\prod_{p \text{ primo}} \left(1 - \frac{1}{p^s}\right) = \frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} \quad (Re\ s > 1)$$

que muestra la estrecha relación de los ceros de la función zeta, con la función acumulativa $M(x) = \sum_{n\leq x} \mu(n)$.

Por otro lado es sabido que:

$$\text{El T. de los Números Primos es equivalente a que } \sum_{n\leq x} \mu(n) = \sum_{n\leq x} \mu(n) \cdot 1 = o(x)$$

ver [Apo76] pag. 92.

Y también que el T. de los números primos es equivalente a:

$$A(x) = \sum_{n\leq x} \frac{\mu(n)}{n} = o(1)$$

ver [Apo76] pag. 97.

Esto lleva a hacer la siguiente definición que Sarnak [**?**] llama Ley de aleatoriedad de Moebius.

Definición 6. *Se dice que la función aritmética $\mu(n)$ es ortogonal a la función $\xi(n)$ si*

$$\sum_{n\leq x}\mu(n)\xi(n)=o(x)$$

para ξ acotada e independientemente definida de μ

Además de esto queremos un poco de control sobre la función ξ, y por esto nos interesan funciones calculables en tiempo polinomial, es decir las funciones ξ tales que pueden ser calculadas en a lo más una combinación lineal de $(\log n)^k$ pasos, con $k\in\mathbb{N}$.

Como ya vimos las funciones $f(n)=\frac{1}{n}$, y $i(n)=1$ son ortogonales a $\mu(n)$, y aparentemente esto sugiere que todas las funciones "decentes", en el sentido de su tiempo de cálculo, que conocemos son ortogonales a μ, por lo tanto el problema es ahora el siguiente:

Construir $\xi\in\mathcal{P}$ acotada, tal que

$$\frac{1}{x}\sum_{n\leq x}\mu(n)\xi(n)\longrightarrow\alpha\neq 0$$

es decir encontrar una función en $\mathcal{P}$ que no sea ortogonal a μ. El problema en principio no es sencillo y podría no tener solución.

Ahora abordemos un punto de vista dinámico.

Definición 7 (Flujo)**.** *Un flujo F, es un par $F=(X,T)$, en donde X es un espacio topológico compacto, y $T:X\to X$ continua.*

Si existe un $x\in X$, $f\in C[x]$, y $\xi(n)=f(T^nx)_{n\geq 1}$ entonces la sucesión $\xi(n)$ se dice realizada en el flujo (F,T).

De hecho toda sucesión puede ser realizada en algún flujo: veamos por ejemplo si $X=\{0,1\}^{\mathbb{N}}$, y sea $\xi(n)\in\{0,1\}$

$$\begin{aligned} T:\quad & X\longrightarrow X \qquad\qquad f(x)=x_1\\ & (x_1,x_2,\dots)\longmapsto(x_2,x_3\dots)\end{aligned}$$

T, es continua y como f es una proyección también es continua. Y sea $x=(0,\xi)\in X$, entonces tenemos:

$$T(x)=\xi\quad T^n(x)=(\xi(n),\xi(n+1),\dots)\Rightarrow f(T^n(x))=\xi(n)$$

Luego ξ es realizada en el flujo (X,T), así definido.

También existe otra realización de ξ, mediante el flujo dado por:
$F_\xi=(X_\xi,T)$, $X_\xi=\overline{\{T^j\xi\}}_{j=1,2\dots}\subseteq X$, en donde T es restringida a esta clausura.

Para medir la complejidad de un flujo F, la medida más adecuada es la entropía topológica $h(F)$, introducida por Adler(1965), y redefinida por Bowen de la siguiente manera:

Sea (X, d) un espacio métrico, una función continua $T : X \longmapsto X$, se define para cada n natural una nueva métrica sobre X de la siguiente manera

$$d_n(x, y) = \text{máx}\{d(T^i(x), T^i(y)) : 0 \leq i < n\}$$

Definición 8. *Sea n un número natural, $\epsilon > 0$ y X un espacio métrico y compacto. Un subconjunto de X se dice (n, ϵ) separado con respecto a la función T si para cualquier $x \neq y$ se tiene que $d_n(x, y) > \epsilon$.*

Definición 9. *Si n es un número natural, $\epsilon > 0$ y X es un conjunto métrico y compacto, entonces denotamos por $N(n, \epsilon)$ la cardinalidad más grande de un subconjunto de X, (n, ϵ) separado con respecto a T.*

Definición 10 (entropía (Bowen))**.** *(X, d) espacio métrico compacto. $F = (X, T)$ un flujo entonces la entropía de F se define como el número:*

$$\lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{\log N(n, \epsilon)}{n} = h(F).$$

Debemos observar que como X es un espacio compacto, entonces el número $N(n, \epsilon)$ es finito, y este límite mide el crecimiento exponencial promedio de subconjuntos distinguibles.

Como primer ejemplo podemos considerar el flujo definido por T como una isometría y un espacio métrico (X, d), para el cual se tiene que $d_n = d$ pues la isometría conserva distancias, y así $N(n, \epsilon) = N(1, \epsilon)$, es decir, $N(n, \epsilon)$ no depende del valor de n, por lo tanto

$$h(F) = \lim_{\epsilon \to 0} \frac{\log N(1, \epsilon)}{n} = 0$$

Como otro ejemplo consideremos el flujo defindo por la función continua $T(x) = 2x$ y el espacio $X = [1, 2]$ con la métrica euclideana. Entonces $T^n(x) = 2^n x$. Calculemos $N(n, \epsilon)$

$$|T^n(x) - T^n(y)| = 2^n |x - y| > \epsilon \Rightarrow |x - y| > \frac{\epsilon}{2^n}$$

y el número aproximado de subconjuntos de $[1, 2]$ que cumplen esta condición es $\frac{2^n}{\epsilon}$, por lo que $N(n, \epsilon) > \frac{2^n}{\epsilon}$. Por lo cual obetenmos

$$h(F) \geq \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{\log \frac{2^n}{\epsilon}}{n} = \lim_{\epsilon \to 0} \limsup_{n \to 0} \frac{n \log 2}{n} - \frac{\epsilon}{n} = \log 2.$$

Mediante la entropía podemos medir que tan irregular es un flujo.

Definición 11. *F es determinístico si $h(F) = 0$. La sucesión $\xi(n)$ es determinística si esta es realizada en algún flujo determinístico.*

Después de esta definición Sarnak da su primier teorema:

Teorema 7. *$\mu(n)$ no es determinística.*

Vamos a dar un bosquejo de lo que Sarnak quisó mostrar. Sean

$$X = \{-1, 0, 1\}^{\mathbb{N}},\ T \text{ como en el ejemplo anterior },\ \omega = (\mu(1), \mu(2), \dots) \in X,$$

$$X_M = \overline{\{T^j(\omega)\}}_{j=0}^{\infty} \subseteq X$$

$$M = (X_M, T_M) \text{ El flujo de Moebius}$$

en donde T_M la definimos mediante T restringida a X_M.

Consideremos ahora el siguiente flujo:

$$\eta = (\mu^2(1), \mu^2(2), \dots) \in Y = \{0, 1\}^{\mathbb{N}}$$

$$Y_S = \overline{\{T^j(\eta)\}}_{j=0}^{\infty}$$

$$S = (Y_S, T_S)$$

El flujo S lo denominanos el flujo libre de cuadrados, pues las componentes de η son ceros y unos, dependiendo de si el valor de la posición n, es o no libre de cuadrados y el punto importante es que S factoriza al flujo M en el siguiente sentido. Definamos la función π como sigue

$$\begin{aligned} \pi :& X_M \to Y_S \\ & (x_1, x_2, \dots) \to (x_1^2, x_2^2, \dots) \\ & \pi(\omega) = \eta \end{aligned}$$

Una observación importante es que π es sobreyectiva

$$\begin{tikzcd} X_M \arrow[r, "T_M"] \arrow[d, "\pi"'] & X_M \arrow[d, "\pi"] \\ Y_S \arrow[r, "T_S"'] & Y_S \end{tikzcd}$$

y obtenemos la factorización deseada pues tenemos una nueva función através de Y_S, así:

$$R : X_M \longmapsto X_M \quad R(x) = \pi^{-1}(T_S(\pi(x)))$$

Vamos a calcular la entropía del flujo S, bajo la suposición de que η es una sucesión aleatoria. Para esto dotemos a S de la métrica de l^∞ pues los elementos de Y_S, son sucesiones acotadas, $||x|| = \sup_{n\in\mathbb{N}} |x_n|$, y observemos que los elementos de Y_S, son simplemente colas del elemento η, y al calular la iteración T^i en este elemento el resultado es otra cola de la sucesión η, también vemos que $||T^i(x) - T^i(y)|| > \epsilon$ implica que en algún punto k, $x_k \neq y_k$, con $x, y \in Y_S$.

Tenemos que el número esperado de unos en n posiciones de un elemento de Y_S es $n\frac{6}{\pi^2}$, además como η es una sucesión infinita y no puede ser constante a partir de algún punto, pues si $\mu^2(t) = 1$, entonces $\mu^2(t^2) = 0$, y si $\mu^2(t) = 0$ entonces $\mu^2(t+1) = 1$. Entonces iterando las veces que sea necesario, podemos encontrar para cada $x \in Y_S$, y por cada uno en las primeras n posiciones de x, dos sucesiones de Y_S, las cuales difieren en la posición en

donde se encontraba el uno de x, resaltamos de nuevo que esto es posible, pues podemos iterar las veces que sea necesario hasta encontar un uno o un cero en una posición dada. Luego $N(n,\epsilon)$ debe ser a mayor o igual a $2^{n\frac{6}{\pi^2}}$ y así

$$h(S) \geq \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{\log 2^{n\frac{6}{\pi^2}}}{n} = \log 2\frac{6}{\pi^2}.$$

que coincide con el valor que Sarnak da para esta entropía, y como S es una factorización de M obtendríamos entonces que $h(M) \geq h(S) > 6/\pi^2$, luego $h(M) > 0$. Claro estamos suponiendo que η es una sucesión aleatoria, algo que puede no ser cierto, pero es una aproximación a las afrimaciones de Sarnak, para concluir que μ no es determinística.

Retomando el punto vista citado al comienzo del capitulo tenemos la siguiente definición.

Definición 12. *$\mu(n)$ es disyunta del flujo F si:*

$$\sum_{n\leq N} \mu(n)\xi(n) = o(N)$$

Para cada $\xi(n)$ realizada en F.

La principal conjetura de Sarnak es la siguiente, ya la habiamos nombrado atrás indirectamente y él la llama, *ley de aleatoriedad de Moebius*:

Ley de aleatoriedad de Moebius: μ es disyunta de cualquier Flujo determinístico, es decir que μ es orotogonal a cualquier ξ determinística.

Y el da sus razones para creer esto, descritas a continuación:

Conjetura de Chowla: para $0 < a_1 < a_2 < \cdots < a_t$

$$\sum_{n\leq N} \mu(n+a_1)\mu(n+a_2)\cdots\mu(n+a_t) = o(N)$$

Teorema 8. *Conjetura de Chowla implica la Ley de aleatoriedad de Moebius.*

Esta conjetura propone algo similar a que la función uno es ortogonal a μ, pero relacionando valores de μ. Según Sarnak la prueba es puramente combinatoria y es verdad para cualquier sucesión sin correlación $\xi(n)$.

Observemos los siguientes casos particulares de la ley de aleatoriedad de Moebius:

1. F es un punto $\Leftrightarrow$ T. de los números primos
2. F finito $\Leftrightarrow$ T. de Dirichlet

A pesar de estos casos son pocos, nos sugieren la posible validez de la ley de aletoriedad de Moebius, y la imposibilidad de encontrar una función determinística ξ tal que

$$\frac{1}{x}\sum_{n\leq x} \mu(n)\xi(n) \longrightarrow \alpha \neq 0,$$

lo cual sugiere en este sentido que μ no es determinística.

Hemos descrito dos posibles definiciones alternas, y que además son “atípicas", por decirlo de alguna manera, en las que Sarnak define lo que podría ser la aleatoriedad de la función μ de Moebius, hemos visto además sus razones para creer que según ambas definiciones la función μ es aleatoria, esto es lo que hace interesante la charla de Sarnak, puesto que lo que hay detrás de la aleatoriedad de la función μ de Moebius es la Hipótesis de Riemann.

APÉNDICE A

Algoritmos

Para el cálculo del índice de Hurst que nos interesa, hemos usado dos algoritmos, uno para calcular los valores de la función de Moebius, y otro para hallar los valores de que se necesitan en la regresión lineal del proceso de Hurst, todos realizados en $c++$, por el autor, finalmente el valor de H, es calculado en scilab en base de estos datos:

A.1. Algoritmo para Moebius

```
#include <math.h>
#include <stdio.h>
#include <stdlib.h>

//--------------------------------
int primo(int n);
int fprimos(int n,int *fp,int &nf);
int comparar(int n, int *x);
int moebius(int n);
int AcumulativaM(int x);
//=============================
int main()
{
  int n;
  int *factp;
  int nf;
  int i,k;

  printf("\n n = ");
  scanf("%d", &n);

  if(n<2){
    printf(" \n Numero n = %d inadecuado\n\n",n);
```

```
    return 0;
  }
  factp = new int[55];
  if( factp == NULL){
    printf(" \n Memoria Insuficiente\n");
    return 0;
  }
  k= fprimos(n,factp,nf);
  if(k==0) return 0;

  printf("\n Los factores primos de %d son :\n\n ",n);

  for(i=0;i<nf; i++){
    printf(" %d  ",factp[i]);
  }
  delete factp;
  printf("\n\n La funcion de moebius calculada en %d es:\n ",n);
  printf("\n  M(%d) = %d \n\n", n, moebius(n));

  printf("\n El valor de la función acumulativa hasta n es :\n");
  printf("\n   %d\n\n", AcumulativaM(n));
  return 0;
}
//--------------------------
int primo(int n)
{
  // Devuelve el primer divisor de n
  // Devuelve n si n es primo
  // Devuelve -1 si n no es mayor que 2

  int i;

  if( n< 2 ){
    printf("Numero %d inadeacuado ",n);
    return -1;
  }
  for( i=2; i*i <= n; i++ ) if ( n%i == 0 ) return i;
  return n;
}
//----------------------------
int fprimos(int n,int *fp,int &nf)
{
  // Devuelve 1 si no hay problemas
  // Devuelve 0 si hay problemas

    int d;
```

```
    if( n<2){
      printf(" Factores Primos: %d inadecuado\n\n  ",n);
      return 0;
    }
    nf =0;
    while(n>1){
      d=primo(n);
      fp[nf] = d;
      nf++;
      n = n/d;
    };
    return 1;
 }
//-------------------------------------
int comparar(int n, int *x)
{
  // Devuelve 0 si algun elemento del vector x se repite
  // Devuelve 1 si todos sus elementos son distintos

    int i;
    int j;

    for(i=0; i<n; i++){
      for(j=i+1; j<= n; j++){
if(x[i]==x[j]) return 0;
      }
    }
    return 1;
}
//----------------------
int moebius(int n)
{
  // Devuelve 1 si n tiene un numero par
  // de factores primos distintos
  // Devuelve -1 si posee un numero impar
  // de factores primos distintos
  // Devuelve 0 en otro caso

  int *fp;
  int nf;

  if (n == 1) return 1;

  fp = new int[50];
  if( fp == NULL){
    printf(" \n Memoria Insuficiente\n");
    return 0;
```

```
  }
  fprimos(n,fp,nf);
  if( comparar(nf, fp) == 0) return 0;
  if( nf%2 == 0) return 1;
  delete fp;
  return -1;
}
//-------------------------------------------
int AcumulativaM(int x)
{
  // Calcula la funcion acumulativa de moebius

  int i,d;

  FILE *archivo;

  archivo=fopen("moebius.txt","w");
  if( archivo == NULL){
    printf(" \n Nombre de archivo para resultados incorrecto.\n");
    return 0;
   }

  d=0;
  for(i=1; i<=x; i++){
    d += moebius(i);
    fprintf(archivo," %d \n",d+300);
  }
  fclose(archivo);
  return d;
}
```

A.2. Algoritmo de Hurst

Recordemos que el siguiente algoritmo solo va a calcular los datos necesarios para la regresión lineal.

```
#include <stdio.h>
#include <math.h>
#include <stdlib.h>

void transpuesta( double *a, double *at, int n, int m );
int leedatos(FILE *arch, int *x,int n);
int divisores(int *d, int n);
double max(double *x, int n);
double min(double *x, int n);
double desv(double *x, int n, int sx);
```

```
double rango(double *x, int d, int sx);
double hurst(double *x, int n, int *d, int m, double *rt);
//double regresion(int *x, int n);
double datregresion(double *x, int );
//==========================================
int main()
{
  int n,k,i;
  int *x;
  double *y;
  FILE *archdat;
  FILE *xnuevo;

  printf("\n\n Digite el numero de datos \n n = ");
  scanf("%d", &n);

  x = new int[n];
  if( x == NULL){
    printf(" \n Memoria Insuficiente\n");
    return 0;
  }

  archdat = fopen("moebius.txt","r");
  if (archdat == NULL){
    printf("\n\n Archivo inexistente. \n\n");
    return 0;
  }
  k= leedatos(archdat,x,n);

  xnuevo = fopen("loga.txt","w");

  y = new double[n-1];
  if( y == NULL){
    printf(" \n Memoria Insuficiente\n");
    return 0;
  }
  for(i=0; i<n-1; i++){
    y[i] = log(double(x[i+1])/double(x[i]));
  }

  delete x;
  for(i=0; i<n-1;i++){
    fprintf(xnuevo," %lf\n ",y[i]);
  }
  printf("\n\n lee = %d\n",k);

  datregresion(y,n-1);
```

```
  return 0;

}
//---------------------------------------------
void transpuesta( double *a, double *at, int n, int m )
{

  // Halla la transpuesta de una matriz mxn

  int i, j, im;

  for( i=0; i<n; i++){
    im = i*m;
    for( j=0; j<m; j++){
      at[im+j] = a[j*n+i];
    }
  }
}
//--------------------------------------------
int leedatos(FILE *arch, int *x,int n)
{
  // Devuelve 1 si se efectuo la lectura normalmente,
  //          0 si se envontro el fin del archivo antes de acabar la lectura
  //         -1 si hubo un error durante la lectura con fscanf

  int i, res;

  for( i=0; i<n ; i++ ){
    res = fscanf(arch,"%d",&x[i]);
    if( feof(arch) ) return 0;
    if(res <= 0) return -1;
  }
  return 1;
}
//-------------------------------------
int divisores(int *d, int n)
{
  //Halla los divisores de n mayores a 10

  int i;
  int k;

  k = 0;
  for(i=10;i<= n/2; i++){
    if(n%i == 0){
      d[k] = i;
```

```
      k++;
    }
  }

  return k;
}
//----------------------------------------
double max(double *x, int n)
{
  // Halla el valor maximo de un vector

  int i;
  double vmax;

  vmax = x[0];
  for(i=0; i< n; i++){
    if(x[i]>vmax) vmax = x[i];
  }
  return vmax;
}
//-------------------------------------
double min(double *x, int n)
{
  // Halla el valor minimo de un vector

  int i;
  double vmin;

  vmin = x[0];
  for(i=0; i< n; i++){
    if(x[i]<vmin) vmin = x[i];
  }
  return vmin;
}
//---------------------------------------
double desv(double *x, int n, int sx)
{
  //Calcula la desviacion estandard

  int i;
  double s,m;

  m=0.0;
  for(i=0; i<n; i++) m += x[i+sx];
  m=m/n;

  s=0.0;
```

```
  for(i=0; i<n; i++){
    s += x[i+sx]*x[i+sx];
  }
  return sqrt(s/n-m*m);
}
//-------------------------------------
double rango(double *x, int d, int sx)
{
  // Calcula el rango - rescalado en un subperiodo
  // de tamaño d

  int i,j,k;
  double p;
  double t;
  double y[d];
  double r,rs;


  p=0.0;
  for(i=sx; i<d+sx; i++){
    p +=x[i];
  }
  p=p/d;

  for(j=0; j<d; j++){
    t=0;
    for(k=sx; k <= sx+j; k++) t += x[k]-p;
    y[j] = t;
    //printf("\n y(%d) = %lf",j,y[j]);
  }
  r = max(y,d)-min(y,d);
  rs= r/desv(x,d,sx);

  return rs;
}
//------------------------------
double hurst(double *x, int n, int *d, int m, double *rt)
{
  // Halla el indice de hurst
  // Usa la funcion rango en cada subperiodo

  int i,j;

  double *r;
  double t;

  for(j=0; j< m; j++){
```

```
    t = 0.0;

    r = new double[n/d[j]];
    if( r == NULL){
      printf(" \n Memoria Insuficiente\n");
      return 0;
    }

    for(i=0; i<n/d[j]; i++){
      r[i]= rango(x,d[j],i*d[j]);
      t += r[i];
    }
    delete r;
    rt[j] = log(t/(n/d[j]));
  }
  return 0;
}
//-------------------------------------
double datregresion(double *x, int n)
{

  FILE *archivo;
  int i,m;
  int *d;
  double *dd,*rt;


  d = new int[n/3];
  if( d == NULL){
    printf(" \n Memoria Insuficiente\n");
    return 0;
  }

  m = divisores(d,n);

  rt = new double[m];
  if( rt == NULL){
    printf(" \n Memoria Insuficiente\n");
    return 0;
  }

  hurst(x,n,d,m,rt);

  dd = new double[m];
  if( dd == NULL){
    printf(" \n Memoria Insuficiente\n");
    return 0;
```

```
  }

  for(i=0; i<m; i++) dd[i] = log(d[i]);
  delete d;

  printf("\n\n hola m = %d \n\n",m);

  archivo=fopen("datosreg.txt","w");
  if( archivo == NULL){
    printf(" \n Nombre de archivo para resultados incorrecto.\n");
    return 0;
   }
  for(i=0; i<m; i++) fprintf(archivo," %lf ",dd[i]);
  fprintf(archivo,"\n\n");
 for(i=0; i<m; i++) fprintf(archivo," %lf ",rt[i]);
  fclose(archivo);
  return 0;
}
```

Al llamar la función de Moebius, nos pedira que entremos el valor de n, para el cual digitaremos el valor que deseamos más una unidad, pues el agoritmo de Hurst convierte la serie original en una serie de $n-1$ elementos, al igual que en el algoritmo de Hurst, también debe introducirse $n+1$ en el algoritmo para Moebius, después debemos ir al archivo datosreg.txt en el cual están los valores listos para hacer la regresión lineal, la cual fue hecha en scilab 5.3.2, mediante la función $regress(x, y)$, en la cual previamente debemos definir el valor de x, y, de la manera usual, es decir $x = [\quad]$, $y = [\quad]$, dentro de los corchetes ponemos los datos y depues damos la orden de la regresión $regress(x, y)$, el segundo valor que nos genera es el valor de la pendiente de la recta que aproxima los datos, es decir el valor de H.

Conclusiones

- La función zeta de Riemann es una poderosa herramienta en el estudio de los números primos. Está demostrado que la menor cota para el error en el teorema de los números primos se obtiene asumiendo la hipótesis de Riemann.
- Uno de los posibles caminos para resolver la Hipótesis de Riemann es estudiar el comportamiento de la función acumulativa de Moebius, debido a la equivalencia conocida y descrita en el trabajo entre el crecimiento de dicha función y la hipótesis de Riemann.
- Asumiendo la independencia en los valores de la función μ el argumento de Denjoy concluye que con probabilidad cercana a uno la Hipótesis de Riemann es cierta. Obtenemos el mismo resultado incluyendo algunas correlaciones obvias que Denjoy no considera, refinando así su resultado.
- Hay razones de más para averiguar el comportamiento de la función de Moebius y en particular de $M(x)$, por lo que es válido estudiar dicha función desde otras ramas de la matemática, como los sistemas dinámicos. Para ello usamos las diversas maneras de aleatoriedad descritas por Sarnak.

Bibliografía

[Ahl79] L.V Ahlfors. *Complex Analysis.* McGraw-Hill, third edition, 1979.

[AK92] S.M Voronin A.A. Karatsuba. *The Riemann zeta-function.* De Gruyter exposition in Mathematics, 1992.

[Apo76] T.M. Apostol. *Introduction to Analytic Number Theory.* Springer - Verlag, New York, 1976.

[Edw74] H.M. Edwards. *Riemann's Zeta Function.* Dover Publications, New York, 1974.

[Fra08] Biagini. Francesca. *Stochastic calculus for fractional Brownian motion and applications.* Springer, 2008.

[Jam03] G.J.O Jameson. *The prime Number Theory.* Cambridge Univ. Press, 2003.

[Pet94] E. Peters. *Fractal Market Analysis.* John Wiley and Sons, 1994.

[Ros96] S. Ross. *Stochastic Processes.* John Wiley and Sons, 1996.

[Sar10] P. Sarnak. Moebius randomness and dynamics. *IAS*, June 2010.

[Ívi85] A. Ívic. *The Riemann Zeta - Function.* Dover Publications, New York, 1985.

[Wal82] P. Walters. *An Introduction to Ergodic Theory.* Springer, 1982.